ADVANCED PLANNING, CONTROL, AND SIGNAL PROCESSING METHODS AND APPLICATIONS IN ROBOTIC SYSTEMS

EDITED BY: Zhan Li, Shuai Li, Yangming Li and Yan Wu
PUBLISHED IN: Frontiers in Neurorobotics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-489-3
DOI 10.3389/978-2-88974-489-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ADVANCED PLANNING, CONTROL, AND SIGNAL PROCESSING METHODS AND APPLICATIONS IN ROBOTIC SYSTEMS

Topic Editors:
Zhan Li, University of Electronic Science and Technology of China, China
Shuai Li, Swansea University, United Kingdom
Yangming Li, Rochester Institute of Technology, United States
Yan Wu, Institute for Infocomm Research (A*STAR), Singapore

Citation: Li, Z., Li, S., Li, Y., Wu, Y., eds. (2022). Advanced Planning, Control, and Signal Processing Methods and Applications in Robotic Systems. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-489-3

Table of Contents

frontiers in Neurorobotics

ORIGINAL RESEARCH
published: 29 June 2020
doi: 10.3389/fnbot.2020.00030

An Enhanced Robot Massage System in Smart Homes Using Force Sensing and a Dynamic Movement Primitive

Chunxu Li[1,2], *Ashraf Fahmy*[3,4], *Shaoxiang Li*[2]* *and Johann Sienz*[3]

[1] Centre for Robotics and Neural Systems, University of Plymouth, Plymouth, United Kingdom, [2] Shandong Marine Corrosion and Safety Protection Engineering Research Center, Qingdao University of Science and Technology, Qingdao, China, [3] ASTUTE 2020 in Future Manufacturing Research Institute, College of Engineering, Swansea University, Swansea, United Kingdom, [4] Department of Electrical Power and Machines, Helwan University, Helwan, Egypt

With requirements to improve life quality, smart homes, and healthcare have gradually become a future lifestyle. In particular, service robots with human behavioral sensing for private or personal use in the home have attracted a lot of research attention thanks to their advantages in relieving high labor costs and the fatigue of human assistance. In this paper, a novel force-sensing- and robotic learning algorithm-based teaching interface for robot massaging has been proposed. For the teaching purposes, a human operator physically holds the end-effector of the robot to perform the demonstration. At this stage, the end position data are outputted and sent to be segmented via the Finite Difference (FD) method. A Dynamic Movement Primitive (DMP) is utilized to model and generalize the human-like movements. In order to learn from multiple demonstrations, Dynamic Time Warping (DTW) is used for the preprocessing of the data recorded on the robot platform, and a Gaussian Mixture Model (GMM) is employed for the evaluation of DMP to generate multiple patterns after the completion of the teaching process. After that, a Gaussian Mixture Regression (GMR) algorithm is applied to generate a synthesized trajectory to minimize position errors. Then a hybrid position/force controller is integrated to track the desired trajectory in the task space while considering the safety of human-robot interaction. The validation of our proposed method has been performed and proved by conducting massage tasks on a KUKA LBR iiwa robot platform.

Keywords: hybrid force/position, teaching by demonstration, dynamic motion primitive, dynamic time warping, gaussian mixture regression

Edited by:
Yangming Li,
Rochester Institute of Technology,
United States

Reviewed by:
Hang Su,
Fondazione Politecnico di Milano, Italy
Jihong Zhu,
UMR5506 Laboratoire d'Informatique,
de Robotique et de Microélectronique
de Montpellier (LIRMM), France

****Correspondence:***
Shaoxiang Li
leeshaoxiang@126.com

Received: *19 March 2020*
Accepted: *04 May 2020*
Published: *29 June 2020*

Citation:
Li C, Fahmy A, Li S and Sienz J (2020) An Enhanced Robot Massage System in Smart Homes Using Force Sensing and a Dynamic Movement Primitive. Front. Neurorobot. 14:30. doi: 10.3389/fnbot.2020.00030

1. INTRODUCTION

With the continuous development of technology, many traditional industries have been gradually replaced by high-tech products. Among them, the development of robot technology plays an important role. Robot technology integrates multiple disciplines, such as machinery, information, materials, intelligent control, and biomedicine, and intelligent service robots are intelligent equipment that provide humans with necessary services in an unstructured environment. The tasks that a robot can accomplish are divided into two categories. One is non-contact operation: the robot carries and operates the target in a free space, and the simple position control can be used for another type of contact operation. For tasks such as assembly, grinding, polishing, debarring, etc., simple position control is no longer sufficient. In such cases where the contact force is required, the slight positional deviation of the robot end may cause the contact force to damage the robot and

FIGURE 1 | Conception of robot massaging in smart homes.

the target. Therefore, the contact force control function must be added to the contact robot control system. The traditional robot control system for free space motion cannot meet the requirements of the extended application of robots. The perception and control of the robots of environmental forces are problems that current robot technology needs to solve. The control strategy combined with the end pose and the contact force for a robot is defined as the robotic force/position control. The key technologies of the current robotic force/position control mainly include the following aspects:

1. the robotic impedance control under the condition that the environmental parameter geometry and dynamic model parameters are unknown or changed;
2. the robotic force/position hybrid control under system interference (model error, measurement noise, and external input interference);
3. the robotic collision contact control;
4. the engineering implementation of robotic force/ position control.

Service robots mimic humans in shape and behavior design; for instance, they have hands, feet, heads, and a torso. This helps them to adapt to human life and a work environment. Service robots replace human beings to complete various tasks and expand human capabilities in many aspects. Recently, research into robot massaging in smart homes (shown in **Figure 1**) has attracted widespread attention. Massaging is an important part of people's realization of a healthy lifestyle that can relieve stress and relax the body. By relaxing the muscles and joints, massages improve the flexibility of the body's joints and muscles, thereby reducing muscle pain caused by poor posture. However, for the manual massage tasks, it takes lots of energy and labor cost. Thus, in recent years, robot massaging has become a research hotspot in the field of robots. In 1996, the Mechatronics Research Center of Japan's Sanyo Electric Co., Ltd. designed a mechanical therapy unit and verified the feasibility of developing an intelligent massage robot (Kume et al., 1996). In 2000s, Toyohashi University of Technology and Japan's Gifu Institute of Technology carried out research on humanoid multi-finger massage robots with four fingers and 13 joints (Zhang and Zhang, 2017). Since 2004, Chinese scholars have carried out research on massage robots based on the theory of traditional Chinese massage and systematically discussed the robotic synthesis of various traditional Chinese medical massage techniques (Ma et al., 2005). However, existing massage robots are complicated to operate with limited functions, a large size with bulky equipment, and they are also expensive. Moreover, most of the existing robot massage technology depends on the function of the robot product itself; in other words, these robots cannot be used for operations other than massage, and their tandem structure stiffness is large, the motion inertia is unstable, and the working space and flexibility are limited (Zhang and Zhang, 2017; Field, 2018). To overcome the abovementioned issues, this paper develops a teaching-by-demonstration-based interface using a hybrid position/force control strategy with adjustable stiffness, which can be implemented onto a general robot manipulator with high accuracy, taught by a human operator.

Goradia et al. (2002) studied robotic forces/positions under different surfaces (i.e., unknown to the environment) to better facilitate the polishing or grinding processes. Scherillo et al. (2003) studied the force/position control of multi-finger grasping, Nguyen et al. (2000) used the sliding mode variable structure to control the force and the position of the robot, Ha et al. (2000) and used the impedance control method to control the force/position of the robot. In Fanaei and Farrokhi (2006), authors used a robust adaptive method to control the force and the position of the robot. Research on the force/position intelligent control of the robot is still mainly theoretical, the technical realization is at the stage of exploration, and there is still a certain distance to go in terms of promotion and practicality.

Movement is necessary to directly produce skill effects. The motion model generates a continuous robot space state representation skill offline or online with clear meanings related to the physical system, such as position, attitude, and contact force. Generally, the motion is assumed to be nonlinear. The motion model mainly includes two categories: trajectory encoding modeling and dynamic system modeling. Trajectory encoding is a compact mathematical model that represents the shape, constraints, and other information of the trajectory. When the human operators teach and store one or more specific fixed trajectories in the robot, then the robot can accurately reproduce the trajectory when the motion skills are executed. Calinon and Billard (2007) used trajectory encoding with a multivariate Gaussian mixture model (GMM), which expresses the trajectory in a fixed coordinate system. However, considering reality, the trajectory often needs to be expressed in different reference coordinate systems, so the transformation relationship of the reference coordinate system may also change. Thus, in Silvério et al. (2015), they proposed a task-parameterized Gaussian Mixture Model (TP-GMM) wherein the origin and rotation transformation matrix of the coordinate system are used as the task parameters in the model, which allows the observation and reproduction of the motion trajectory in different reference coordinate systems. However, the trajectory encoding modeling method only explicitly generates a fixed trajectory. When the robot is disturbed and deviates from the trajectory, the trajectory

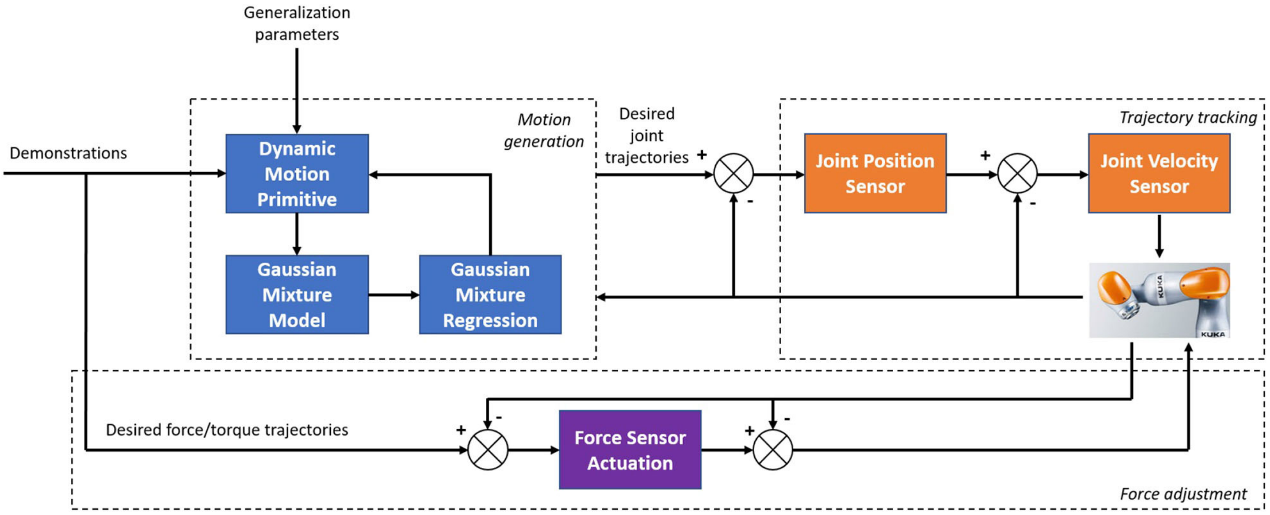


FIGURE 2 | Diagram of the proposed control architecture.

cannot be adjusted in real time, and a new trajectory needs to be regenerated.

The dynamic system is automatically evolved according to certain rules. Compared with the trajectory encoding, there are two main differences: the dynamic system does not explicitly depend on the time variable–only the relationship between the spatial state and its time derivative; the dynamic system can be online. The trajectory generated by such a dynamic system obtains the online adaptive ability for disturbance. DMP was proposed by Schaal (2003) in 2002 as a dynamic system model that can generate trajectories of arbitrary shapes where its basic idea is to drive a transform system with a canonical system. Nanayakkara et al. (2004) proposed a Mixture of motor primitives (MoMP) using a Gaiting network to calculate the weight of each DMP in the current state. Ideally, the Gaiting network should only choose one DMP implementation. However, the reality is more complicated; in order to achieve a good generalized performance, MoMP weights the output of all DMPs based on the weight of the output of the gated network to obtain the final output. In Matsubara et al. (2011), the coordination matrix is used to represent the coupling relationship between multiple DMPs corresponding to multiple degrees of freedom, and the iterative dimensionality reduction method is used to reduce the unnecessary degrees of freedom in the cooperation matrix, which is beneficial to enhancing the learning efficiently. DMP mainly depends on the target state and weight coefficient. The former is determined by the environment or setting. The latter can learn from the teaching trajectory based on the linear weighted regression (LWR) method. However, the LWR can only learn the DMP model parameters from a single demonstration. This paper mainly presents two aspects of research: the theory of hybrid force/position control with direct human–robot interaction and the experimental studies on a real robotic platform. The motion planning is performed in 3D task space, where the GMM is employed to evaluate the DMP to learn from multiple demonstrations, and GMR is used for the reproduction of the generalized trajectory with a smaller error. For the force input aspects, a hybrid force/position controller is introduced to ensure the safety of direct human–robot interaction. An overview of the presented control architecture is shown in **Figure 2**. The contributions of this paper are summarized.

1. This paper employs a teaching interface to perform the robot massaging under the demonstrations of the human operator, and the experimental studies show that, after the teaching process, the robot can generate an even smoother trajectory that strictly adheres to what is requested to be followed.
2. The application of the hybrid force/position control scheme takes care of the safety issues and achieves massage services performing without knowing subject profiles.
3. The generalization functions of our proposed method supplies a more flexible and convenient option with only once teaching for multiple tasks to the carers and patients, which promotes the user experiences.

2. DATA PREPROCESSING

2.1. Motion Skills Segmentation

FD approximates the derivative by the limitable variance and finds an estimated solution for the differential equation (Chelikowsky et al., 1994). The differential form, especially, is to substitute the differential with a finite difference and the derivative with a finite differential quotient so as to roughly change the fundamental formula and limit state (usually the differential equation) into the differential equation (Algebraic equation) (Gear, 1988). The solution of the differential equations problem is updated so that the algebraic equations problem is resolved.

Segmentation of skills is usually a complicated and systematic process that requires more time and efforts on its algorithms designing. For such a complex method, there are often difficulties in setting the priori parameters. Considering the situation that the massaging motion is planed on a horizontal plane, and the force is applied into vertical direction of that plane, during the massaging tasks, it is easy to track for the system at which point the position being massaged has been changed. This is because its changing position represents the value varying in z coordinates. Consequently, a simple segmentation method (FD) is employed in this paper. In view of y_i=$f(z_i)$ relies upon the z_i variables, where z_i denotes the coordinate values in the vertical direction at the i_{th} time series, and y_i represents the corresponding spatial sequence of the whole dataset. If z_i is changed into z_{i+1}, the corresponding changes in the whole dataset is $df(z_i) = f(z_{i+1}) - f(z_i)$ and d is the differentiation operator. Difference has a differential-like arithmetic value. This displays the following equation (Li et al., 2018):

$$f'(z_i) = df(z_i) = y_i' \approx \frac{y_{i+1} - y_i}{z_{i+1} - z_i} \tag{1}$$

where one significant massaging aspect is that the endpoint would be lifted once every single massaging task is done. The "z" alignment values of the experimental data are thus viewed as the segmentation reference. We have received, pursuing the FD,

$$\xi(y_i^{'}) = sign(\left|y_i^{'}\right| - \theta) \tag{2}$$

Where ξ is the gaping variable, abd θ is a constant. We could modify the segmented characters, such as, θ, here θ=0.5, by giving different values of θ. Sign is the Signum function, and, for each component of ξ, the formulation could be described as following:

$$sign(\xi) = \begin{cases} -1 & if\ \xi < 0, \\ 0 & if\ \xi = 0, \\ 1 & if\ \xi > 0. \end{cases} \tag{3}$$

Up to now, the segmented motion trajectories $sign(\xi)$ for massaging have been outputted to different local text files for the use of GMM and DMP generalization, which correspond to the "z" coordinate information in the robot space, suddenly and sharply rising, which implies every time the massaging for one position was done. The flowchart of the segmentation is shown in **Figure 3**.

2.2. Alignment of Time Series

The DTW employed in this paper aligned the outputted trajectories curves in the form of $W = \{w_1, w_2, \cdots, w_p, \cdots, w(p)\}$ in which $w(p) = (i_p, j_p)$ denotes the match variable in Petitjean et al. (2014). In this situation, the skewed characteristic W is required to reduce the discrepancy between the pending trajectory and the reference trajectory. The formula is therefore described as

$$D = min \sum_{k=1}^{K} d[w(p)] \tag{4}$$

where $d[w(p)] = d[T_i(p), R_j(p)]$ represents the measured distance from the $i(p)$th featured point of the pending trajectory to the $j(p)$th featured point of the reference trajectory, which is normally described by a square measure specified as follows:

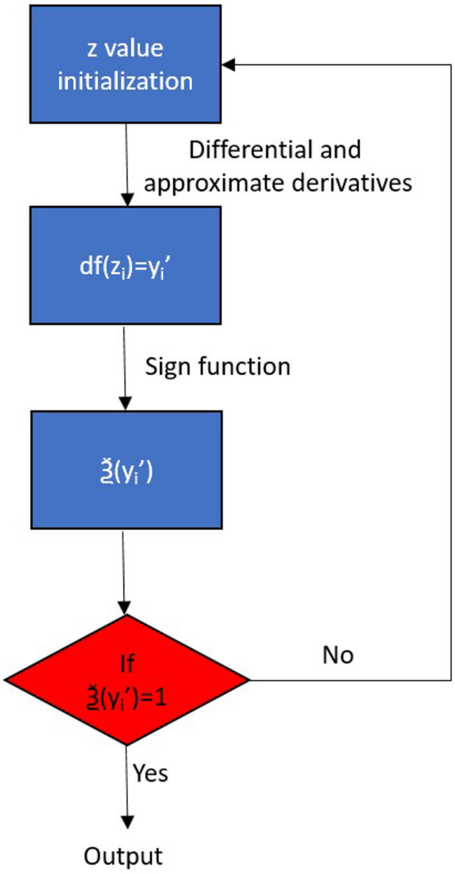


FIGURE 3 | The flowchart of the segmentation.

$$d[w(p)] = [T_i(p) - R_j(p)]^2 \tag{5}$$

We need to construct a matrix grid of $m \times n$ in order to coordinate the two specimens; the matrix element (I, j) represents the range $d(T_i, R_j)$ between the two points T_i and R_j, and each matrix element (i, j) represents the alignment of points R_j and R_j. DTW aims to find a direction which reflects the matched points for both samples to be determined via several grid points. First we illustrate, as $D_{Acc}(i, j)$, the total minimum range between the two trajectories, then we consider (Senin, 2008)

$$D_{Acc}(i, j) = d(T_i, R_j) + min_{(q_i, q_j)}[D_{Acc}(q_i, q_j)] \tag{6}$$

where (q_i, q_j) belongs to the set of points between $(1, 1)$ and (i, j) within a certain direction. From the above components it can be seen that the average cumulative distance of the (i, j) element is linked not only to the regional distance $d(T_i, R_j)$ of the own values T_i, R_j but also to the total cumulative distance earlier than this stage in the coordinate system.

We thus assume that $(i, j-1)$, $(i-1, j)$, and $(i-1, j-1)$ for any point $c(p) = (i, j)$ within the coordinate system may enter the preceding point of $c(p)$, so the choice of the preceding variable also needs to agree with the above three factors. We may measure the corresponding DTW range between the test pattern vector and the comparison model vector, as shown below, according to the formula.

$$D' = D_{Acc}(L_1, L_2). \tag{7}$$

3. TRAJECTORY GENERATION

DMP is an innovative model (Schaal, 2006) dynamic system learned from biological research that learns from motor primitives. The definition of dynamic primitive can be separated into two groups. One is to use different formulas based on dynamic system to represent the state, and the other is to produce the track by interpolation through the interpolation points (Li et al., 2017). DMP is made up of two parts: the transformed model r and the canonical framework h. The equation is shown as follows:

$$\dot{s} = h(s) \tag{8}$$

$$\dot{t} = r(t, s, w) \tag{9}$$

where t and s are the transformed process states and the canonical function, and the canonical system output variable h is referred to as w.

The canonical model is defined by an exponential differential equation, which is given:

$$\tau\dot{s} = -\alpha_f s \tag{10}$$

where s is a step function varying from 0 to 1, $\tau > 0$, and α_f is a temporal scaling variable and a balanced component.

The transformed model consists of two nonlinear term sections and a Cartesian space spring damping mechanism, the formulas are defined as (Schaal, 2006):

$$\tau\dot{v} = k(g - p) - cv + X(g - p_0) \tag{11}$$

$$\tau\dot{p} = v \tag{12}$$

where p_0 is the starting position, $p \in R$ is the Cartesian position, $v \in R$ is the end-effector velocity of the robot, g is the goal, k is the spring variable, and c is the damping factor. X is a conversion method of dynamic nonlinear structures capable of transforming the outcomes of the canonical model found in the following formula:

$$X = \sum_{i=1}^{N} w_i l(s) \tag{13}$$

Where the GMM number is N, $w_i \in R$ is the weights, and l is the uniform radial variable value that can be supplied as follows:

$$l(s) = \frac{exp(-h_i(s - c_i)^2)}{\sum_{m=1}^{N} exp(-h_m(s - c_m)^2)} \tag{14}$$

where $c_i > 0$ are the centers, and $h_i > 0$ are the widths of the functions of the Gaussian foundation. N is the number of functions in Gaussian.

In addition, we can use the weight variable to produce motions by specifying the starting point of the canonical process (s=0) X_0 and aim g, which is the canonical system integration. The theory of DMP is to measure the nonlinear transition feature X by observing the presenter's movements. Nonetheless, there are drawbacks in developing a multi-demonstration conversion model, which is why the GMM is used to solve the above problems.

GMM's parameter estimation is the method by which the design parameters are obtained under certain conditions. In fact, it is the process of knowing the parameters of the model, namely, the method of resolving $\lambda = \{\mu_i, \sum_i, \omega_i\}$ to bring the GMM sequence of observation (Tong and Huang, 2008). The most commonly employed parameter estimation is the total probability approximation process. The basic idea is to consider the system parameter λ when the peak likelihood of GMM is obtained by providing the observation sequence X obtained by DMP from the previous chapter, then λ is the model. The optimum function, λ, defines to the maximum extent practicable the distribution of the observed string.

The end goal of the total probability calculation after providing the training information is to seek a template variable that maximizes the GMM's likelihood. For a training vector series of $X = \{x_1, x_2 \ldots x_D\}$ of duration D, it is possible to describe the likelihood of GMM as

$$P(X \mid \lambda) = \prod_{t=1}^{D} P(X_t \mid \lambda) \tag{15}$$

The parameter λ is then continuously updated until a set of parameters λ is found to maximize $P(X \mid \lambda)$:

$$\hat{\lambda} = \underset{\lambda}{argmax} P(X \mid \lambda) \tag{16}$$

For the convenience of analysis, $P(X \mid \lambda)$ usually takes its log likelihood, giving us

$$log(P(X \mid \lambda)) = \sum_{t=1}^{D} logP(X_t \mid \lambda) \tag{17}$$

The Expectation Maximization (EM) algorithm can be used for parameter estimation provided that there is a relatively complex nonlinear interaction between the probability function and the template parameters, and the peak value cannot be determined according to the standard probability estimation process. The EM algorithm is in essence an iterative method for calculating the probability model's maximum likelihood. The process of each iteration is to estimate the unknown data distribution based on the parameters that have been acquired and then calculate the new model parameters under the maximum likelihood condition. Let the initial model parameter be λ, which satisfies

$$P(X \mid \lambda') \geqslant P(X \mid \lambda) \tag{18}$$

Next, we calculate the new model variable λ' according to the equation above and then use the λ' parameter as the original parameter for the next iteration. It iteratively iterates until the state of convergence is met. Here we assume a Q function that represents the E phase of the EM process shown below:

$$Q(\lambda, \lambda') = \sum_{i=1}^{M} P(X, i \mid \lambda) log P(X, i \mid \lambda) \tag{19}$$

where i is an elusive and unpredictable secret country. $Q(\lambda, \lambda')$ corresponds to all observable data's log likelihood assumptions. Calculating the maximum value of $Q(\lambda, \lambda')$ increasing give the maximum log likelihood of the observed data, which is the M stage of the EM process. It is possible to obtain replacement formulas (15) and (16) for equations (6):

$$Q(\lambda, \lambda') = \sum_{i=1}^{M} \sum_{t=1}^{T} r_t(i) log \omega_i b_i(x) \tag{20}$$

$$r_t(i) = P(X_t, i \mid \lambda) \tag{21}$$

The approximate values of each variable are then computed according to E and M. Phase E calculates the posterior likelihood of the t_{th} test X_t of the training data according to the Bayesian equation in the i_{th} state; phase M first uses the Q method to extract the three parameters independently and then evaluate the corresponding figures. To re-evaluate the variables, we iteratively perform measures E and M. The loop is halted when the peak value of the likelihood function is reached.

The first step when using the EM method to calculate the GMM parameters is to determine the number of Gaussian components in the GMM, such as system M order and model initial parameter λ (Howlett et al., 2009). Based on the actual situation, such as the sum of training data, it is appropriate to choose the order M of the template. The most widely employed approach for the model's initial variable λ is the K-means algorithm. Currently, the K-means algorithm is the simplest and most effective classification algorithm, commonly used in different models (Nazeer and Sebastian, 2009). The GMM used in this paper chooses the basic parameters using the K-means method. The K-means algorithm partitions the information into K clusters according to the in-cluster number of squares in the category theory. After using the K-means method to cluster the feature vectors, the mean and variance of each group are determined and the percentage of each class 'feature vectors is calculated as the blending weight (Tatiraju and Mehta, 2008). The average, variance, and combined weight can be then collected as the predicted values. Finally, by applying the GMR, the same theory as our previous work (Li et al., 2018), the reconstructed motion trajectories can be then obtained.

4. HYBRID FORCE/POSITION CONTROL

The robot kinematics model of n DOF robot is presented in the subsequent form

$$x(t) = \Psi(\Theta) \tag{22}$$

where x(t) $\in$ R^n represents the position and direction vector, and $\Theta \in$ R^n represents the joint angle vector. The inverse kinematics are

$$\Theta(t) = \Psi^{-1}(x). \tag{23}$$

The derivative of (23) can thus be rewritten:

$$\dot{x}(t) = J(\Theta)\dot{\Theta} \tag{24}$$

where $J(\Theta)$ is the Jacobian matrix of the robot. Moreover, differentiating (24), we can get

$$\ddot{x}(t) = \dot{J}(\Theta)\dot{\Theta} + J(\Theta)\ddot{\Theta}. \tag{25}$$

The relationship between wrench and joint force can be described:

$$T_{ext} = J^{\mathrm{T}}(\Theta) f. \tag{26}$$

In addition, the robot manipulator dynamics in joint space is

$$M_{\Theta}(\Theta)\ddot{\Theta} + C_{\Theta}(\Theta, \dot{\Theta})\dot{\Theta} + G_{\Theta}(\Theta) + T_{fric} = T + T_{ext} \tag{27}$$

where $\dot{\Theta}$ and $\ddot{\Theta}$ are the vectors of velocity and accelerations, respectively. $Mq(\Theta) \in$ R^n is the inertia matrix; $C_{\Theta}(\Theta, \dot{\Theta})$ is the Coriolis and centripetal torque; $G_{\Theta}(\Theta)$ is the gravity; T is the robot torque; T_{fric} is the friction torque and T_{ext} is the external torque.

- *Property 1*: Matrix $M_{\Theta}(\Theta)$ is bounded above and below and positive definite symmetric.
- *Property 2*: Matrix $M_{\Theta}(\Theta)\ddot{\Theta} - 2C_{\Theta}(\Theta, \dot{\Theta})$ is skew symmetric. matrix.

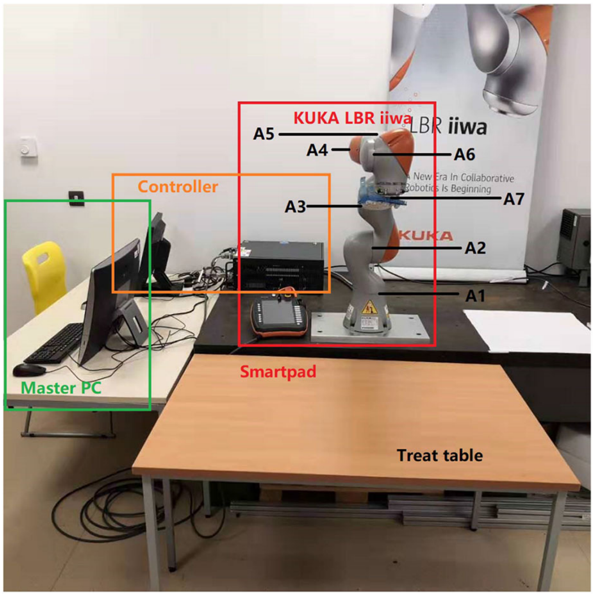


FIGURE 4 | Illustration of the experimental system.

A force-position model of the relationship between the external force and position in joint space is

$$D_g J(\Theta)(\ddot{\Theta}_r - |\ddot{\Theta}_g) + (C_g J(\Theta) + M_g \dot{J}(\Theta))(\dot{\Theta}_r - \dot{\Theta}_g) + K_g(\phi\left(q_r\right) - \phi\left(q_g\right)) = -J^{-T} T_{ext} \tag{28}$$

where $q_g \in \mathrm{R}^n$ and $q_r \in \mathrm{R}^n$ are the desired trajectory generated from GMR algorithm and virtual desired trajectory, respectively. The M_g, C_g, and K_g are gain matrix of the mass, damping and stiffness matrices designed by the controller, respectively.

- *Assumption 1*: c_1; c_2; and c_3 are positive constants, and both q_g and q_r are differentiable and bounded: $q_g, q_r \le c_1, \dot{\Theta}_g, \dot{\Theta}_r \le c_2, \ddot{\Theta}_g, \ddot{\Theta}_r \le c_3$.
- *Remark 1*: In the specific cases, force-position models such as the damping-stiffness model and stiffness model are applied

$$\begin{aligned} C_g\left(\dot{\Theta}_r - \dot{\Theta}_g\right) + K_g\left(q_r - q_g\right) &= -T_{ext} \\ K_g\left(q_r - q_g\right) &= -T_{ext} \end{aligned} \tag{29}$$

In the case that the desired manipulator's motion is free and no external collisions are generated, we can get $q_r = q_g$; $T_{ext} = 0$. In the opposite, while there is an external collision, the robot will generate and follow the new trajectory, which is the adaptation to the force-position model and the external torque specified in Bernhardt et al. (2005) illustrates the relationship.

Regarding to the safety consideration, a moveable limit has been added to the KUKA iiwa platform in both of the Cartesian spaces to make sure the manipulator can only reach the areas in front of it with a radian of 0.6 m; in this case, the robot manipulator cannot fully stretch. Meanwhile, the stiffness of its endpoint in all directions is set upon a reasonable level, hence participants can easily afford the force from the robot. Furthermore, by adjusting the threshold

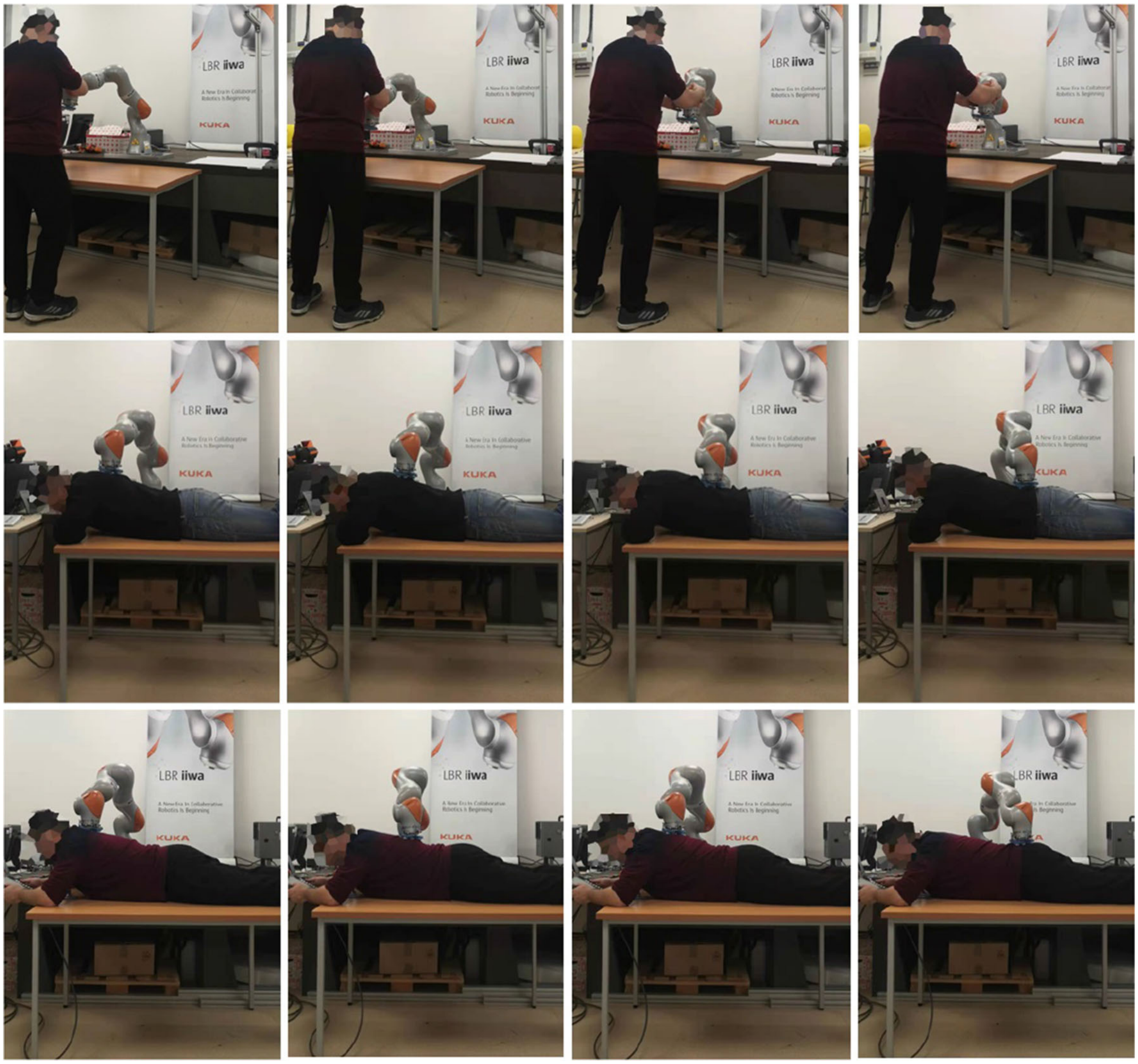


FIGURE 5 | Experiment snapshots of the Kuka LBR iiwa manipulator for massage tasks by the proposed hybrid position/force control method.

in **SafetyConfiguration.sconf** of the KUKA controller, the Collision Detection Framework can be activated. It is relevant to decide at what levels the external force is to lock down the robot to protect the participants. Here, we set up this parameter at a low level, which avoids all the touching with high force.

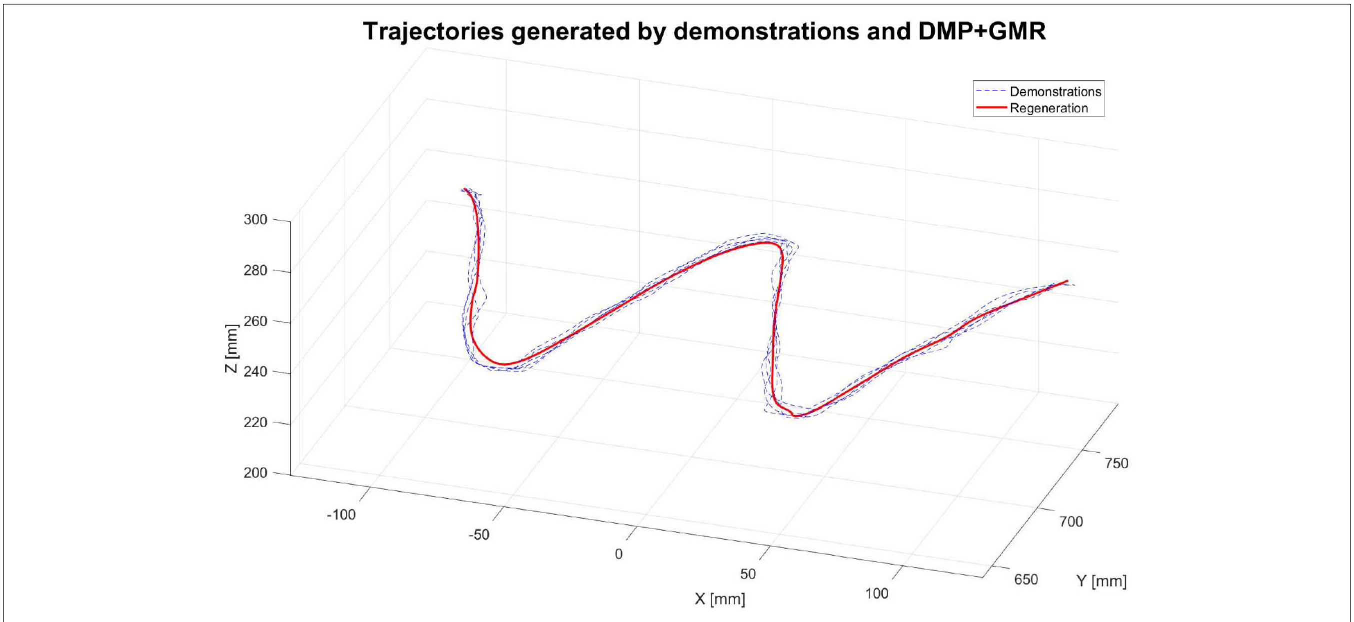


FIGURE 6 | The trajectories of the robot endpoint in Cartesian space generated by demonstrations and the proposed teaching interface.

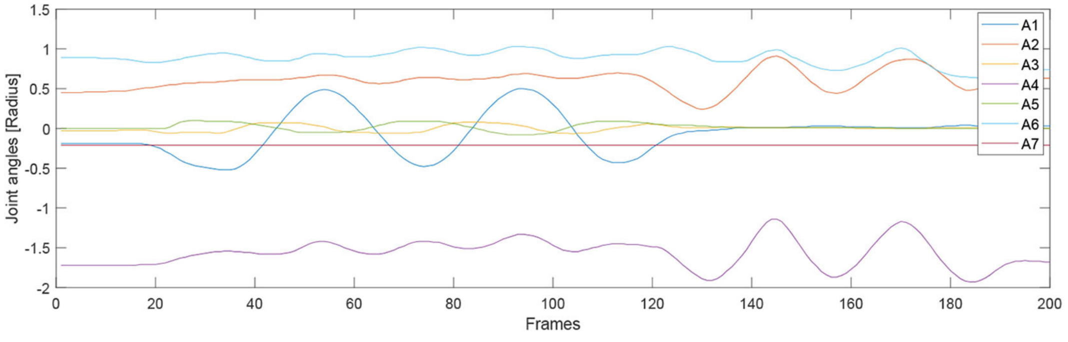


FIGURE 7 | Angular joint values of the robot while massaging the first participant.

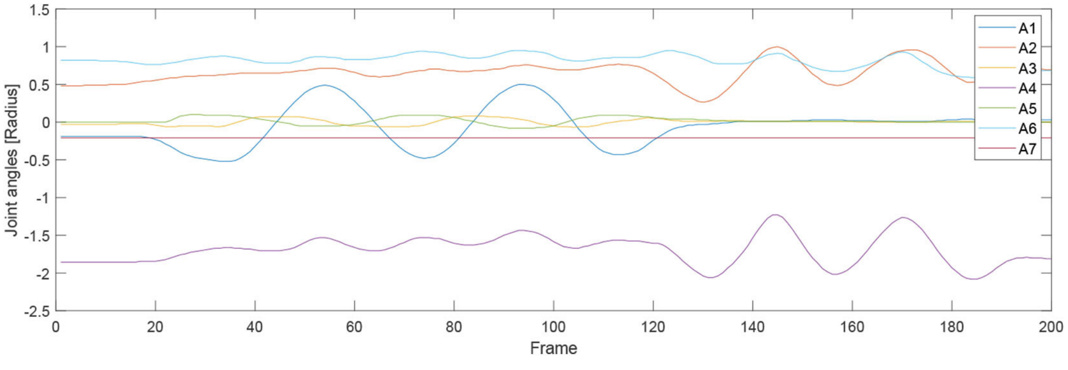


FIGURE 8 | Angular joint values of the robot while massaging the second participant.

5. EXPERIMENTAL STUDIES

5.1. Experiment Setup

A KUKA LBR iiwa robot, which has 7 DOFs of flexible joints, is implemented in our experimental studies to validate the proposed method. It is controlled utilizing the KUKA Smartpad. A massage glove was attached to the end-effector of the robot. As can be seen from **Figure 4**, there is a master computer lying by the robotic controller for offline data training purposes, and it is linked to the robot controller by an Ethernet cable. The labels A1–A7 in **Figure 4** show the joint actuator's position with force sensor each. In addition, there is a treat table placed in front of the robot base, which is in the workspace of the robot manipulator.

The control frequency is set as 10Hz for the KUKA LBR iiwa manipulator, and the running time is limited to 30 s for the massage path tracking, and thus 300 time samplings are executed for the control loop. The endpoint stiffness in X, Y, and Z directions, Roll, Pitch, and Yaw orientations are set as 1,000, 1,000, 100, 300, 300, and 300 N/m, respectively. The control gains M_g, C_g and K_g are gain matrix of the mass, damping and stiffness and respectively set as $diag[1.0]$, $diag[1.0]$, and 0.5. The reason for setting those values are that we planned the motion path in horizontal plane using GMM-evoluted DMP algorithms with implementing the force in Z directions, which resulted in the impedance effectiveness in the vertical direction of the patient's back.

First of all, we conducted an experiment to test the accuracy of our proposed robotic teaching interface where only position control was considered. As in some particular situations during the massage, the service robot may require to manipulate as accurately as our human carer; there are specific areas of the patients required to be massaged. This is the reason behind the accuracy of the teaching interface matters. In order to test the accuracy, firstly the position of the KUKA iiwa robot end-effector in the Cartesian space was chosen as the performance index. The human operator physically guided the robot to draw a sine curve for five times in the treat table by holding its end-effector. After training process, the robot could regenerate a new smooth trajectory. For the second experiment, one human operator physically taught the robot to perform the massage movements on the first participant by holding the end-effector of the robot. After the robot was taught, a participant as well as the operator himself were massaged by the robot; the participants slowly and smoothly lifted their back up and down. **Figure 5** shows the teaching-based massage process and we can observe that the KUKA LBR iiwa robot successfully accomplished the desired massage task with only one time teaching. Consequently, our proposed massage system could automatically fit different body shapes. Meanwhile, the instantaneous external force and torque of the robot endpoint in Cartesian space were outputted to the master PC for data analysis. The real position in both the Joint space and Cartesian space were plotted by MATLAB.

In addition, the third experiment has been conducted to validate the spatial generalization functions of our proposed robotic teaching interface. To do this, three participants were

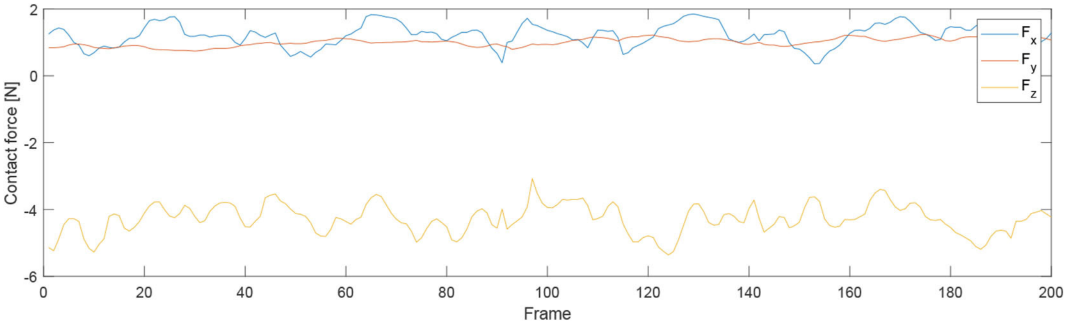


FIGURE 9 | Contact force variables of the end-effector of the robot in X Y Z directions during the massage for the first participant.

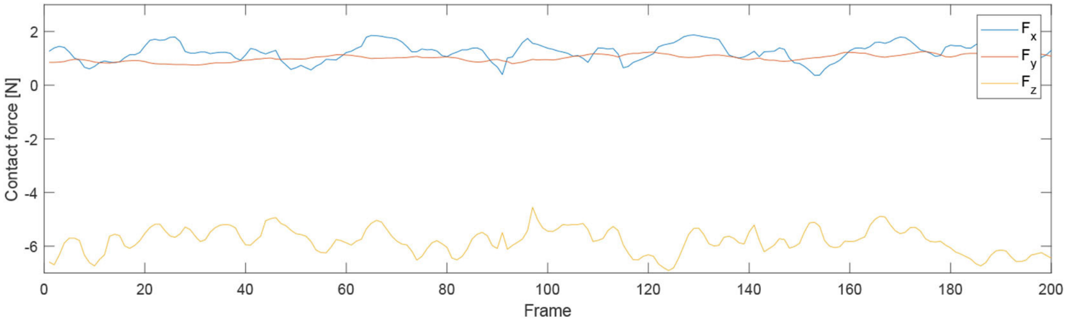


FIGURE 10 | Contact force variables of the end-effector of the robot in X Y Z directions during the massage for the second participant.

FIGURE 11 | Illustration of the third experimental process.

sought to be undertaken the massage services of their shoulder by seating under the robot manipulator. The operator taught the robot to do the massage task on all the three participants with only once teaching. All the parameters kept the same as those from the second experiment. The massage services were reproduced by the robot for all the participants one by one with different orders.

5.2. Experimental Results

The first group of experiments aimed to verify the learning performance of our proposed teaching interface when the demonstrations are defective. To verify the learning performance of the modified DMP better, we designed a drawing task for the robot, and the experiment setup is shown in **Figure 5**. In this experiment, the robot is required to draw an image of sinusoid on the paper after the human operator demonstrates the task five times. The parameters of the DMP model are set as $\tau = 1$, k = 25, c = 10, and $\alpha_f = 8$. As is shown in **Figure 6**, the demonstrations are defective and the curves are irregular. One of the reasons is that the demonstrator is drawing on the paper indirectly by holding the wrist of the robot, which affects the exertion of drawing skill. The demonstrations are modeled in the task space. As is shown in **Figure 6**, comparing the performance index of the demonstrations and the generated trajectory, a smooth curve is accurately retrieved from multiple demonstrations using the modified DMP without any unexpected drawing. The robot performs the drawing task after learning, and the curve that the robot draws is smoother than the demonstrations.

For the second experiment, two participants are massaged by the robot by only teaching once. During the massaging, the A7 joint of the robot is set as fixed value because it is only related to the end-effector's orientation. **Figures 7**, **8** illustrate the robot's angular data of all the seven joints when the robot was reproducing the massage on the first and second participant, respectively. Because the massage task was first along the direction of the back of the participants and then pounded the back; from the figures, we can thus see that the first joint of the robot A1 was firstly fluctuating and then kept stable; the A2, A4, and A6 joints were firstly kept stable and then fluctuated. This is because the robot via A1 joint moves left and right; via A2, A4, and A6, joints move up and down. Comparing **Figures 7**, **8**, we can notice that the A2 and A4 joints were increasing, and the A6 joint was decreasing during the whole massage process. This is caused by the fact that the second participant has a thicker body shape, and when the robot is in a lift-up configuration, its 6th joint will be more folded. **Figures 9**, **10** show the contact force variables of the end-effector of the robot in X, Y, and Z directions during the massage for the first and second participants. Here we define the moving direction is the positive direction of the robot endpoint. We can notice that the contact forces in X and Y directions are the positive values while those in Z direction are negative values with bigger figures. This is caused by the fact that, during the massage paths playback process, the endpoint of the robot in Z direction met resistance comparing to its original teaching configuration. In addition, the contact force vertical to the massage paths (Z direction) also differed with two participants. Contact force variables when massaging the first participant are in the interval $[-3.5\ -5]$ N, while the variables when the robot was massaging the second participant are in $[-5\ -6.5]$ N. This is owing to the two participants differing in body thickness. The third test has validated the generalization ability of our proposed massaging system. The training results are shown in **Figure 11**. The motions of the robot are regenerated from an one time teaching based demonstration, which synthesize the features of the demonstration and enable the robot to perform the massage task successfully as shown in **Figure 11**. The target order of the massage service is then modulated to be changed. A conclusion can be drawn from the video abstract; the profile of the reproduction is obtained.

5.3. Remark

Through the above conducted three experiments, we can note several things.

1. Our proposed hybrid position/force mode teaching interface was able to generate an accurate path after being taught, which reduces the errors in 3D space.
2. Our proposed hybrid position/force mode teaching interface was able to automatically and adaptively fit all body shapes with smooth force implementing.
3. The spatial generalization ability was validated, where the whole massage tasks can be segmented into several unit movement primitives, which can be regrouped into different

orders with only one-time teaching, and it promoted the working flexibility.

6. CONCLUSION

In this paper, an enhanced force-sensing and robotic-learning algorithm-based robotic teaching interface has been developed to perform the massaging tasks. In the motion generation part, the discrete DMP is selected as the basic motion model, which can generalize the motions. To improve the learning performance of the DMP model, the GMM, and GMR are employed for the estimation of the unknown function of the motion model. With this modification, the DMP model is able to retrieve a better motion from multiple demonstrations of a specific task. For the force input aspects, a hybrid force/position controller is introduced to ensure the safety of direct human-robot interaction. Several experiments have been performed on the KUKA LBR iiwa robot to test the performance of our proposed methods, which has proven that our proposed method can be used to establish a novel robot learning framework for massaging and facilitate the robot learning at a higher level. Our future work will focus on combining with visual monitoring technology, where the acupuncture point and bones of the patient's back will be clearly recognized and tracked, which results in better demonstrations for the robot to learn from.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

The written informed consent was obtained from the individuals for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

CL, AF, JS, and SL: conceptualization. CL and AF: methodology and writing–review and editing. CL: validation and formal analysis. AF and JS: resources. CL and SL: writing–original draft preparation. JS and SL: supervision and funding acquisition. AF: project administration. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbot.2020.00030/full#supplementary-material

REFERENCES

Bernhardt, M., Frey, M., Colombo, G., and Riener, R. (2005). "Hybrid force-position control yields cooperative behaviour of the rehabilitation robot lokomat," in *9th International Conference on Rehabilitation Robotics, 2005* (Chicago, IL), 536–539. doi: 10.1109/ICORR.2005.1501159

Calinon, S., and Billard, A. (2007). "Incremental learning of gestures by imitation in a humanoid robot," in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction* (Washington, DC: ACM), 255-262. doi: 10.1145/1228716.1228751

Chelikowsky, J. R., Troullier, N., and Saad, Y. (1994). Finite-difference-pseudopotential method: electronic structure calculations without a basis. *Phys. Rev. Lett.* 72:1240. doi: 10.1103/PhysRevLett.72.1240

Fanaei, A., and Farrokhi, M. (2006). Robust adaptive neuro-fuzzy controller for hybrid position/force control of robot manipulators in contact with unknown environment. *J. Intell. Fuzzy Syst.* 17, 125–144.

Field, T. (2018). Pain and massage therapy: a narrative review. *Curr. Res. Complement. Altern. Med.* doi: 10.29011/CRCAM-125/100025

Gear, C. W. (1988). Differential-algebraic equation index transformations. *SIAM J. Sci. Stat. Comput.* 9, 39–47. doi: 10.1137/0909004

Goradia, A., Xi, N., and Tan, J. (2002). "Hybrid force/position control in moving hand coordinate frame," in *7th International Conference on Control, Automation, Robotics and Vision, 2002. ICARCV 2002, Vol. 3* (Guangzhou), 1126–1131. doi: 10.1109/ICARCV.2002.1234931

Ha, Q., Nguyen, Q., Rye, D., and Durrant-Whyte, H. (2000). Impedance control of a hydraulically actuated robotic excavator. *Autom. Constr.* 9, 421–435. doi: 10.1016/S0926-5805(00)00056-X

Howlett, P. G., Pudney, P. J., and Vu, X. (2009). Local energy minimization in optimal train control. *Automatica* 45, 2692–2698. doi: 10.1016/j.automatica.2009.07.028

Kume, M., Morita, Y., Yamauchi, Y., Aoki, H., Yamada, M., and Tsukamoto, K. (1996). "Development of a mechanotherapy unit for examining the possibility of an intelligent massage robot," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS'96, Vol. 1* (Osaka), 346–353. doi: 10.1109/IROS.1996.570698

Li, C., Yang, C., Ju, Z., and Annamalai, A. S. (2018). An enhanced teaching interface for a robot using DMP and GMR. *Int. J. Intell. Robot. Appl.* 2, 110–121. doi: 10.1007/s41315-018-0046-x

Li, C., Yang, C., Wan, J., Annamalai, A., and Cangelosi, A. (2017). "Neural learning and kalman filtering enhanced teaching by demonstration for a baxter robot," in *2017 23rd International Conference on Automation and Computing (ICAC)* (Huddersfield, UK), 1–6. doi: 10.23919/IConAC.2017.8081985

Ma, L.-Z., Yao, G.-Y., Ni, Q.-L., and Zhu, Z. (2005). Study of the traditional Chinese medicine (TCM) massage robot based on the hybrid mechanism to obtain rolling treatment. *Mach. Des. Res.* 21, 43–46.

Matsubara, T., Hyon, S.-H., and Morimoto, J. (2011). Learning parametric dynamic movement primitives from multiple demonstrations. *Neural Netw*. 24, 493–500. doi: 10.1016/j.neunet.2011.02.004

Nanayakkara, T., Watanabe, K., Kiguchi, K., and Izumi, K. (2004). Evolving a multiobjective obstacle avoidance skill of a seven-link manipulator subject to constraints. *Int. J. Syst. Sci.* 35, 167–178. doi: 10.1080/00207720410001684530

Nazeer, K. A., and Sebastian, M. (2009). "Improving the accuracy and efficiency of the K-means clustering algorithm," in *Proceedings of the World Congress on Engineering, Vol. 1*, 1–3 (London, UK: Association of Engineers).

Nguyen, Q., Ha, Q., Rye, D., and Durrant-Whyte, H. (2000). "Force/position tracking for electrohydraulic systems of a robotic excavator," in *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No. 00CH37187), Vol. 5* (Sydney, NSW), 5224–5229. doi: 10.1109/CDC.2001.914787

Petitjean, F., Forestier, G., Webb, G. I., Nicholson, A. E., Chen, Y., and Keogh, E. (2014). "Dynamic time warping averaging of time series allows faster and more accurate classification," in *2014 IEEE International Conference on Data Mining* (Shenzhen), 470–479. doi: 10.1109/ICDM.2014.27

Schaal, S. (2003). "Movement planning and imitation by shaping nonlinear attractors," in *Proceedings of the 12th Yale Workshop on Adaptive and Learning Systems* (New Haven, CT: Citeseer).

Schaal, S. (2006). Dynamic movement primitives-a framework for motor control in humans and humanoid robotics, *Adaptive Motion of Animals and Machines* (Springer), 261–280. doi: 10.1007/4-431-31381-8_23

Scherillo, P., Siciliano, B., Zollo, L., Carrozza, M., Guglielmelli, E., and Dario, P. (2003). "Parallel force/position control of a novel biomechatronic hand prosthesis," in *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003)* (Kobe), *Vol. 2*, 920–925. doi: 10.1109/AIM.2003.1225465

Senin, P. (2008). *Dynamic Time Warping Algorithm Review*. Honolulu, HI: Information and Computer Science Department University of Hawaii, 855.

Silvério, J., Rozo, L., Calinon, S., and Caldwell, D. G. (2015). "Learning bimanual end-effector poses from demonstrations using task-parameterized dynamical systems," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Hamburg), 464–470. doi: 10.1109/IROS.2015.7353413

Tatiraju, S., and Mehta, A. (2008). Image segmentation using K-means clustering, EM and normalized cuts. *Depart. EECS* 1, 1–7.

Tong, K.-F., and Huang, J. (2008). New proximity coupled feeding method for reconfigurable circularly polarized microstrip ring antennas. *IEEE Trans. Antenn. Propag.* 56, 1860–1866. doi: 10.1109/TAP.2008.924736

Zhang, Y., and Zhang, W. (2017). Recent patents on Chinese massage robot. *Recent Patents Eng.* 11, 156–161. doi: 10.2174/2212797610666170127171713

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurorobotics

ORIGINAL RESEARCH
published: 05 November 2020
doi: 10.3389/fnbot.2020.566172

Muscle Fatigue Analysis With Optimized Complementary Ensemble Empirical Mode Decomposition and Multi-Scale Envelope Spectral Entropy

Juan Zhao[1,2], Jinhua She[3], Edwardo F. Fukushima[3], Dianhong Wang[1], Min Wu[1,2] and Katherine Pan[4]*

[1] School of Automation, China University of Geosciences, Wuhan, China, [2] Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan, China, [3] School of Engineering, Tokyo University of Technology, Tokyo, Japan, [4] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, United States

The preprocessing of surface electromyography (sEMG) signals with complementary ensemble empirical mode decomposition (CEEMD) improves frequency identification precision and temporal resolution, and lays a good foundation for feature extraction. However, a mode-mixing problem often occurs when the CEEMD decomposes an sEMG signal that exhibits intermittency and contains components with a near-by spectrum into intrinsic mode functions (IMFs). This paper presents a method called optimized CEEMD (OCEEMD) to solve this problem. The method integrates the least-squares mutual information (LSMI) and the chaotic quantum particle swarm optimization (CQPSO) algorithm in signal decomposition. It uses the LSMI to calculate the correlation between IMFs so as to reduce mode mixing and uses the CQPSO to optimize the standard deviation of Gaussian white noise so as to improve iteration efficiency. Then, useful IMFs are selected and added to reconstruct a de-noised signal. Finally, considering that the IMFs contain abundant frequency and envelope information, this paper extracts the multi-scale envelope spectral entropy (MSESEn) from the reconstructed sEMG signal. Some original sEMG signals, which were collected from experiments, were used to validate the methods. Compared with the CEEMD and complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), the OCEEMD effectively suppresses mode mixing between IMFs with rapid iteration. Compared with approximate entropy (ApEn) and sample entropy (SampEn), the MSESEn clearly shows a declining tendency with time and is sensitive to muscle fatigue. This suggests a potential use of this approach for sEMG signal preprocessing and the analysis of muscle fatigue.

Keywords: surface electromyography, complementary ensemble empirical mode decomposition, least-squares mutual information, multi-scale envelope spectral entropy, muscle fatigue

Edited by:
Zhan Li,
University of Electronic Science and Technology of China, China

Reviewed by:
Xinjun Sheng,
Shanghai Jiao Tong University, China
Suncheol Kwon,
National Rehabilitation Center, South Korea

***Correspondence:**
Jinhua She
she@stf.teu.ac.jp

Received: *27 May 2020*
Accepted: *18 September 2020*
Published: *05 November 2020*

Citation:
Zhao J, She J, Fukushima EF, Wang D, Wu M and Pan K (2020) Muscle Fatigue Analysis With Optimized Complementary Ensemble Empirical Mode Decomposition and Multi-Scale Envelope Spectral Entropy. Front. Neurorobot. 14:566172. doi: 10.3389/fnbot.2020.566172

1. INTRODUCTION

Muscle fatigue is defined as a temporary decrease in the physical force during exercise (Liu et al., 2014; Kyranou et al., 2018). During a continuous contraction, a localized muscle gradually undergoes biological changes and enters the state of muscle fatigue. While the mechanism of muscle fatigue is complicated, accurate detection of fatigue is of great significance for assessing functional impairment, planning training programs, and evaluating rehabilitation effect (Gandevia, 2013). For these reasons, the detection of muscle fatigue has been a hot topic in the field of rehabilitation and sports medicine over the last couple of decades.

A surface electromyography (sEMG) signal captures the state of muscle activity and motor function and is considered as an effective tool to evaluate local muscle fatigue (Chowdhury et al., 2013). The changes in sEMG signals correlate to the number of motor units, activity patterns, metabolic situations, and other factors (Srhoj-Egekher et al., 2011). An sEMG signal is non-stationary during muscle dynamic contractions and shows a high degree of complexity (Zhang Q. et al., 2017). Thus, it is a reliable approach to processing sEMG signals by a non-linear method and to extract muscle fatigue features from the complexity.

An essential part of processing a non-stationary signal is to find a way to represent the oscillatory modes of the signal. An sEMG signal consists of many single harmonic signals, and each contains only one oscillatory mode. Introduced by Huang et al., an intrinsic mode function (IMF) is a single harmonic signal model that gives sharp identifications of embedded structures through producing instantaneous frequencies as functions of time (Huang, 2000). With the empirical mode decomposition (EMD), a non-stationary signal is decomposed into IMFs and performed time-frequency analysis (Huang, 2000). The EMD was first used to filter the activity attenuation of sEMG signals (Andrade et al., 2006). Later, it was widely used for the artifact removal and feature extraction of sEMG signals (Pilkar et al., 2017). Unfortunately, the mode-mixing problem often occurs for the reason that the IMFs contain signals of different scales or the signals of a similar scale spread in different IMFs (Hu et al., 2012). To solve this problem, a method of analyzing noise-assisted signals called ensemble EMD (EEMD) was presented (Wu and Huang, 2009). The EEMD has been used not only to filter various noises (Zhang and Zhou, 2013) but also to quantitatively analyze the features of sEMG signals (Wu et al., 2017; Zhang Y. et al., 2017). This method has a good effect on suppressing mode mixing decomposed from weak discontinuous signals with high-frequency noise, but the suppression effect is very limited for signals with similar frequencies. Therefore, the EEMD may leave a small amount of noise in reconstructed signals.

Many methods have been proposed to further reduce noise for the EEMD, such as a complementary EEMD (CEEMD) (Yeh et al., 2010) and a complete EEMD with adaptive noise (CEEMDAN) (Torres et al., 2011). They improved the EEMD in different ways. The CEEMDAN adds adaptive white noise to the original signals and obtains IMFs by averaging the modal components at each stage of signal decomposition. However, the CEEMDAN still has some problems like spurious modes and high computational cost (Rezaie-Balf et al., 2019; Li et al., 2020). The CEEMD adds a pair of Gaussian white noises with equal amplitudes and a relative phase difference of 180° to the original signal, and then performs the EMD decomposition of the two groups of signals.

The CEEMD decomposes a signal into IMFs based on the characteristics of the signal itself, which is very important for analyzing a non-stationary signal. It was used in some fields and its applications demonstrate successful results (Zhao et al., 2014; Lu et al., 2019). Based on the advantages of the CEEMD, this paper employs it to preprocess sEMG signals, which lays a good foundation for next signal reconstruction, de-noising, and feature extraction. However, the CEEMD also has a small number of mode-mixing components and long iterative times when it decomposes a signal that exhibits intermittency and contains components with proximity spectrum into intrinsic mode functions (IMFs) (Zheng et al., 2014; Chen and Wang, 2017).

This paper analyzes the causes of these problems. On one hand, there is information coupling between IMFs, which leads to the existence of similar signal components in different IMFs. On the other hand, the standard deviation of added Gaussian white noise is not suitable. As a result, the signal decomposition is not complete, or the number of iterations increases in order to achieve a specific decomposition effect. Therefore, this paper presents an improved CEEMD to overcome these defects.

After the sEMG signal is preprocessed, the feature of evaluating muscle fatigue needs to be extracted. The quality of the feature has a crucial influence on the classification and predicting of muscle fatigue. The complexity of sEMG is a commonly used index to reflect the physiological characteristics of muscle activity (Talebinejad et al., 2011; Cashaback et al., 2013). Extracting entropy from sEMG is an effective method for analyzing muscle fatigue. Approximate entropy (ApEn) and sample entropy (SampEn) are the earliest entropy parameters and are still in use today (Pethick et al., 2019; Xie et al., 2019). However, they are subject to the problem of tolerance selection. Furthermore, Costa et al. found that the single-scale entropy of a healthy person often conflicts with that of a heart-disease patient for the use of their heartbeat fluctuations. To solve these problems, they proposed the concept of multi-scale entropy (MSEn), which is calculated on multiple time scales of a signal (Gao et al., 2015). Different time scales are obtained through a coarse granulation process, and provide more complete information on signals than a single scale does. Note that the evolution of muscle fatigue needs to be described in this study, the selected entropy-based indicator should be related to the characteristics of sEMG signals.

The purpose of this study was to devise a method, which is called optimized CEEMD (OCEEMD), to process sEMG signals with high precision and high efficiency, and to present a new entropy-based index for muscle fatigue. The OCEEMD integrates the least-squares mutual information (LSMI) (Kimura and Sugiyama, 2011) and the chaotic quantum particle swarm optimization (CQPSO) algorithm (Valdez et al., 2017) with the conventional CEEMD. The LSMI is used to calculate the correlation between IMFs to reduce their coupling. And

TABLE 1 | Information of 10 subjects involved in experiments (Min., Minimum; Max., Maximum; Avg., Average; Std. Dev., Standard Deviation).

	Min.	Max.	Avg.	Std. Dev.
Age (years)	23	39	29.7	12.1
Height (cm)	155.6	182.9	170.3	7.8
Weight (kg)	44	86.4	68.5	2.7

the CQPSO is used to search an optimal solution of the standard deviation of Gaussian white noise to optimize a related parameter. The advantage of the OCEEMD is that it suppresses mode mixing and improves decomposition efficiency by making the best use of the algorithms. After the OCEEMD decomposes an sEMG signal into a set of IMFs adaptively, the useful IMFs are reconstructed to a de-noised signal, eliminating most of the interference. Then, multi-scale envelope spectral entropy (MSESEn) is calculated. This index improves the reliability of analyzing muscle fatigue based on the rich frequency and envelope information of IMFs. Verification was carried out for collected sEMG signals in a pedaling experiment for the presented and the related methods. The results show that, compared with other methods, the OCEEMD suppresses the mode mixing with high efficiency for the decomposition of sEMG signals, and the extracted feature shows a high sensitivity to the sEMG changes with muscle fatigue.

2. MATERIALS AND METHODS

This section first introduces the experimental protocol and tests. Then, it describes signal preprocessing and feature extraction of an sEMG signal.

2.1. Subjects and sEMG Data Acquisition

Ten healthy subjects (two females and eight males) performed the experiments at the Advanced Mechatronics Laboratory, School of Engineering, Tokyo University of Technology, Japan (**Table 1**). They were covered by personal accident insurance for students pursuing education and research provided by Japan Educational Exchanges and Services. The advisability of students' involvement in experiments and the experimental protocol were first assessed by the ethical committee of Tokyo University of Technology. All participants signed informed consent.

Each subject sat in front of a pedaling machine that was developed for the rehabilitation training of lower limbs (**Figure 1**). The training load, inclined pedal angle, seat height, distance from the machine, and other parameters were calibrated to suit the subject before running the experiments. Each subject placed his/her left foot on the pedal for 5 min of up-down pedaling, with appropriate training load so as to feel local muscle fatigue after training. The sEMG data were collected from four muscles of lower limbs: rectus femoris, biceps femoris, tibialis anterior, and gastrocnemius (She et al., 2017a,b). Four electrodes (Model: Biometrics SX230-1000; Origin: UK) were used to collect sEMG signals, which were attached on the muscle belly along the muscle fibers. The sEMG signals were sampled at 1,000 Hz with a fixed gain of 60 dB (amplification factor: 1000). The data were stored in a laptop computer [Model: DELL Precision M3800; OS: Windows 8.1 Pro 64 bits; CPU: Intel(R) Core(TM) i7-4712HQ; RAM: 16.0 GB] (Zhao et al., 2018).

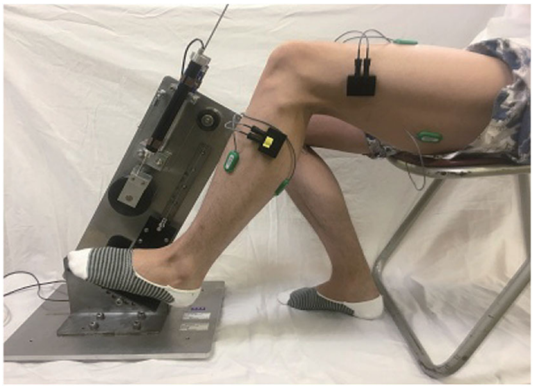


FIGURE 1 | A shot of pedaling experiment.

2.2. Optimized Complementary Ensemble Empirical Mode Decomposition

An original sEMG signal contains noise and invalid frequency components. So, we need to perform signal preprocessing to extract true information. An improved CEEMD called optimized CEEMD (OCEEMD) is used to decompose the original sEMG signal into a set of IMFs, solving the mode-mixing problem with high iteration efficiency. Then according to their frequency characteristics, the useful IMFs are selected to reconstruct a new signal, achieving the effect of de-noising.

The CEEMD decomposition has three main processes. First, a pair of Gaussian white noises that have the same amplitudes and a relative phase difference of 180° are added to an sEMG signal to form two new signals by making use of the uniform distribution property of the white-noise spectrum. Then, the two new signals are decomposed and their IMF candidates are obtained, separately. Finally, every two IMF candidates at the same level are averaged as the IMF component of the original signal, and thus the corresponding residual function is calculated (Yeh et al., 2010).

Since two similar white-noise signals are added to the sEMG signal with opposite phases, the residual noises in the IMFs are only different in their signs. They cancel with each other when the IMFs are added to form a new IMF that contains little noise. This method effectively suppresses mode mixing. However, it is difficult to completely eliminate mode mixing in actual applications. It is necessary to investigate the causes of mode mixing so that we can further reduce it.

2.2.1. Least-Squares Mutual Information

Mode mixing occurs when the components of similar scales reside in different IMFs. At this time, there is information coupling between the IMFs, that is, the IMFs are not completely orthogonal to each other. So, the existence of information coupling between the IMFs is a primary cause of mode mixing.

Theoretically speaking, eliminating the coupling between IMF components ensures the perfect orthogonality of the IMF components. Thus, it solves the mode-mixing problem completely. Mutual information (MI) is a non-parametric and non-linear measure indicator that quantitatively represents the statistical correlation between two random variables in information theory (Valdez et al., 2017). Since an sEMG signal has a zero-mean characteristic, the MI can be used to measure the coupling degree of IMF components obtained by the CEEMD according to the equivalence principle of irrelevance and orthogonality of zero-mean random signals. That is, it is possible to use the MI to measure whether or not there is mode mixing and the level of mixing. Thus, the first improvement of the OCEEMD on the CEEMD is the embedding of the least-squares mutual information (LSMI) in the CEEMD.

For two given IMFs, $\{x_i\}_{i=1}^{n}$ and $\{y_j\}_{j=1}^{n}$, their LSMI is calculated as follows.

Let the marginal probabilities of x_i be and y_j be $p(x_i)$ and $p(y_j)$, respectively; and the joint probability of x_i and y_j be $p(x_i, y_j)$. The MI of two IMFs is defined to be

$$\mathrm{MI} = \sum_{i=1}^{n}\sum_{j=1}^{n} p(x_i, y_j) \ln \psi(x_i, y_j), \tag{1}$$

where $\psi(x_i, y_j)$ is a density-ratio function

$$\psi(x_i, y_j) = \frac{p(x_i, y_j)}{p(x_i)p(y_j)}. \tag{2}$$

MI is non-negative. It is zero if x_i and y_j are statistically uncorrelated, that is,

$$\psi(x_i, y_j) = 1. \tag{3}$$

The logarithmic function in (1) is sensitive to a change in $\psi(x_i, y_j)$ around one. This degrades the accuracy of MI around zero. Note that based on the Taylor's series approximation, when $\psi(x_i, y_j) - 1 \to 0$,

$$\ln\left\{1 + [\psi(x_i, y_j) - 1]\right\} \approx \psi(x_i, y_j) - 1 \tag{4}$$

and

$$\left\{1 + [\psi(x_i, y_j) - 1]\right\}^2 \approx 1 + 2[\psi(x_i, y_j) - 1] \tag{5}$$

hold. Combining (4) and (5) yields

$$\ln \psi(x_i, y_j) \approx \frac{1}{2}[\psi^2(x_i, y_j) - 1] \tag{6}$$

Thus,

$$p(x_i, y_j) \ln \psi(x_i, y_j) \approx \frac{1}{2} p(x_i)p(y_j)[\psi(x_i, y_j) - 1]^2 \tag{7}$$

when $\psi(x_i, y_j) \to 1$.

Substituting (7) into (1) gives a squared-loss mutual information (SMI) (Kimura and Sugiyama, 2011)

$$\mathrm{SMI} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}[\psi(x_i, y_j) - 1]^2 p(x_i)p(y_j). \tag{8}$$

An equivalent form of (8) is used to simplify the calculation:

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} p(x_i, y_j)\psi(x_i, y_j) - \frac{1}{2}. \tag{9}$$

Since neither $p(x_i)$, $p(y_j)$, nor $p(x_i, y_j)$ is known, LSMI is used to further approximately calculate $\psi(x_i, y_j)$ so as to approximate SMI from a paired data set of x_i and y_j: (x_1, y_1), $\cdots$, (x_n, y_n) (Kimura and Sugiyama, 2011). In this study, $\psi(x_i, y_j)$ was approximated by a Gaussian-radial-basis-function model

$$\psi_\Theta = \sum_{i=1}^{n} \Theta_i K(x_i, y_j), \tag{10}$$

where Θ_i $(i = 1, \ldots, n)$ are estimated parameters and

$$K(x_i, y_j) = \exp\left(-\frac{\|x_i - y_i\|_2^2}{2h^2}\right) \tag{11}$$

is a kernel function. h in (11) is the width of the kernel.

The problem of optimizing the squared error of the estimated parameters using an empirical approximation is defined to be

$$\min_{\Theta}\left\{\frac{1}{2}\Theta^T\hat{H}\Theta - \Theta^T\hat{h} + \frac{\lambda}{2}\|\Theta\|^2\right\}, \tag{12}$$

where λ is a regularization parameter; and $\hat{H}$ and $\hat{h}$ are an $n \times n$ matrix and an n-dimensional vector, respectively. They are

$$\begin{cases} \hat{H} = \dfrac{1}{n^2}\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i, y_j)K(x_i, y_j)^T, \\ \hat{h} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i, y_j). \end{cases} \tag{13}$$

The analytical solution of (12) is

$$\hat{\Theta} = (\hat{H} + \lambda I)^{-1}\hat{h}, \tag{14}$$

where I is the identity matrix. Thus, LSMI is given by

$$\mathrm{LSMI} = \frac{1}{2}\hat{h}^T(\hat{H} + \lambda I)^{-1}\hat{h} - \frac{1}{2}, \tag{15}$$

which is used to approximate SMI. Note that LSMI between two IMFs can be calculated directly. If all LSMIs for an sEMG signal are sufficiently small, all of the IMFs are orthogonal to each other. This means that there is no mode mixing.

2.2.2. Chaotic Quantum Particle Swarm Optimization

The standard deviation of the added Gaussian white noise is an important parameter in the CEEMD. An unsuitable value may result in a large number of iterations and IMF modes and may cause mode mixing. So, the unsuitable standard deviation of Gaussian white noise is another cause of mode mixing.

Let d be the standard deviation of the added white noise. The selection of d requires careful consideration. As a rule of thumb, d is usually chosen to be a value between 0.1 and 0.2 (Torres et al., 2011). But choosing a different number between 0.1 and 0.2 inevitably leads to different decomposition results. The chaotic quantum particle swarm optimization (CQPSO) is an evolutionary computational algorithm (Huang, 2016). The main purpose of the algorithm is to obtain an optimal solution by sharing information among individuals of particle populations. It can be used here to search for the optimal solution of d. Thus, the second improvement of the OCEEMD on CEEMD is the embedding of CQPSO in the CEEMD to search for an optimal d.

LSMI is used as a fitness function for the CQPSO:

$$\arg \min_{d} \text{LSMI}. \tag{16}$$

A minimum LSMI corresponds to an optimal d, which is used to determine the amplitudes of the white noise.

At the beginning of the tth iteration for a M-dimensional space with N individuals, let the location of the ith particle be $X_i = \{X_{ij}\}_{j=1}^{M}$ $(i = 1, 2, \ldots, N)$, the historical location of the ith particle be $P_i = \{P_{ij}\}_{j=1}^{M}$, and the searched optimal locations of particle populations be $P_k^{(t)} = \{P_{kj}^{(t)}\}_{j=1}^{M}$ $(k = 1, 2, \ldots, N)$. The CQPSO of searching for an optimal d has the following steps.

Step 1: Initialize the particle population by chaos and randomly generate a series of parameters for d with N individuals. Initialize the locations of the particles.

Step 2: On the tth iteration, calculate the fitness of each particle, LSMI, at different locations, and compare each with the corresponding historical optimal fitness. If the current fitness is smaller than the historical one, replace the location vector P_i by the current one, X_i. Otherwise, keep P_i unchanged.

Step 3: Determine the optimal locations $P_k^{(t)}$ by comparing all particles with their corresponding optimal fitness—the minimum of LSMI.

Step 4: Update the location of each particle and calculate the corresponding fitness like Step 2. Retain the particle with the best performance in the population according to the fitness function. Then, update P_i and $P_k^{(t)}$.

Step 5: Check whether or not the current d meets the preset accuracy requirement and if the number of iteration is larger than a prescribed one. If not, let $t = t + 1$ and go to Step 2. Otherwise, stop and output the solution of d that corresponds to the optimal location $P_k^{(t)}$.

The OCEEMD incorporating the LSMI and CQPSO reduces mode mixing and improves iteration efficiency. It is used to decompose an sEMG signal to IMFs with real physical significance and a final residual according to a time scale. Then, all useful IMFs are selected and added to reconstruct a new signal in which most interference is removed.

2.3. Multi-Scale Envelope Spectral Entropy

A multi-scale entropy is a sample entropy measurement on multiple time scales and can analyze the complexity of a signal on different scales (Costa et al., 2008). Consequently, it provides complete information about an sEMG signal. The envelope spectrum is a spectral analysis method that is sensitive to shock components of a signal and reflects sudden changes in an sEMG signal (Lv et al., 2015). Since IMF components after the OCEEMD decomposition preserve rich information about frequency and envelope, we use multi-scale envelope spectral entropy (MSESEn) as an index for the detection of muscle fatigue.

The calculation of the MSESEn has two main steps: First, calculate the envelope spectrum of a signal. Second, calculate the multi-scale entropy (MSEn) of the envelope spectrum. The steps are explained as follows:

In the first step, we carry out the Hilbert-Huang transformation (HHT) on a reconstructed signal, $f(t)$,

$$H(f(t)) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{f(s)}{t - s} ds. \tag{17}$$

Structuring an analytic function yields the envelope of the signal

$$B(t) = \sqrt{f(t) + H^2(f(t))}. \tag{18}$$

Then, we use the fast Fourier transform (FFT) to obtain the demodulation spectrum

$$B(\omega) = \text{FFT}(B(t)). \tag{19}$$

In the second step, dividing a sequence of an envelope spectrum that contains N data points, $\{z_i\}_{i=1}^{N}$, into K segments with a scale τ, that is, $\{z_j, z_{j+1}, \ldots, z_{j+\tau}\}_{j=1}^{K}$, where K is the integer part of N/τ. The series of continuous coarse granulation is $\{w_j^{(\tau)}\}_{j=1}^{K}$ and

$$w_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} z_j, \; 1 \leq j \leq K. \tag{20}$$

Then, for the sequence $\{w_j^{(\tau)}\}_{j=1}^{K}$, embedding m adjacent points constructs a new sequence, $\{w_k^{(\tau)}, w_{k+1}^{(\tau)}, \ldots, w_{k+m}^{(\tau)}\}_{k=1}^{K-m+1}$. The distance $d_{i,j}^{(\tau)}$ between different elements, $(w_i^{(\tau)}, w_{i+1}^{(\tau)}, \ldots, w_{i+m}^{(\tau)})$ and $(w_j^{(\tau)}, w_{j+1}^{(\tau)}, \ldots, w_{j+m}^{(\tau)})$, is calculated for all $i \neq j$. For a threshold γ, let the number of the distances satisfying $d_{i,j}^{(\tau)} \leq \gamma$ be $n^{(\tau,m)}$. The ratio of $n^{(\tau,m)}$ to $K - m + 1$ is

$$C_i^{(\tau,m)}(\gamma) = \frac{n^{(\tau,m)}}{K - m + 1}. \tag{21}$$

The correlation degree between different vector elements is

$$C^{(\tau,m)} = \frac{1}{K - m + 1} \sum_{i=1}^{K-m+1} C_i^{(\tau,m)}(\gamma). \tag{22}$$

Similarly, $C^{(\tau,m+1)}$ is obtained for the sequence $\{w_j^{(\tau)}\}_{j=1}^{K}$ that embedding $m+1$ adjacent points into the sequence $\{w_k^{(\tau)}, w_{k+1}^{(\tau)}, \dots, w_{k+m+1}^{(\tau)}\}_{k=1}^{K-m}$.

Finally, the MSEn of envelope spectrum, MSESEn, is defined to be the set of sample entropy (SampEn) for the scale τ

$$\text{MSESEn} = \text{SampEn}(\tau, m, \gamma) = -\ln \frac{C^{(\tau,m+1)}}{C^{(\tau,m)}}. \tag{23}$$

2.4. Flow of Signal Processing

After the OCEEMD is used to adaptively decompose and reconstruct an sEMG signal, MSESEn is applied to evaluate the state of muscle fatigue. In the OCEEMD, finding an optimal solution of the standard deviation of Gaussian white noise helps the sEMG signal to be decomposed quickly in an appropriate way. Nevertheless, not all sub-signals can be decomposed thoroughly. Thus, we calculate coupling degrees between IMFs to detect whether there is mode mixing between IMFs. When mode mixing is detected, we adjust the residual and then perform the decomposition process again.

Note that the decomposition principle of the CEEMD allows us to use the correlation between an IMF component and a residual at the same level to replace the correlation between adjacent IMF components. This is due to the fact that IMF components is determined by the residual at the previous lever (Lu et al., 2019). This treatment features small number of iteration and high computational efficiency.

Summarizing the above explanation gives the following sEMG signal processing steps based on the OCEEMD and MSESEn (**Figure 2**).

Step 1: Calculate the standard deviation σ of an sEMG signal $x(t)$. Use the CQPSO to search for an optimal d to determine the amplitude of added white noise.

Step 2: Decompose the signal with white noise using the CEEMD and obtain an IMF candidate $c(t)$ and a residual $r(t)$.

Step 3: Calculate LSMI between $c(t)$ and $r(t)$. Compare LSMI with the selected threshold θ empirically. If LSMI $\leq \theta$, keep this $c(t)$ as an effective IMF and go to Step 4. Otherwise, discard this $c(t)$ as an invalid IMF and go to Step 5 to adjust the residual $r(t)$.

Step 4: Check whether or not $r(t)$ meets the terminal condition, that is, whether or not $r(t)$ is a monotonous function (Yeh et al., 2010). If it is, go to Step 6. Otherwise, take this $r(t)$ as a new signal $x(t)$ and go to Step 1 to start another level of decomposition.

Step 5: Remove the overlapping residual information using

$$\widetilde{r}(t) = r(t) - \text{LSMI} \times c(t), \tag{24}$$

take this new residual $\widetilde{r}(t)$ as $x(t)$, and go to Step 1.

Step 6: Calculate the spectrum of each IMF component, reconstruct a new sEMG signal, $f(t)$.

Step 7: Obtain the envelope of the reconstructed signal, $f(t)$, by calculating the demodulation spectrum based on the HHT.

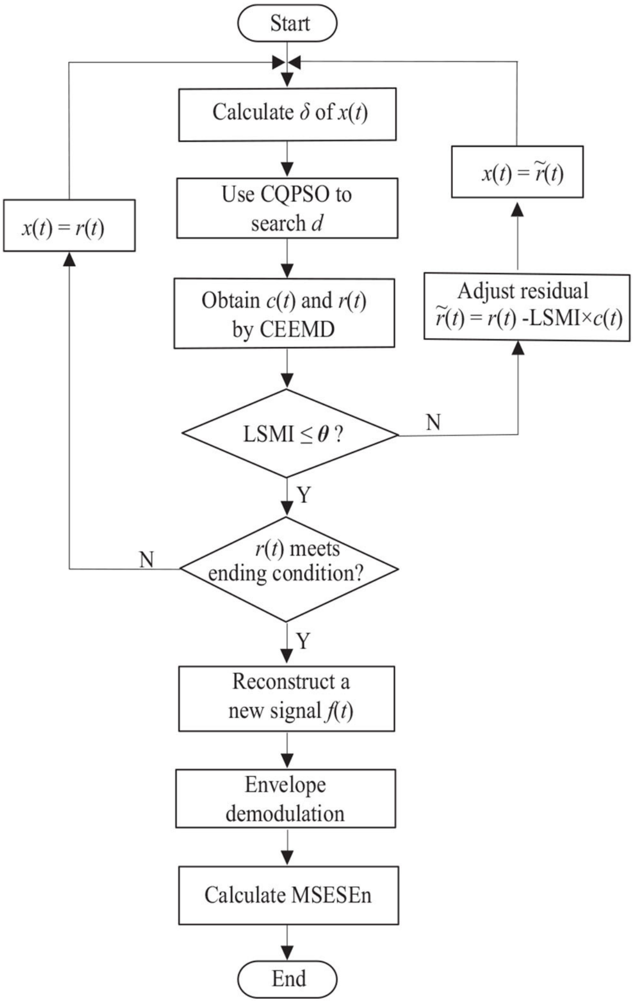


FIGURE 2 | Flow chart of sEMG signal processing.

Step 8: Calculate the MSESEn, and thus extract the feature.

In Step 3, if there is LSMI $\leq \theta$, which means there is little overlapping between $c(t)$ and $r(t)$, $c(t)$ and $r(t)$ are orthogonal. On the other hand, if there is LSMI $> \theta$, it means that modes are mixing between $c(t)$ and $r(t)$.

3. RESULTS

The presented method was used to process sEMG signals of all the ten subjects. The processing results of a sEMG signal sampled from the rectus femoris of one subject (age: 39, height: 163 cm, weight: 60 kg) is used as an example to show the typical results.

3.1. Acquired Signal

sEMG signals were sampled at a frequency of 1000 Hz. An original signal is shown in **Figure 3**. The horizontal axis is the sampling time, and the vertical axis is the amplitude of the sEMG. **Figure 3A** shows the sEMG signal for the whole length

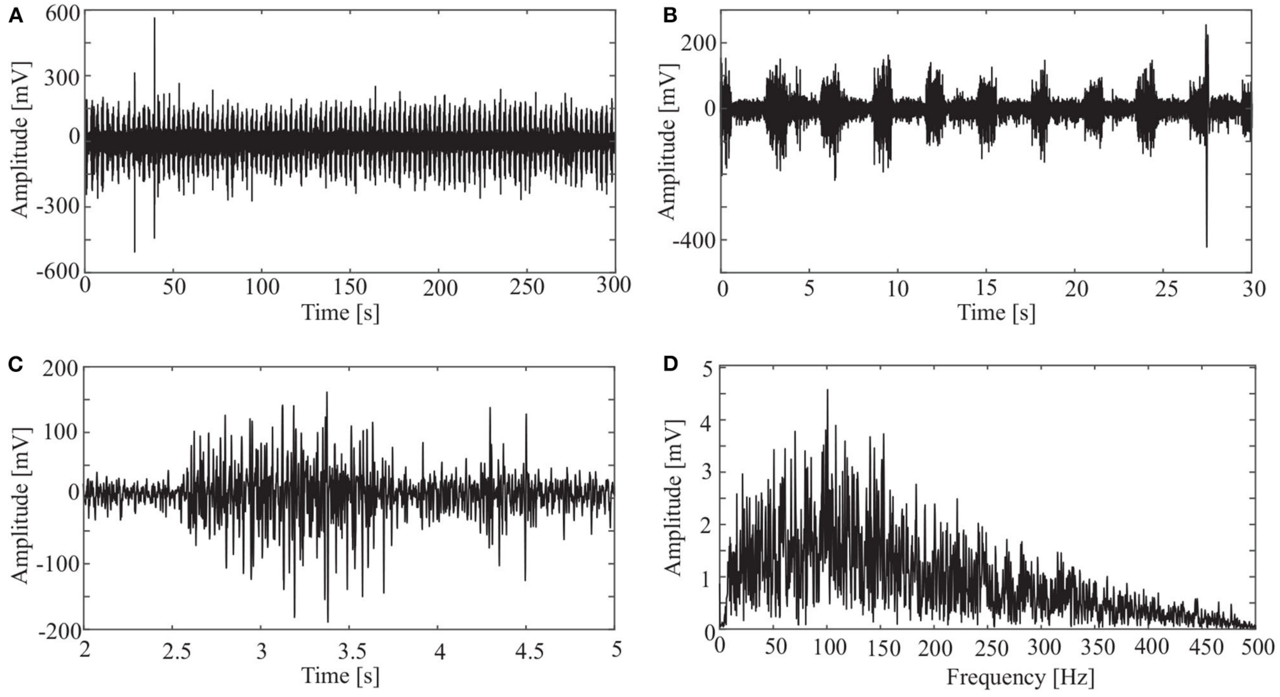


FIGURE 3 | Original sEMG signal and processing result: **(A)** sEMG for the whole length, **(B)** sEMG for the first 30 s, **(C)** sEMG from 2 to 5 s, and **(D)** spectrum of sEMG from 2 to 5 s.

($N = 300,000$). **Figure 3B** shows the signal for the first 30 s. It is clear from the figure that the period of the wave is about 3 s. Considering that a period of wave reveals the pedal movement, we chose the signal from 2 to 299 s for processing. Thus, the first period is from 2 to 5 s (**Figure 3C**), and there are 99 periods for the whole data set.

It is clear from **Figure 3D** that most frequency components of the sEMG signal are in the range [10, 200] Hz.

3.2. Signal Decomposition and Reconstruction

In addition to the OCEEMD, we also used the CEEMD and the CEEMDAN for comparison. Decomposing the signal in **Figure 3D** and obtained a series of IMFs (**Figures 4A**, **5A**, **6A**). These IMFs represent the characteristics of the original signal on different scales. Next, we got the spectrum of each IMF by fast Fourier transform (FFT) (**Figures 4B**, **5B**, **6B**). Then, we reconstructed a signal using effective IMFs extracted based on useful spectral distributions (**Figures 4C**, **5C**, **6C**), and obtained the corresponding spectrums (**Figures 4D**, **5D**, **6D**).

The spectrum of the second IMF in **Figure 4B** and that of the second and third IMFs in **Figure 5B** are mixed with other modes. But this does not appear in **Figure 6B**. This shows that mode mixing is suppressed more successfully for the OCEEMD than that for the CEEMD and the CEEMDAN.

sEMG signals are concentrated within a certain frequency. The highest frequency is about 200 Hz in **Figure 4D**, 180 Hz in **Figure 5D**, and 150 Hz in **Figure 6D**. It suggests that the reconstructed signal in **Figure 6C** has the minimum noise, which shows that the OCEEMD outperforms the CEEMD and the CEEMDAN.

The qualitative and quantitative analysis was carried out to compare the three methods using the root mean square error (RMSE), the number of IMF components, and the standard deviation of the amplitude ratio, d (**Table 2**). The results reveal the follows:

(1) RMSE of the reconstructed signal is 2.8986×10^{-4} for the OCEEMD. It is much smaller than those for the CEEMD and the CEEMDAN. Note that the smaller RMSE is, the difference between the de-noised sEMG signal and the original signal is. So, it means that the reconstructed signal using the OCEEMD has the highest accuracy.
(2) The number of IMFs is 6 for the OCEEMD. It is much smaller than those for the other two methods. This suggests that the computational efficiency of the OCEEMD is higher than that of the CEEMD and the CEEMDAN.
(3) d was adjusted in the OCCEMD and the CEEMDAN to adaptively process decomposition, while it was fixed in the CEEMD. A comparison between the CEEMD and the OCEEMD shows that only the first value of d is larger for the OCEEMD than for the CEEMD, and the others are all smaller for the OCEEMD than for the CEEMD. A comparison between the CEEMDAN and the OCEEMD shows that only the first value of d is the

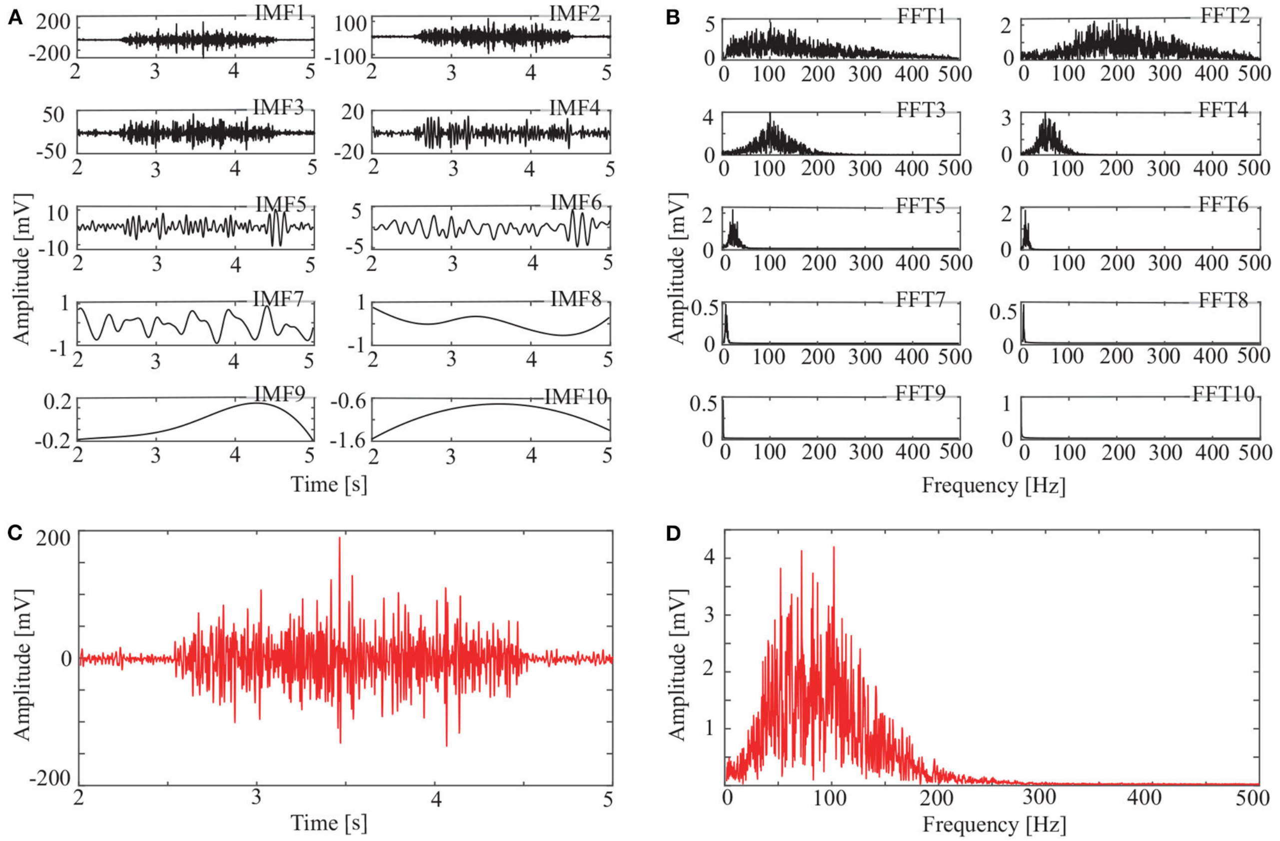


FIGURE 4 | IMFs and spectrums using CEEMD: **(A)** IMFs of signal in **Figure 3C**, **(B)** Corresponding FFTs, **(C)** Reconstructed signal, and **(D)** Spectrum of reconstructed signal.

same for both methods, and the others are all smaller for the OCEEMD than for the CEEMDAN. It shows that the OCEEMD has stronger adaptability than the other two methods.

(4) The LSMIs of adjacent IMFs are less than 0.1 for the OCEEMD. They are much smaller than those for the other two. The adjacent IMFs of the OCEEMD have the minimum coupling degrees, indicating that the OCEEMD suppresses the mode mixing very well.

In summary, **Table 2** reveals that the OCEEMD is superior to the CEEMD and the CEEMDAN while de-noising sEMG signals.

3.3. Extracted Entropy

This study used the MSEEn to quantify the envelope spectrum of a signal containing 99 periods in the whole pedaling process. We took a coarse-graining processing to each data sequence with different τ ($\tau \in [1, 20]$) (**Figure 7**).

In **Figure 7**, the relationships between the MSESEn, the time, and the scale τ show that the MSESEn is low for τ in the range [1, 3] and [15, 20]. The MSESEn is the largest when $\tau = 8$. It means that this scale has the largest correlation with the sEMG signals. Picking out the relationship between the MSESEn and the time for $\tau = 8$ gives **Figure 8D**. It shows that the MSESEn decreases from 0.7565 to 0.1038 along time. This shows a clear declining trend.

Then we calculated the MSESEn of the original signal for $\tau = 8$ in **Figure 3A**. The distribution of the MSESEn (**Figure 8A**) is scattered and is hard to judge the tendency. The MSESEn of the original signal does not show a distinctive feature. However, if we calculated the MSESEn of the reconstructed signals in **Figures 4C**, **5C**, **6C**, we can easily observe the correlation of the data and the trend of the muscle fatigue from the distributions (**Figures 8B–D**). Moreover, we can observe that the regularity and the decreasing tendency of the MSESEn in **Figure 8D** is more obvious than that in **Figures 8B,C**. MSESEn clearly shows the changing process of the muscle from fresh to fatigue. It has the advantages of high concentricity, good monotonicity, and relative consistency. These results also show that OCEEMD is an effective preprocessing method for extracting the MSESEn.

Then, we quantitatively analyzed two other indicators, ApEn and SampEn, and compared them with the MSESEn for the reconstructed sEMG signal processed by the OCEEMD. **Figures 9A–C** show the distributions and the best fitting straight lines for the three indicators. A comparison of them shows that MSESEn in **Figure 9C** declines more significantly and its distribution is more concentrated than ApEn in

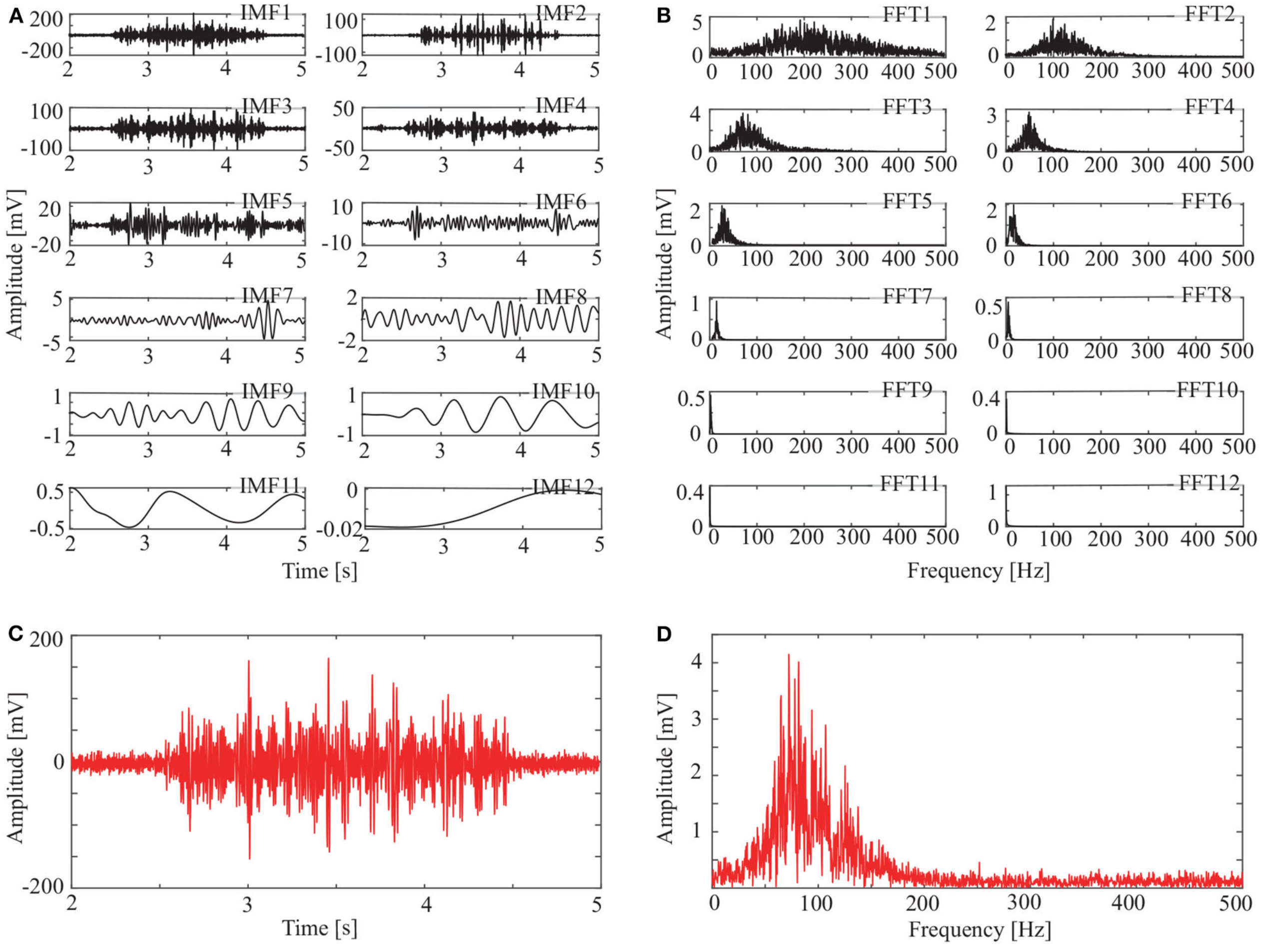


FIGURE 5 | IMFs and spectrums using CEEMDAN: **(A)** IMFs of signal in **Figure 3C**, **(B)** corresponding FFTs, **(C)** reconstructed signal, and **(D)** spectrum of reconstructed signal.

Figure 9A and SampEn in **Figure 9B** during the development of muscle fatigue.

Table 3 lists the gradients and intercepts of the trend lines in **Figures 9A–C**. The absolute gradient of MSESEn is 0.0051. It is 1.5 times larger than that of ApEn and more than 1.3 times larger than that of SampEn. The intercept of MSESEn is 0.57. It is larger than those of the ApEn and SampEn. These values figure that, among these three indicators, the MSESEn dropped the most, suggesting the MSESEn are the most sensitive to the generation of muscle fatigue. Furthermore, **Figure 9D** shows a box diagram of the three indicators. Clearly, the MSESEn has the least data dispersion and thus the best robustness among the three indicators.

3.4. Statistical Analysis

The goodness of fit, R^2, was used to test 40 sEMG signals (four muscles for each of the ten subjects). It measured whether our method effectively extracted the features of the sEMG signals. It is defined to be

$$R^2 = \frac{\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} y_i^2}, \tag{25}$$

where y_i is the ith extracted entropy and $\hat{y}_i$ is the corresponding value on the fitting line. We calculated R^2 with the three preprocessing methods and three feature extraction methods. The average of R^2, $\overline{R^2}$, with different methods is shown in **Table 4**.

It is clear from **Table 4** that, compared with the CEEMD and the CEEMDAN, the OCEEMD yields the largest $\overline{R^2}$ for three different definitions of entropy. This indicates that the OCEEMD extracts the entropy more effectively than the other two methods do. Moreover, the MSESEn, which is the new one used in this study, has the largest $\overline{R^2}$ than the ApEn and the SampEn. This shows the advantage of using this entropy in this study.

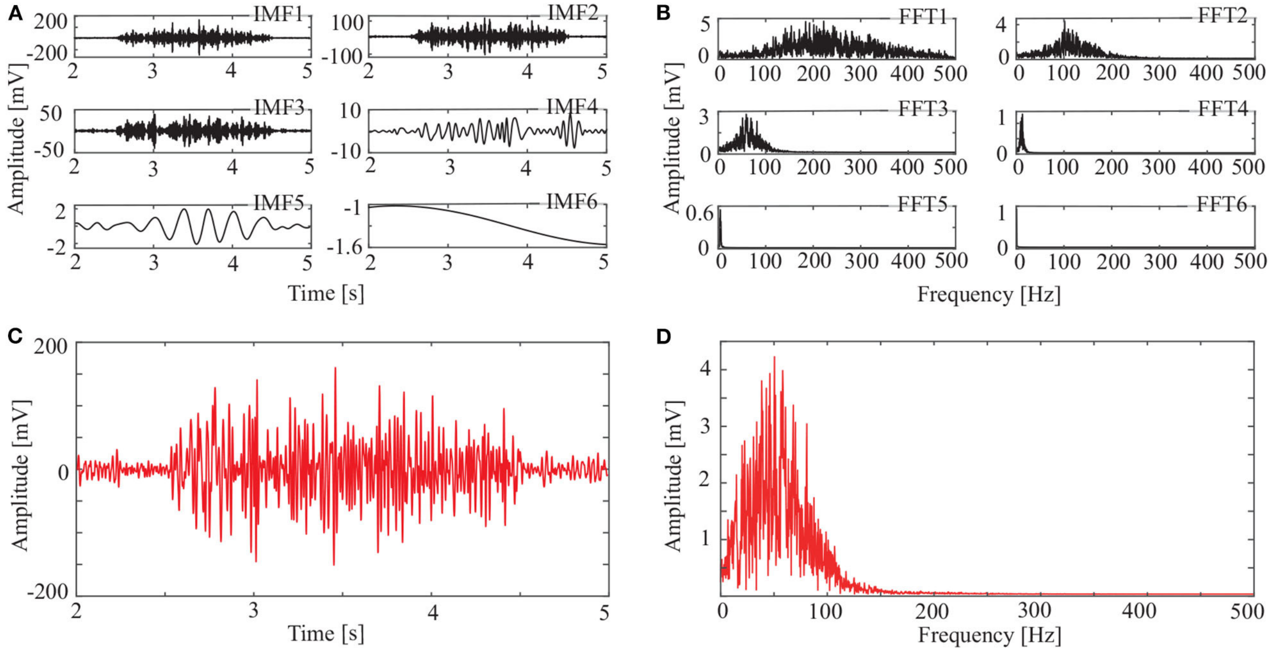


FIGURE 6 | IMFs and spectrums using OCEEMD: **(A)** IMFs of signal in **Figure 3C**, **(B)** corresponding FFTs, **(C)** reconstructed signal, and **(D)** spectrum of reconstructed signal.

TABLE 2 | Comparison of CEEMD, CEEMDAN, and OCEEMD.

	CEEMD	CEEMDAN	OCEEMD
RMSE	4.5347×10^{-2}	1.3275×10^{-3}	2.8986×10^{-4}
No. of IMFs	10	12	6
d	±0.27	0.32 0.37 0.28 0.22 0.21 0.16 0.11 0.11 0.09 0.07 0.03	±0.32 ±0.26 ±0.18 ±0.12 ±0.12
LSMI of adjacent IMFs	0.5751 0.5826 0.3529 0.2529 0.1721 0.0819 0.0525 0.0427	0.7238 0.6147 0.4629 0.4588 0.3025 0.3126 0.1196 0.1064 0.1012 0.0828	0.0911 0.0866 0.0562 0.0104

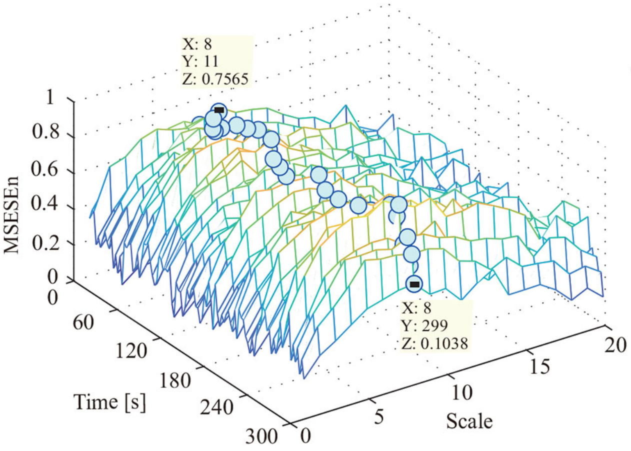


FIGURE 7 | MSESEn of reconstructed signal using OCEEMD with different scales.

4. DISCUSSION

The non-linear methods were explored to process the sEMG signals. **Figures 4**, **5**, **6**, and **Table 2** show that the CEEMD, the CEEMDAN, and the OCEEMD have their merits. While the number of IMFs is smaller for the CEEMD than for the CEEMDAN, the de-noising effect is better for the CEEMDAN than for the CEEMD. The features closely relate to their decomposition processes as well as the intermittent and spectral characteristics of sEMG signals. The CEEMD adds *N* sets of Gauss white noise with the same amplitudes but opposite signs to an original sEMG signal, and separately decomposes the 2*N* new signals into mode components (Yeh et al., 2010). Since there may have asymmetric mode components in sEMG signals, some

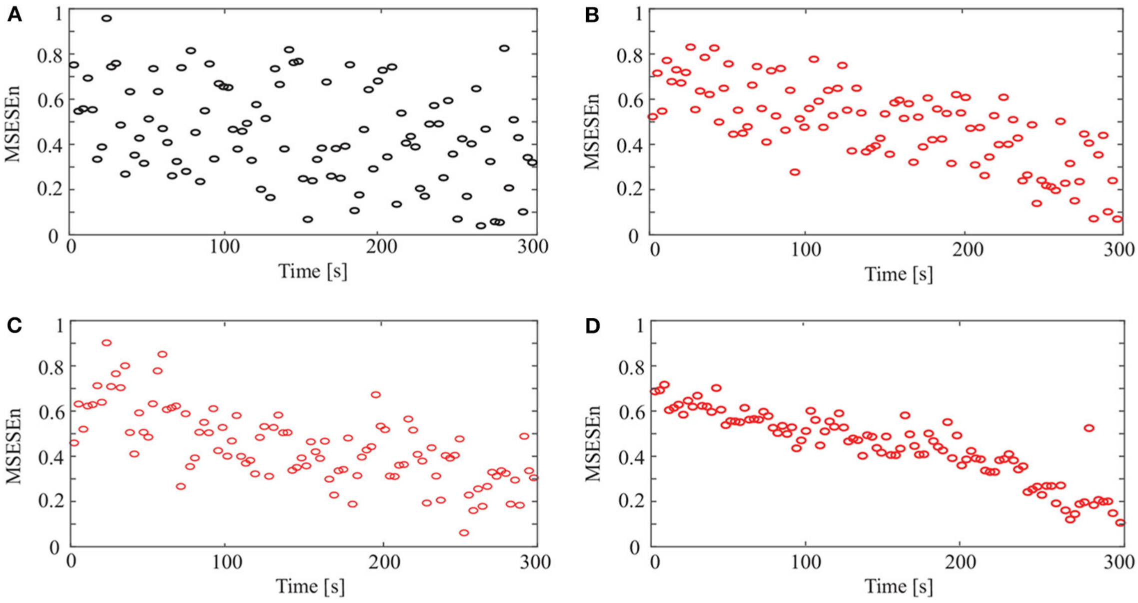


FIGURE 8 | MSESEn of original and reconstructed signals for different methods: **(A)** original signal, **(B)** CEEMD, **(C)** CEEMDAN, and **(D)** OCEEMD.

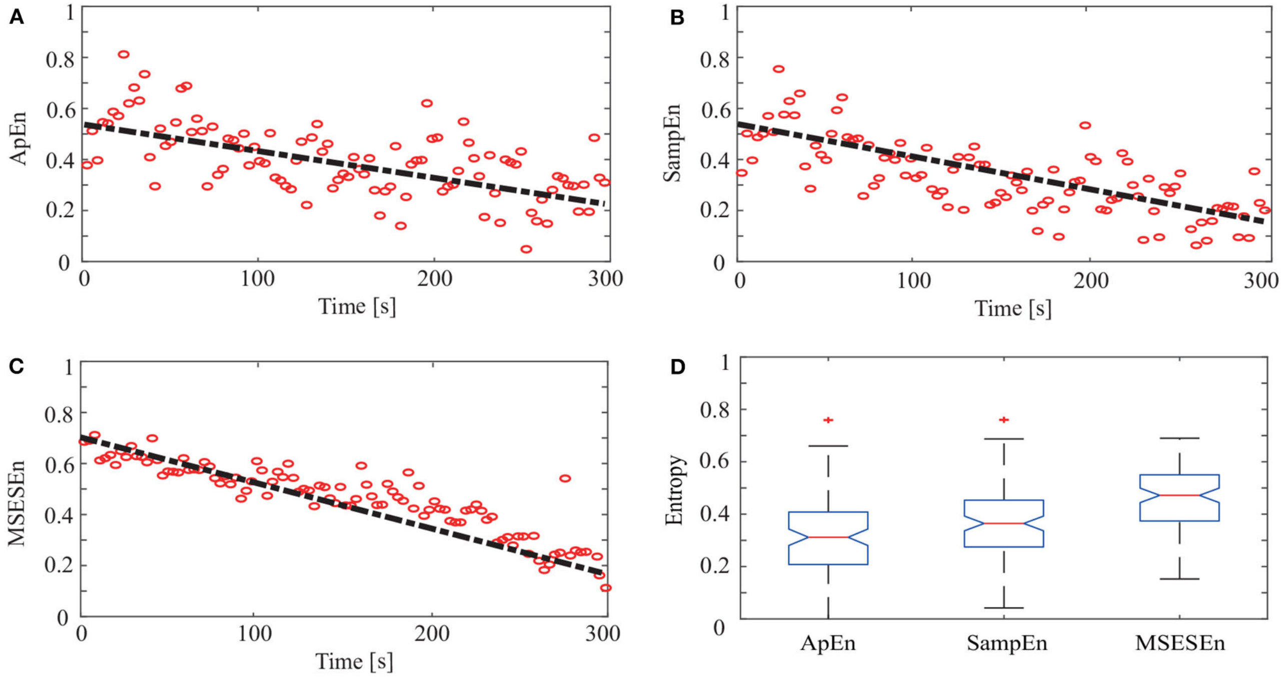


FIGURE 9 | ApEn, SampEn, and MSESEn of OCEEMD: **(A)** ApEn (gradient: −0.0034), **(B)** SampEn (gradient: −0.0038), **(C)** MSESEn (gradient: −0.0051), and **(D)** box diagram.

noise may remain in the signals when summing up and averaging the components. The CEEMDAN adaptively adds white noise to an original signal and yields the IMFs by obtaining a unique residual at each decomposition level (Torres et al., 2011). This method ensures that the noise does not transfer from the present decomposition level to the next level, but the computational expense is high and this method has some dummy IMFs.

Then, the OCEEMD further improves the performance. For example, the number of IMFs for the OCEEMD is 38% less than that for the CEEMD. This indicates the OCEEMD has high

TABLE 3 | Gradients and intercepts for different indicators in **Figure 8**.

	Gradient	Intercept
ApEn	−0.0034	0.32
SampEn	−0.0038	0.39
MSESEn	−0.0051	0.57

TABLE 4 | $\overline{R^2}$ with different methods.

	CEEMD	CEEMDAN	OCEEMD
ApEn	0.8408	0.8613	0.8996
SampEn	0.8914	0.9039	0.9372
MSESEn	0.9423	0.9771	0.9806

decomposition efficiency. It might be due to that the OCEEMD embedded the CQPSO and thus quickly found the optimal solution of d. However, the d at the first level is larger. The reason is that the standard deviation of the reconstructed sEMG signal after using the OCEEMD is very small, which indicates the reconstructed sEMG signal after using the OCEEMD was de-noised very well. Other values of d are smaller, validating the adaptability of the OCEEMD. For the LSMI, the correlations between each IMFs are reduced by more than 80%. The results are inseparable from the threshold criterion of the LSMI on each level. The crucial step was only IMFs with very small LSMI could be further decomposed to the next level. So, every adjacent IMFs were irrelevant finally. Thus, there was no mode mixing in the whole decomposition.

Figure 8 and **Table 4** show that the OCEEMD is superior to the CEEMD and the CEEMDAN. The MSESEn reveals the trend of muscle fatigue. The trend can hardly be observed from the MSESEn in **Figure 8A** because the original signal was not preprocessed. Using the CEEMD, the CEEMDAN, and the OCEEMD to preprocess the signal explicitly displays the trend. As shown in **Figures 8B–D**, the MSESEn in **Figures 8B–D** decreases over time. This indicates that these methods are suitable for processing non-stationary sEMG signals. In particular, the MSESEn for the OCEEMD is well-lumped to show the trend (**Figure 8D** and **Table 4**). This is because that the OCEEMD greatly reduces the mode mixing and residual noise of IMFs.

The results in **Figures 8B–D** shows high consistency with the conclusions in (Pincus, 2006; Pethick et al., 2019). The MSESEn, like ApEn and SampEn, reflects the complexity and the power distribution of sEMG signals in a frequency range (Pincus, 2006; Pethick et al., 2019). When a muscle begins contraction during physical activity, the muscle fibers are activated and show disordered discharge, producing a signal with widely distributed power. At this point, the components of the sEMG signal are complex, and the ratio of the power to total power is large. Thus, the MSESEn is large. After a period of muscle contraction, some muscle fibers are "tired" and the main muscle fibers still participate in the activity. So, the power decreases. It is clear and simple to detect the main components of the signal. Thereby, the MSESEn becomes small. Thus, the decrease of the MSESEn reveals the degree of muscle fatigue.

Compared to the ApEn and the SampEn, the MSESEn has better centrality and robustness. **Table 3** shows that the gradient of the MSESEn increased by 33% compared with the ApEn and 25% compared with the SampEn. **Tables 3**, **4** also show that the MSESEn has the best fitting effect. There are two reasons. First, the OCEEMD decomposes the sEMG signals according to their envelope spectrum characteristics. The MSESEn, which takes into consideration "envelope spectrum entropy," is sensitive to the changing trend in sEMG signals. Second, sEMG signals have unknown potential sequence patterns and related time scales. The MSESEn, which uses multiple scales, provides additional spatial statistics. So, the MSESEn decreases the most, suggesting that it has the greatest correlation with muscle fatigue levels, and thus it best reflects the changes of muscle fatigue. Furthermore, this obvious downward trend indicates that it has good anti-interference ability and facilitates the feature classification of muscle fatigue in the next step. Therefore, MSESEn is a feasible feature for analyzing muscle fatigue based on sEMG.

Overall, the OCEEMD, which integrates the LSMI and CQPSO, is used to decompose sEMG signals to obtain some IMFs with no mode mixing. This algorithm, along with the next signal reconstruction by useful IMFs, aims to de-noise and purify the signals. Then, the MSESEn of the reconstructed sEMG is calculated to detect the process of muscle fatigue. The test on the decomposition and reconstruction demonstrates that the OCEEMD effectively suppressed mode mixing between IMFs with fast iteration. Experiments on the extracted envelope illustrate that the MSESEn displays muscle fatigue clearly. These results show that our method has the potential to process sEMG signals and measure muscle fatigue.

In this study, we devised the OCEEMD and MSESEn to capture the features of sEMG signals reflecting the changing trend of muscle fatigue. A 5-min pedaling experiment was designed to record the sEMG signals and to verify the effectiveness of our method. On the other hand, experiments for different conditions, such as time duration and pedaling load, may provide us a more comprehensive understanding of the relationship between sEMG signals and muscle fatigue. They will be carried out in the future.

DATA AVAILABILITY STATEMENT

The raw data were recorded at the Advanced Mechatronics Laboratory, School of Engineering, Tokyo University of Technology, Japan. The data supporting the findings of this study are available from zhaojuan0859@cug.edu.cn or she@stf.teu.ac.jp on request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The ethical committee of Tokyo University of Technology. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JZ and JS conceived this study. JZ and KP performed the experiments and wrote the manuscript. JS, EF, DW, and MW contributed to the methodology. JZ, JS, and KP revised the manuscript. All the co-authors agreed with the present version.

FUNDING

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB1300900; the National Natural Science Foundation of China under Grant 61873348; the Natural Science Foundation of Hubei Province, China, under Grant 2020CFA031; Wuhan Applied Foundational Frontier Project under Grant 2020010601012175; the 111 Project, China, under Grant B17040; and JSPS KAKENHI, Japan, under Grant 20H04566.

ACKNOWLEDGMENTS

The authors would like to thank the research team at the Advanced Mechatronics Laboratory, School of Engineering, Tokyo University of Technology, Japan, and Mr. Ruoyu Jiang, Mr. Yujian Zhou, Mr. Wangyang Ge, Mr. Qicheng Mei, and Mr. Zewen Wang at School of Automation, China University of Geosciences, Wuhan, China, for their contributions to this project.

REFERENCES

Andrade, A. O., Nasuto, S. J., Kyberd, P. J., Sweeney-Reed, C. M., and Van-Kanijn, F. R. (2006). EMG signal filtering based on empirical mode decomposition. *Biomed. Signal Process. Control* 1, 44–55. doi: 10.1016/j.bspc.2006.03.003

Cashaback, J. G., Cluff, T., and Potvin, J. R. (2013). Muscle fatigue and contraction intensity modulates the complexity of surface electromyography. *J. Electromyogr. Kinesiol*. 23, 78–83. doi: 10.1016/j.jelekin.2012.08.004

Chen, X., and Wang, W. (2017). Extracting and compensating for fog vibration error based on improved empirical mode decomposition with masking signal. *Appl. Opt*. 56:3848–3856. doi: 10.1364/AO.56.003848

Chowdhury, S. K., Nimbarte, A. D., Jaridi, M., and Creese, R. C. (2013). Discrete wavelet transform analysis of surface electromyography for the fatigue assessment of neck and shoulder muscles. *J. Electromyogr. Kines*. 23, 995–1003. doi: 10.1016/j.jelekin.2013.05.001

Costa, M. D., Peng, C. K., and Goldberger, A. L. (2008). Multiscale analysis of heart rate dynamics: entropy and time irreversibility measures. *Cardiovasc. Eng*. 8, 88–93. doi: 10.1007/s10558-007-9049-1

Gandevia, S. C. (2013). Spinal and supraspinal factors in human muscle fatigue. *Physiol. Rev*. 81, 10127–10134. doi: 10.1152/physrev.2001.81.4.1725

Gao, J., Hu, J., Liu, F., and Cao, Y. (2015). Multiscale entropy analysis of biological signals: a fundamental bi-scaling law. *Front. Comput. Neurosci*. 2015, 2039–2049. doi: 10.3389/fncom.2015.00064

Hu, X., Peng, S., and WenLiang, H. (2012). EMD revisited: a new understanding of the envelope and resolving the mode-mixing problem in AM-FM signals. *IEEE Trans. Signal Process*. 60, 1075–1086. doi: 10.1109/TSP.2011.2179650

Huang, M. L. (2016). Hybridization of chaotic quantum particle swarm optimization with SVR in electric demand forecasting. *Energies* 9, 426–441. doi: 10.3390/en9060426

Huang, N. E. (2000). New method for nonlinear and nonstationary time series analysis: empirical mode decomposition and Hilbert spectral analysis. *Proc. SPIE Int. Soc. Opt. Eng*. 4056, 197–209. doi: 10.1117/12.381681

Kimura, M., and Sugiyama, M. (2011). Dependence-maximization clustering with least-squares mutual information. *J. Adv. Comput. Intell. Intell. Informat*. 15, 800–805. doi: 10.20965/jaciii.2011.p0800

Kyranou, I., Vijayakumar, S., and Erden, M. S. (2018). Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses. *Front. Neurorobot*. 12:58. doi: 10.3389/fnbot.2018.00058

Li, T., Qian, Z., and He, T. (2020). Short-term load forecasting with improved CEEMDAN and GWO-based multiple kernel ELM. *Complexity* 2020:1209547. doi: 10.1155/2020/1209547

Liu, S., Chang, K., and Cheng, D. (2014). The progression of muscle fatigue during exercise estimation with the aid of high-frequency component parameters derived from ensemble empirical mode decomposition. *IEEE J. Biomed. Health Informatics* 18, 1647–1658. doi: 10.1109/JBHI.2013.2286408

Lu, Y., Xie, R., and Liang, S. (2019). CEEMD-assisted bearing degradation assessment using tight clustering. *Intern. J. Adv. Manuf. Technol*. 104, 1259–1267. doi: 10.1007/s00170-019-04078-2

Lv, Y., Q., Z., and R., Y. (2015). Fault diagnosis of rolling bearing based on fast nonlocal means and envelop spectrum. *Cardiovasc. Eng*. 15, 1182–1198. doi: 10.3390/s150101182

Pethick, J., Winter, S. L., and Burnley, M. (2019). Fatigue reduces the complexity of knee extensor torque during fatiguing sustained isometric contractions. *Eur. J. Sport. Sci*. 19, 1349–1358. doi: 10.1080/17461391.2019.1599450

Pilkar, R., Ramanujam, A., Yarossi, M., Rajagopalan, V., and Bayra, M. B. (2017). Application of empirical mode decomposition combined with notch filtering for interpretation of surface electromyograms during functional electrical stimulation. *IEEE Trans. Neural System Rehabil. Eng.g*. 25, 1268–1277. doi: 10.1109/TNSRE.2016.2624763

Pincus, S. M. (2006). Approximate entropy as a measure of irregularity for psychiatric serial metrics. *Bipolar Disord*. 8, 430–440. doi: 10.1111/j.1399-5618.2006.00375.x

Rezaie-Balf, M., Maleki, N., Kim, S., Ashrafian, A., Babaie-Miri, F., Kim, N. M., et al. (2019). Forecasting daily solar radiation using CEEMDAN decomposition-based MARS model trained by crow search algorithm. *Energies* 12:1416. doi: 10.3390/en12081416

She, J., Ohyama, Y., Wu, M., and Hashimoto, H. (2017a). Development of electric cart for improving walking ability-application of control theory to assistive technology. *Sci. China Inf. Sci*. 60:123201. doi: 10.1007/s11432-017-9261-1

She, J., Wu, F., Mita, T., Hashimoto, H., Wu, M., and Iliyasu, A. M. (2017b). Design of a new lower-limb rehabilitation machine. *J. Adv. Comput. Intell. Intell. Informat*. 21, 409–416. doi: 10.20965/jaciii.2017.p0409

Srhoj-Egekher, V., Cifrek, M., and Medved, V. (2011). The application of Hilbert-Huang transform in the analysis of muscle fatigue during cyclic dynamic contractions. *Med. Biol. Eng. Computing* 49, 659–669. doi: 10.1007/s11517-010-0718-7

Talebinejad, M., Chan, A. D., and Miri, A. (2011). A Lempel-Ziv complexity measure for muscle fatigue estimation. *J. Electromyogr. Kinesiol*. 21, 236–241. doi: 10.1016/j.jelekin.2010.12.003

Torres, M. E., Colominas, M. A., Schlotthauer, G., and Flandrin, P. (2011). "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (Prague), 4144–4147. doi: 10.1109/ICASSP.2011.5947265

Valdez, M. A., Jaschke, D. L., and Vargas, L. D. (2017). Quantifying complexity in quantum phase transitions via mutual information complex networks. *Phys. Rev. Lett*. 119:225301. doi: 10.1103/PhysRevLett.119.225301

Wu, Q., Xi, C., Ding, L., Wei, C., Ren, H., Law, R., et al. (2017). Classification of EMG signals by BFA-optimized GSVCM for diagnosis of fatigue status. *IEEE Trans. Auto. Sci*. 14, 915–930. doi: 10.1109/TASE.2016.2564419

Wu, Z., and Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Data Anal*. 1, 1–41. doi: 10.1142/S1793536909000047

Xie, H., Guo, J., and Zheng, Y. (2019). Fuzzy approximate entropy analysis of chaotic and natural complex systems: detecting muscle fatigue using electromyography signals. *Ann. Biomed. Eng.* 38, 1483–1496. doi: 10.1007/s10439-010-9933-5

Yeh, J. R., Shieh, J. S., and Huang, N. E. (2010). Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. *Adv. Data Anal.* 2, 135–156. doi: 10.1142/S1793536910000422

Zhang, Q., Liu, R., Chen, W., and Xiong, C. (2017). Simultaneous and continuous estimation of shoulder and elbow kinematics from surface EMG signals. *Front. Neurosci.* 11:280. doi: 10.3389/fnins.2017.00280

Zhang, X., and Zhou, P. (2013). Filtering of surface EMG using ensemble empirical mode decomposition. *Med. Eng. Phys.* 35, 537–542. doi: 10.1016/j.medengphy.2012.10.009

Zhang, Y., Su, S. W., Xu, P., and Yao, D. (2017). "Performance evaluation of noise-assisted multivariate empirical mode decomposition and its application to multichannel EMG signals," in *The 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Jeju Island), 3457–3460. doi: 10.1109/EMBC.2017.8037600

Zhao, J., She, J., Wang, D., Fukushima, E. F., Wu, M., and Hashimoto, H. (2018). "Evaluation of lower-limb rehabilitation based on muscle fatigue," in *The 27th IEEE International Symposium on Industrial Electronics (ISIE)* (Cairns), 1106–1111. doi: 10.1109/ISIE.2018.8433835

Zhao, L., Yu, W., and Yan, R. (2014). Rolling bearing fault diagnosis based on CEEMD and time series modeling. *Math. Probl. Eng.* 2014:101867. doi: 10.1155/2014/101867

Zheng, J., Cheng, J., and Yang, Y. (2014). Partly ensemble empirical mode decomposition: an improved noise-assisted method for eliminating mode mixing. *Signal Process.* 96(Pt. B), 362–374. doi: 10.1016/j.sigpro.2013.09.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurorobotics

ORIGINAL RESEARCH
published: 12 November 2020
doi: 10.3389/fnbot.2020.599889

Phase-Synchronized Learning of Periodic Compliant Movement Primitives (P-CMPs)

*Tadej Petrič**

CoBoTaT Lab, Department of Automatics, Biocybernetics and Robotics, Jožef Stean Institute (JSI), Ljubljana, Slovenia

Autonomous trajectory and torque profile synthesis through modulation and generalization require a database of motion with accompanying dynamics, which is typically difficult and time-consuming to obtain. Inspired by adaptive control strategies, this paper presents a novel method for learning and synthesizing Periodic Compliant Movement Primitives (P-CMPs). P-CMPs combine periodic trajectories encoded as Periodic Dynamic Movement Primitives (P-DMPs) with accompanying task-specific Periodic Torque Primitives (P-TPs). The state-of-the-art approach requires to learn TPs for each variation of the task, e.g., modulation of frequency. Comparatively, in this paper, we propose a novel P-TPs framework, which is both frequency and phase-dependent. Thereby, the executed P-CMPs can be easily modulated, and consequently, the learning rate can be improved. Moreover, both the kinematic and the dynamic profiles are parameterized, thus enabling the representation of skills using corresponding parameters. The proposed framework was evaluated on two robot systems, i.e., Kuka LWR-4 and Franka Emika Panda. The evaluation of the proposed approach on a Kuka LWR-4 robot performing a swinging motion and on Franka Emika Panda performing an exercise for elbow rehabilitation shows fast P-CTPs acquisition and accurate and compliant motion in real-world scenarios.

Keywords: adaptive control, autonomous learning, human-robot interaction, internal dynamic model, compliant movement primitives

Edited by:
Shuai Li,
Swansea University, United Kingdom

Reviewed by:
Tomas Kulvicius,
University of Göttingen, Germany
Luca Patanè,
University of Catania, Italy

****Correspondence:***
Tadej Petrič
tadej.petric@ijs.si

Received: *28 August 2020*
Accepted: *07 October 2020*
Published: *12 November 2020*

Citation:
Petrič T (2020) Phase-Synchronized Learning of Periodic Compliant Movement Primitives (P-CMPs). Front. Neurorobot. 14:599889. doi: 10.3389/fnbot.2020.599889

1. INTRODUCTION

Programming by demonstration (PbD) is a typical approach for transferring skills to robots by mirroring human actions (Billard et al., 2008; Argall et al., 2009; Calinon et al., 2014; Peternel et al., 2018). For simple tasks, human demonstrations are typically recorded using vision-based systems (Welschehold et al., 2016) or motion tracking suites (Filippeschi et al., 2017). For a more challenging task where force constraints and compliance strategies are required, the kinesthetic guidance or multi-modal human-in-the-loop skill transfer approaches can be used (Peternel et al., 2014; Rozo et al., 2016). Besides, such learning has the advantage of already being adapted to the kinematic and dynamic parameters of the robotic system. Here, typically a well-established inverse dynamic control approach is in use (Sciavicco and Siciliano, 2012). However, due to the increasing complexity of robot mechanisms and tasks, they are performing, accurate dynamical models' of both the robot and the task are often difficult to obtain. To bridge the gap, machine learning algorithms were adopted because of their ability of learning complex models. Although learning algorithms are powerful enough to learn the inverse dynamics of both the robot and the task (Nguyen-Tuong and Peters, 2011), they still require a large amount of data for the learning processes, which makes them unsuitable for on-line learning of tasks-specific dynamics.

Knowing the exact dynamical model is crucial to achieving compliant robot behavior, which is needed when robots are operating in an unstructured environment. Hence, exact dynamical models of both, the robot and the task makes it possible to either adjust the controller feedback gains to obtain the desirable compliance or to prescribe the desired dynamic behavior (Buchli et al., 2011; Kronander and Billard, 2013; Žlajpah and Petrič, 2019). Skill learning approaches that can expand the database can be time-consuming. For example, reinforcement learning might take a long time to tune the skill because a high number of repetitions is needed (Kober et al., 2012). Such exploitative learning methods were successfully used for learning force profiles for in-contact tasks (Racca et al., 2016). On the other hand, supervised learning methods are typically faster but require a reference for the optimization process (Wang et al., 2009). Nevertheless, even these methods might take too much time to produce a large enough database enabling statistical methods to generate an accurate dynamic model for a given task. However, when using PbD methods, dynamic models of both the robot and the task, are usually not known and can not be easily learned from imitation. Since modeling of system dynamics is typically a difficult and time-consuming task, this work instead addresses the problem of how to obtain the task-specific dynamics through autonomous learning and thereby avoid the need for an expert to define them.

Learning of task-specific dynamics was proposed in Deniša et al. (2016), where Compliant Movement Primitives (CMPs) were introduced. CMPs encode both the kinematic trajectory in the form of Dynamic Movement Primitives (DMP) (Ijspeert et al., 2013) and accompanying dynamics called Torque Primitives (TPs), i.e., joint-torques encoded with weighted radial-basis functions. In Deniša et al. (2016) TPs were obtained through execution of the desired movement trajectories using high-gain feedback control. This limitation was mitigated in Petrič et al. (2018), where TPs were learned iteratively until the error of compliant tracking was reduced below a predefined threshold.

The main contribution of this paper is a two-layered system that combines Phase-synchronized Adaptive Fourier Series (P-AFO) with Periodic Compliant Movement Primitives (P-CMPs). The P-AFO is an incremental improvement of AFO proposed in Petrič et al. (2011), which guaranties unambiguous frequency and phase synchronization to an arbitrary input signal, which is crucial for P-CMPs. Furthermore, the P-CMPs is a periodic extension of CMPs proposed in Deniša et al. (2016) and Petrič et al. (2018). Here the kinematic trajectory is encoded in Periodic Dynamic Movement Primitives (Gams et al., 2009) and the corresponding task-specific dynamics with Periodic Torque Primitives (P-TPs). For the P-TPs we propose a novel combination of weighted kernel functions that are frequency and phase-dependent. The novel P-TPs framework allows direct modulation of frequency, which was not possible before (Deniša et al., 2016). Inspired by human sensorimotor learning (Kawato, 1990), the P-TPs are learned on-line using a feed-back error learning approach. The learning is active until the tracking error of a compliant controller robot is reduced below a predefined threshold.

This paper is organized as follows. In the next section, we describe related work detailing the topics of learning of robot torque profiles and their modulation and generalization. In section 3 we describe the main contributions of this paper, i.e., unambiguous phase synchronization (P-AFO), periodic torque primitives (P-TPs), and the integration of feedback error learning. Results of experimental evaluation on a Kuka LWR-4 robot arm learning to perform a swinging task and evaluation on Franka Emika Panda robot learning to rehabilitate the elbow by a stretching task are presented in section 4. A discussion concludes the paper in section 5.

2. RELATED WORK

2.1. Torque Learning

For accurate and compliant execution of tasks, the task-space dynamics is required (Del Prete and Mansard, 2016), whereby a dynamic model of the task might be hard to obtain (Petrič et al., 2010). To mitigate mathematical modeling, different biologically inspired methods were proposed to enhance robot control (Franklin and Wolpert, 2011). Merging them with robots that have joint-torque sensors led to the development of Compliant movement Primitives (CMPs), first reported in Petrič et al. (2014). Originally, CMPs recorded feed-forward torques during initial task execution with stiff robot behavior that ensures accurate motion tracking. Once torques were recorded, they were used as feed-forward components of the CMPs in the next repetitions of the same task. Since the torque profiles had to be recorded for each variation of the tasks, even for different execution speeds, a statistical generalization method was proposed in Deniša et al. (2016). They showed that generalization between CMPs can successfully be used to generate CMPs for tasks where kinematic or dynamic parameters were changed. Besides generalization, a statistical-graph search was shown to effectively generate new CMPs by joining together different parts of several CMPs (Deniša et al., 2013).

Exploiting the feed-forward torque was also utilized when the possibility of measuring joint torque was available. For example in Calandra et al. (2015) they use tactile sensors to compute joint torques on an iCub humanoid robot. The computed joint-torques were used as feed-forward signals similar to the CMPs framework. Learning of joint torques together with kinematic trajectory was also implemented in Steinmetz et al. (2015), where the recorded torques were used as a feed-forward signal to increase the motion accuracy of the in-contact task. Originally the learning of CMPs torque signals was performed during an exact motion execution, whereby the robot was stiff due to the high feedback gains. As a consequence, the application of CMPs during learning was limited and potentially dangerous when interacting with the environment or humans. To mitigate this issue, an approach using autonomously learning of torque profiles while using compliant robot behavior, i.e., low feed-back gains, was introduced in Petrič et al. (2018). However, the approach was not suitable for periodic tasks.

Other approaches for torque learning not directly related to CMPs were also proposed. Gaussian process regression for on-line learning of the dynamical model was proposed in Nguyen-Tuong and Peters (2011), where the accuracy of the dynamical model was improved while keeping compliant robot behavior. While results were promising, this approach required a large amount of data, hence it was not focused on learning only task-specific torques. For learning only task-specific torques iterative learning control (ILC) was utilized in Schwarz and Behnke (2013). Here ILC was used to identify model parameters for motor and friction models. Similarly, in Kronander et al. (2015) ILC was used to update the dynamical model. Inspired by human sensorimotor learning, Kawato (1990) introduced a feedback error learning approach for learning task-specific dynamics for a given kinematic trajectory. The feedback error learning was later adopted in Gopalan et al. (2013), where it was used to stabilize the controller's output for adapting the gait of an under-actuated bipedal robotic system.

2.2. Modulation and Generalization

Trajectory modulation and generalization is a wide topic that can be considered from different domains of application. Mostly, methods for modulation and generalization were focused on the kinematic trajectory and only a few dealt with dynamics. The modulation and generalization ability of kinematic and dynamic parameters are specifically important for the P-CMPs framework proposed in this paper. The kinematic part of P-CMPs is encoded with P-DMPs, which already allow a certain degree of modulation and generalization. In Gams et al. (2009), the modulation abilities of DMPs to change the goal and the frequency was demonstrated. Furthermore, the DMPs were also used as means to represent results of statistical generalization using locally weighted regression in Ude et al. (2010) and generalization between weights of DMPs using Gaussian process regression (GPR) in Forte et al. (2012). For both approaches, a task parameter is required to generate a new trajectory from a motion database. Similarly, in Stulp et al. (2013) the task parameter was used to learn the weights of DMPs of a single demonstration. Instead of rallying on one demonstration, Matsubara et al. (2011) used several demonstrations to create a parametric attractor landscape in a set of differential equations. Similarly, a variation of DMPs as a Mixture of Motor Primitives (MoMP) was introduced in Mülling et al. (2013), where they proposed an algorithm that can autonomously update the weights. By exploiting the external inputs the on-line modulation and adaptation of DMPs are also possible as shown in Gams et al. (2010) and Kulvicius et al. (2013).

The DMPs are not the only trajectory representation method or even the only dynamical systems used for modulation and generalization. However, because our proposed approach in this paper is composed also of DMPs, other possible alternatives are only briefly listed below. The task-specific Gaussian Mixture Models (TP-GMM) were proposed by Khansari-Zadeh and Billard (2011) and were also used in Calinon (2016). Another possibility is also Hidden Markov Models (HMM) that were used in Lee and Ott (2011). While these approaches rely on generating trajectories based on existing database entries, the trajectory generation based on extrapolation and the database expansion is still an open research topic. Extrapolation was mentioned in Calinon et al. (2013), where statistical methods were used to encode the movements. The algorithm for autonomous database expansion was proposed in Petrič et al. (2018), where the new compliant motion trajectories were generated also by extrapolating the database. The literature related to modulation and generalization of dynamic parameters is even more sparse. Besides already mentioned (Calinon et al., 2013; Deniša et al., 2013; Petrič et al., 2018), modulation, and generalization of dynamic parameters, such as forces and torques, was also researched in Gams et al. (2015), where a statistical generalization was used on force-based coupling terms. However, their approach was limited, since it requires user interaction to generate new database entries. Similarly, in Koropouli et al. (2015) a new policy was proposed where the input was motion data and the output was a force.

The generalization of both kinematic trajectories and torque profiles has been reported with the aforementioned CMPs in Deniša et al. (2013) and later extend with an approach enabling autonomous learning in Petrič et al. (2018). Our paper extends the approach in Deniša et al. (2013) and Petrič et al. (2010) first by introducing the Periodic-CMPs framework and second by proposing novel P-TPs formulation which includes frequency modulation capabilities.

3. PERIODIC COMPLIANT MOVEMENT PRIMITIVES

The inspiration for the P-CMPs multi-layered framework has been taken from the two-layered imitation system reported in Gams et al. (2009) and Petrič et al. (2011). In their work, the authors introduced a system that can be used for imitation learning, because it allows autonomous frequency adaption and learning of kinematic trajectories. The extension of kinematic trajectory with corresponding dynamic parameters in the form of P-CMPs is proposed in this paper and illustrated in **Figure 1**.

Periodic Compliant Movement Primitives (P-CMPs) $h(\Omega, \phi)$ are defined as a combination of the adaptive oscillators (P-AFO), kinematic trajectories encoded in periodic Dynamic Movement Primitives (P-DMPs) and corresponding task-specific dynamics encoded in Periodic Torque Primitives (P-TPs).

$$h(\Omega, \phi) = [p_d(\Omega, \phi); \dot{p}_d(\Omega, \phi); \ddot{p}_d(\Omega, \phi); \tau_f(\Omega, \phi)]. \quad (1)$$

Here Ω and ϕ are the desired motion frequency and phase, respectively. $\ddot{p}_d(\Omega, \phi)$, $p_d(\Omega, \phi)$ and $\dot{p}_d(\Omega, \phi)$ are the desired acceleration, velocity and position trajectories, respectively, encoded within P-DMPs. $\tau_f(\Omega, \phi)$ are the corresponding joint torques encoded in P-TPs.

Similar to the discrete CMPs, the two-stage process is used to obtain the P-CMPs. First, the kinematic motion trajectories are obtained typically by imitation learning (Gams et al., 2009). Next, the corresponding periodic torques are obtained using the feedback error learning approach (Kawato, 1990).

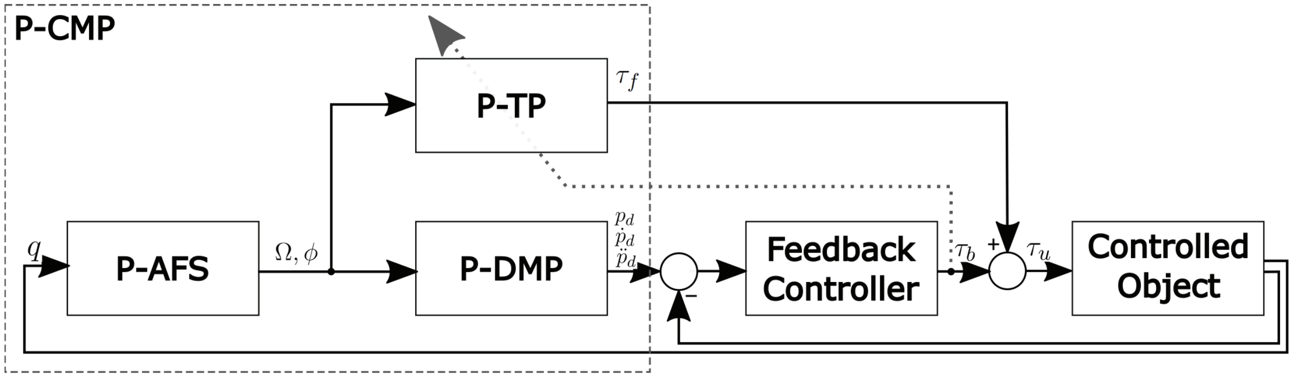


FIGURE 1 | The multi-layered structure of the control system based on P-CMPs. The input q is a measured quantity and the output is the desired kinematic trajectory p_d; $\dot{p}_d$; $\ddot{p}_d$ and the corresponding torque profiles τ_f. Note that the system can work in parallel for an arbitrary number of dimensions.

3.1. Phase-Synchronization

The adaptive phase oscillator with the adaptive Fourier series was originally proposed in Petrič et al. (2011). The core is a second-order system of differential equations governed by

$$\dot{\phi} = \Omega - \kappa\epsilon \sin(\phi), \tag{2}$$

$$\dot{\Omega} = -\kappa\epsilon \sin(\phi). \tag{3}$$

Here Ω is the estimated motion frequency, κ is the coupling strength, ϕ is the corresponding phase and ϵ is governed by

$$\epsilon = q - \hat{q}, \tag{4}$$

where q is the input signal and $\hat{q}$ is the estimated oscillator feedback. It is governed by

$$\hat{q} = \sum_{i=0}^{m} A_i \cos(i\phi) + \sum_{j=2}^{m} B_j \sin(j\phi) \tag{5}$$

where m denotes the size of the modified Fourier series. Note that in this paper the sinus part of the series starts with index $j = 2$, which ensures unambiguously defined phase synchronization. The parameters A_i and B_j are updated as in Petrič et al. (2011).

$$\dot{A}_i = \eta\epsilon \cos(i\phi),\ i \in [0, 1, ..., m], \tag{6}$$

$$\dot{B}_j = \eta\epsilon \sin(j\phi),\ j \in [2, 3, ..., m], \tag{7}$$

where η is the parameter update rate. By skipping the first parameter of the sinusoidal part of the Fourier series, i.e., $j = 1$, we ensure the phase is always well-defined with respect to the input signal q. This is because the main frequency component and corresponding phase is now only related to cosine and not from a combination of cosine and sine as it was in the original system. The novel system is denoted with Phase-synchronized Adaptive Fourier Series (P-AFO). The evaluation results, focused on the novelty, i.e., phase synchronization, are shown in section 4.1.

3.2. Motion Trajectories

The second layer ensures the proper waveform of the kinematic trajectories. It is encoded by P-DMPS, which are anchored to the phase signal ϕ of the adaptive oscillator as in Petrič et al. (2011). The equations for a single-degree-of-freedom are summarized from Ijspeert et al. (2013). The second-order dynamic system of P-DMP is governed by

$$\dot{z} = \Omega\left(\alpha_z\left(\beta_z(g - y) - z\right) + f\right), \tag{8}$$

$$\dot{y} = \Omega z, \tag{9}$$

where α_z and β_z are the positive constants, which guarantee that the system monotonically converges, g is the center of oscillation, and f is the non-linear forcing term that determines the shape of the trajectory. It is given by

$$f(\phi) = \frac{\sum_{i=1}^{N} w_i \psi_i(\phi)}{\sum_{i=1}^{N} \psi_i(\phi)}. \tag{10}$$

Here w is the vector determining the shape and ψ are the Gaussian-like kernel functions given by

$$\psi_i(\phi) = \exp\left(h\left(\cos\left(\phi - c_i\right) - 1\right)\right), \tag{11}$$

where N is the number of kernels, h are the kernels width and c_i is their distribution concerning the phase. Typically they are spread equally between 0 and 2π.

To learn the shape of the trajectory different methods where proposed. When data is available upfront, a batch regression can be used as in Ude et al. (2010). Alternatively, when learning on-line, recursive locally weighted regression can be used. The equations summarizing online learning for the incremental learning approach are summarized from Gams et al. (2009). By rewriting Equations (8) and (9) as one second-order differential equation we get

$$f_d = \frac{\ddot{y}_d}{\Omega^2} - \alpha_z\left(\beta_z(g - y_d) - \frac{\dot{y}_d}{\Omega}\right). \tag{12}$$

Here the triplet of y_d, $\dot{y}_d$ and $\ddot{y}_d$ denotes the desired position, the velocity and the acceleration. To update the weights w_i of the kernel function ψ_i, we use the flowing recursive least-squares method.

$$P_i(t+1) = \frac{1}{\lambda}\left(P_i(t) - \frac{P_i(t)^2}{\frac{\lambda}{\psi_i(\phi(t))} + P_i(t)}\right), \tag{13}$$

$$w_i(t+1) = w_i(t) + \psi_i(\phi(t))P_i(t+1)e_r(t), \tag{14}$$

$$e_r(t) = f_d(t) - w_i(t). \tag{15}$$

The regression typically starts with $w_i = 0$ and $P_i = 0$. Note that P_i is the inverse covariance. λ is the forgetting factor.

Essentially the combination of P-AFO and P-DMP ensures robustness against perturbations and allows frequency modulation of the trajectory. Especially frequency modulation is crucial when performing human-robot cooperative tasks.

3.3. Torque Trajectories

The third layer encodes the corresponding torque trajectories $\tau_f(\Omega, \phi)$ and it is denoted by P-TPs. Note that torques are task-specific, which means they are dependent on the dynamic properties of the task including the execution speed, e.g., frequency. Therefore we propose that P-TPs $\tau_f(\Omega, \phi)$ are both, phase ϕ and frequency Ω dependent. They are governed by

$$\tau_f(\Omega, \phi) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{K} \nu_{i,j}\psi_i(\phi)\Psi_j(\Omega)}{\sum_{i=1}^{M}\sum_{j=1}^{K} \psi_i(\phi)\Psi_j(\Omega)} \tag{16}$$

where ν is a $M \times K$ matrix that encodes the torque profiles and ψ and Ψ are the Gaussian like kernel functions given by

$$\psi_i(\phi) = \exp\left(h^{\phi}\left(\cos\left(\phi - c_i^{\phi}\right) - 1\right)\right), \tag{17}$$

$$\Psi_j(\Omega) = \exp(-h^{\Omega}(\Omega - c_j^{\Omega})^2). \tag{18}$$

Here, h^{ϕ} are the width of the kernel and c_i^{ϕ} is their distribution concerning the phase spread equally between 0 and 2π. h^{Ω} are the kernels width and c_j^{Ω} is their distribution concerning the frequency. Typically c_j^{Ω} is equal between 0 and 4π. Note that M is the number of phase kernels, and K is the number of frequency kernels.

The P-TPs are learned on-line while executing the encoded DMP motion with low gain impedance control using the following law

$$\tau_u = \tau_b + \tau_f, \tag{19}$$

$$\tau_b = K_p e + K_d \dot{e} + K_i \ddot{e} \tag{20}$$

Here, e, $\dot{e}$, and $\ddot{e}$ are the differences between desired p_d, $\dot{p}_d$, and $\ddot{p}_d$ and actual position p, velocity $\dot{p}$, and acceleration $\ddot{p}$, respectively. K_p, K_d, and K_i are the constants selected to ensure robot behaves compliantly, i.e., set to match the low impedance control requirements.

To learn task-specific torque profiles, we used the feedback error learning approach (Nakanishi and Schaal, 2004). It is governed by

$$\dot{\nu}_{i,j} = \iota\tau_b, \tag{21}$$

where ι is a positive constant determining the rate of learning. Note that stability analysis was given in Nakanishi and Schaal (2004).

Because the torques are updated on-line, the task performance, i.e., tracking accuracy, improves over time even if the feedback gains are low. The main idea used in the proposed P-CMPs framework approach is to assure the nominal behavior of the robot for the given periodic task even if compliant robot control is used, i.e., using low feedback gains. In this way, we can assure both, the good tracking accuracy and the compliant behavior. This increases safety aspects for robots working in an unstructured environment or with humans.

4. EXPERIMENTAL VALIDATION

In this section we describe the simulations used to compare the P-AFO phase and frequency synchronization performance with the original AFO (Petrič et al., 2011); and two examples of P-CMPs applications with real-world robots, i.e., Kuka LWR-4 and Franka Emika Panda. Note that stability proofs of CMPs system and the AFO systems were already shown in the above-mentioned research (Nakanishi and Schaal, 2004; Petrič et al., 2011, 2018; Deniša et al., 2016). We therefore focused the evaluation on the system improvements and innovations.

4.1. P-AFO Evaluation

In this numerical simulation example, we compare the phase and frequency synchronization abilities of the original AFO system with the proposed P-AFO system. Note that in both cases the adaptation is done without any signal processing since the entire process of frequency and phase synchronization is completely embedded in the dynamics of the oscillator. In the following example we used for both, AFO and P-AFO, the flowing parameters: $\kappa = 20$, $\mu = 2$, $m = 4$, and $A_i(0) = B_j(0) = 0.5$. The input q was a sinusoidal signal with a frequency of 1 Hz.

Frequency and phase adaption results are illustrated in **Figure 2**. We can see in the top plot that there is no difference between AFO and P-AFO systems performance regarding the adaptation toward the input signal. Similarly, we can also see in the bottom plot that there is no difference between AFO and P-AFO in frequency extraction performance. This shows and confirms that both systems can correctly adapt to extract the frequency of the input signal. However, the crucial difference is in the ability to unambiguously extract the phase signal from the input signal. Clearly, the original AFO signal can extract the phase, which is synchronized to the input signal. However, due to the sum of the first sinusoidal and cosinusoidal elements in the adaptive Fourier series in the original AFO system, a phase shift between the input signal and the extracted phase might appear.

Figure 3 shows the adaptation to the periodic signal with different initial conditions of parameters A_i and B_j. The results

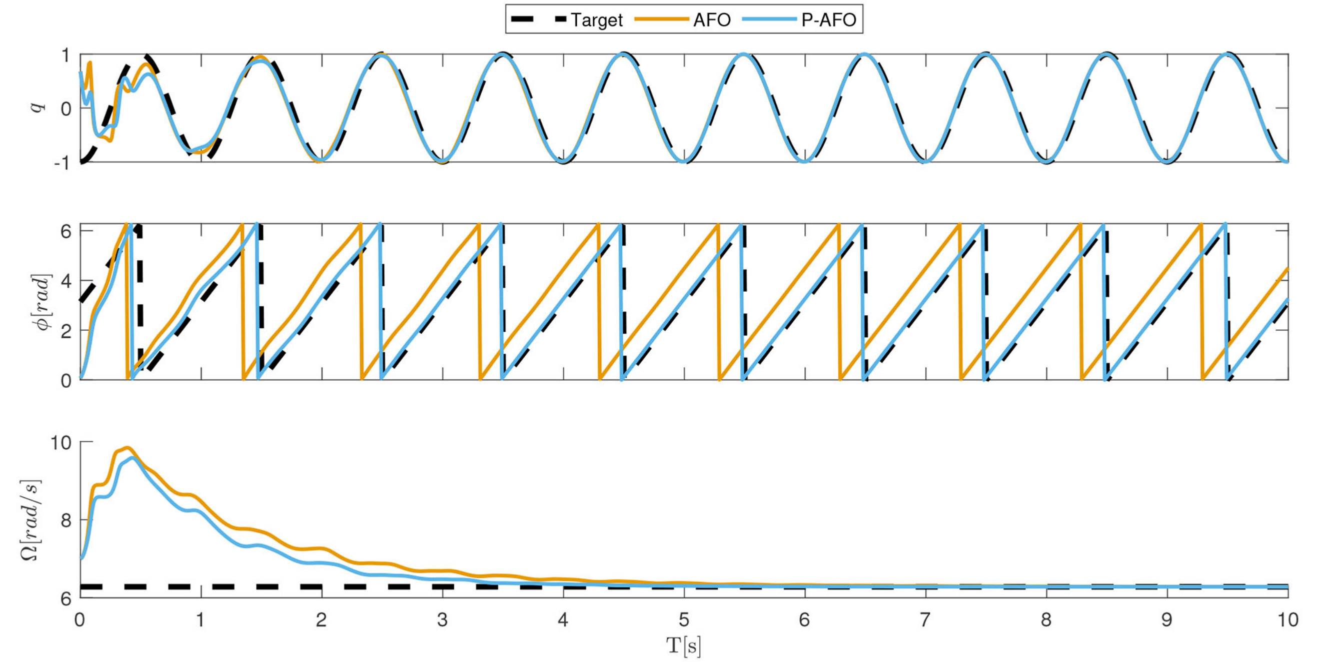


FIGURE 2 | Typical convergence of an AFO and P-AFO systems driven by a sinusoidal periodic signal. In the top plot, the comparison between the input signal and the approximation of the system is shown. The middle plot shows the phase synchronization and the bottom plot shows the frequency adaption.

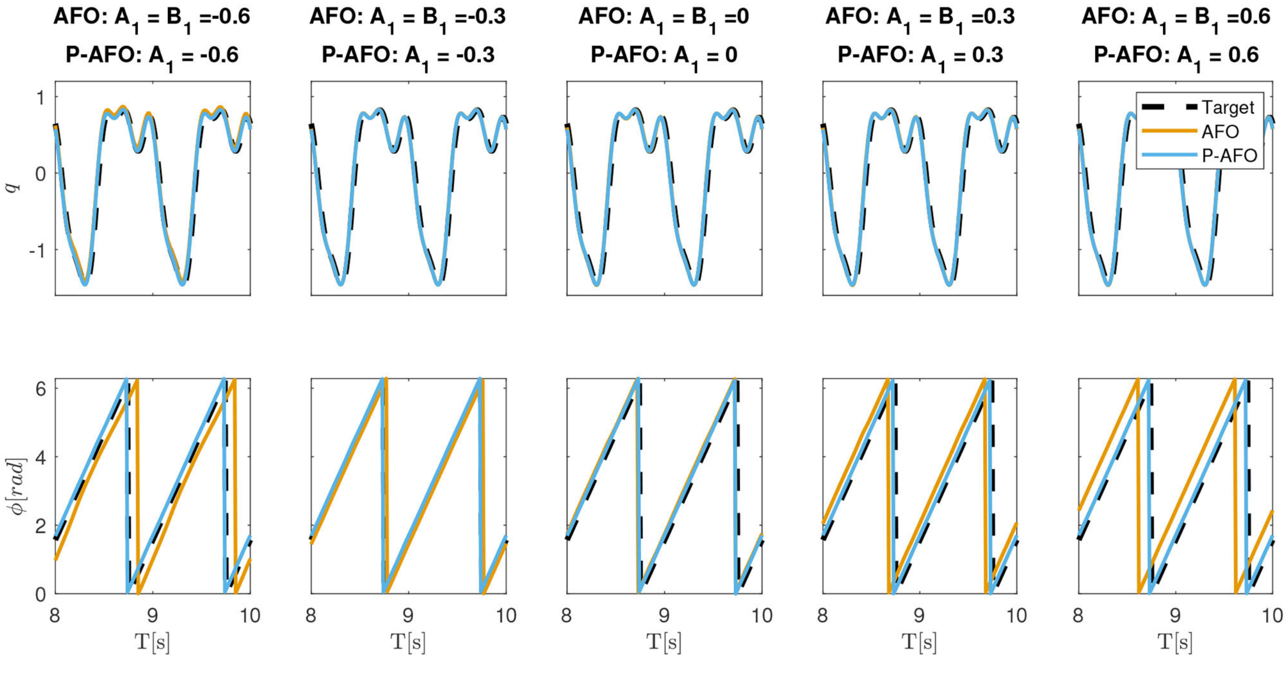


FIGURE 3 | Typical convergence of an AFO and P-AFO systems driven by a periodic signal with different initial conditions. Top plots, shows the comparison between the input signal and the approximation of the system and middle plots shows the phase synchronization.

shows that the phase synchronization of the original AFO concerning the input signal is not repeatable. Note, that if we change the initial parameters or the start of the input signal, the phase shift between the input signal and the extracted phase of AFO will be different. Extracting the exact phase of the input signal is crucial for the P-CMPs. In the middle plot of

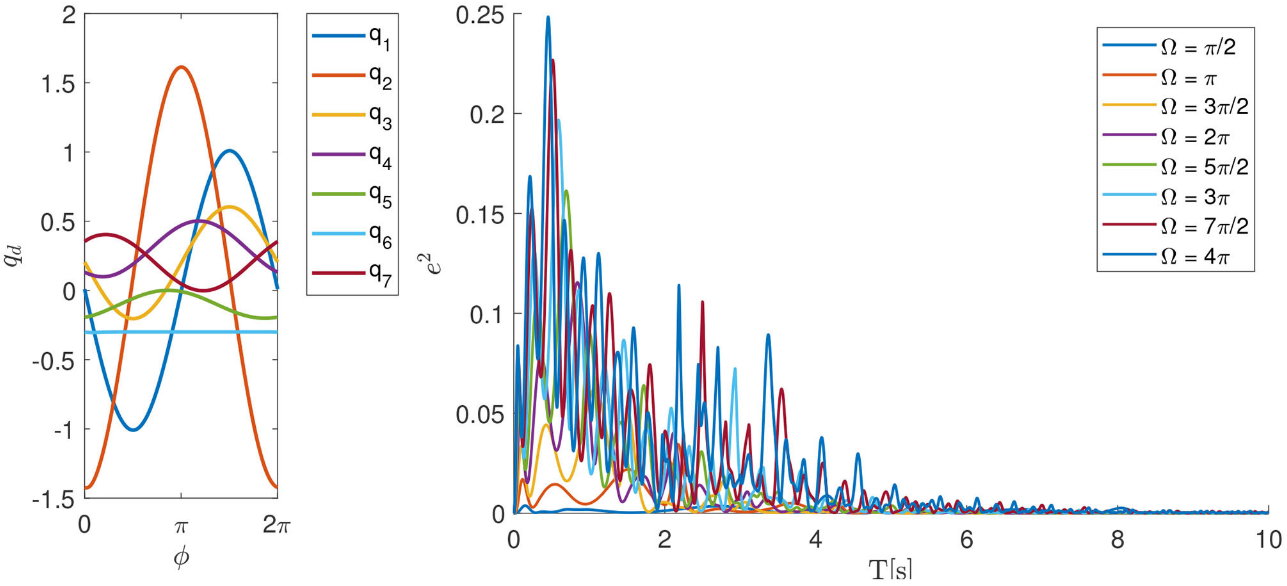


FIGURE 4 | Learning of internal dynamical models for different motion frequencies on 7 degrees of freedom Kuka LWR-4 robot. The left plot shows the desired kinematic motion q_d dependent on the phase parameter ϕ and the right plot shows the sum of square motion tracking error during the leaning process.

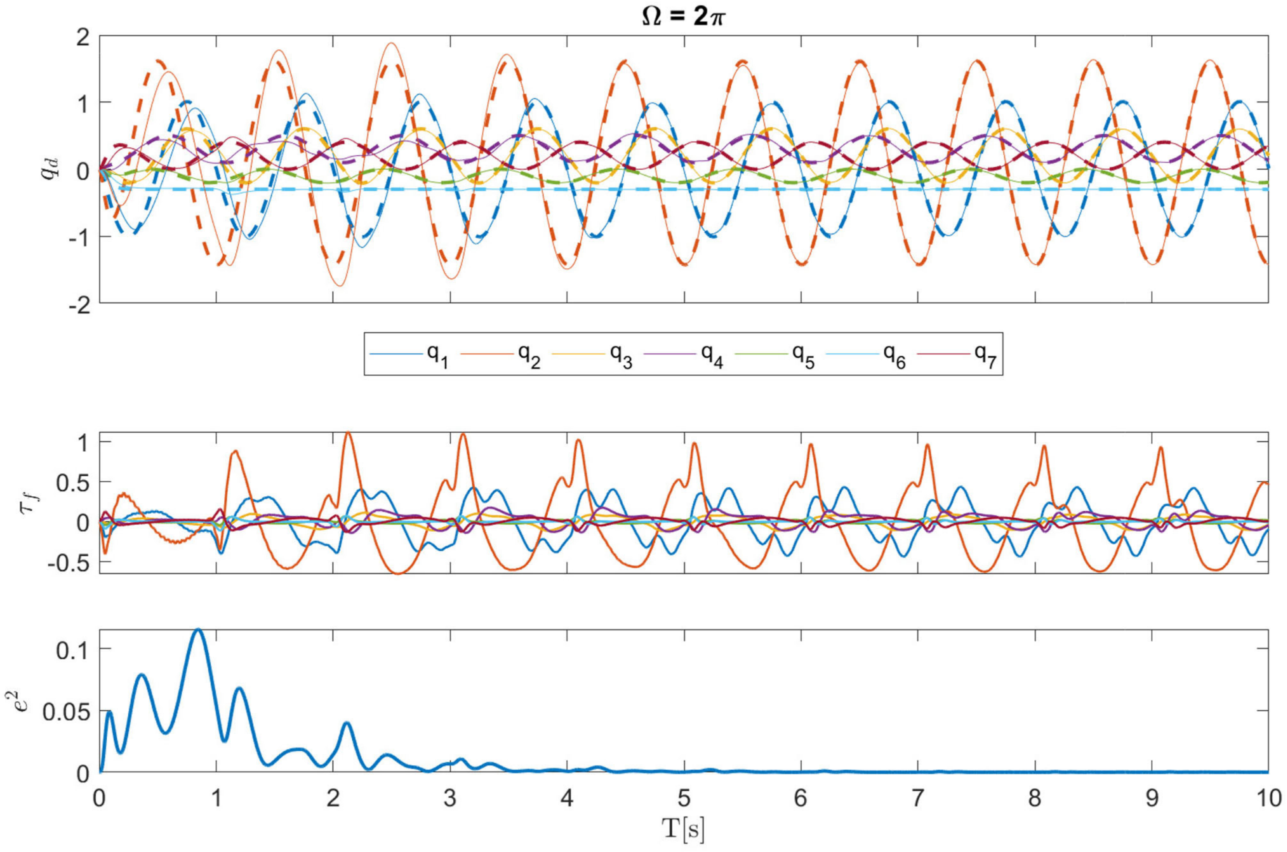


FIGURE 5 | Top and middle plots show example joint and torque trajectories, respectively ($\Omega = 2\pi$ example). The bottom plot shows the sum of the square motion tracking error during the leaning process.

Figure 2 and on the bottom plot of **Figure 3**, we can see that the proposed P-AFO system ensures that the phase is always unambiguously defined concerning the input signal. This allows us to precisely anchor the P-TPs to the corresponding P-DMPs, which, therefore, provides all the aforementioned advantages of a P-CMPs system.

4.2. Robot Dynamics Learning Example

To illustrate the ability to learn the internal dynamical model, we implemented the P-CMPs approach on a real robot Kuka LWR-4. In this example, the goal was to learn the corresponding dynamical model in P-TPs using the approach proposed in section 3. The kinematic trajectory for this task was predefined for all 7 degrees of freedom and it is shown on the left hand side of **Figure 4**. The robot feedback loop gains for all joints were set to 50 Nm/rad and the feed-back error learning parameter ι was set to 10. Note that in general the dynamical model of the robot is not strictly necessary for the proposed approach, however, we made use of the dynamical model provided by the Kuka controller. Even so, by using the provided dynamical model the tracking accuracy is still poor with selected feed-back gain as shown in Deniša et al. (2016).

By using the proposed P-CMPs system we can see that the tracking error, and hence the learning of the internal dynamical model, is rapid and successful. In the left plot in **Figure 5**, we can see several examples of learning dynamics with a different frequency of motion. Despite the fact that the robot axes are not fully dynamically decoupled, the proposed P-CMPs system is able to successfully learn the internal dynamic models, i.e., corresponding feed-forward parameters, and thereby significantly reduce the tracking error. Hence learning was successful for all desired frequencies of motion. Note that the rate of learning does not depend on the frequency of movement, as can be seen on the right-hand side-plot in **Figure 4**.

The kinematic motion improvements and the evolution of the corresponding internal dynamical models, i.e., torque profiles, is for a $\Omega = 2\pi$ example shown in **Figure 5**. Here we can see

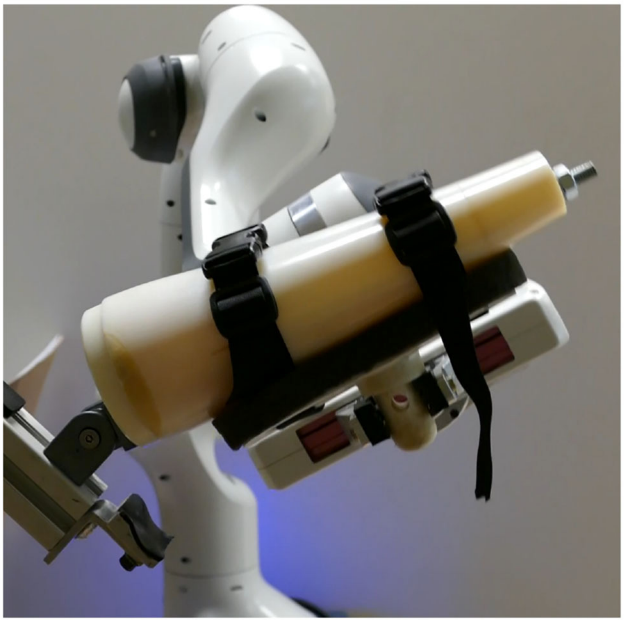


FIGURE 7 | Experimental setup for physically simulated human elbow stretching tasks.

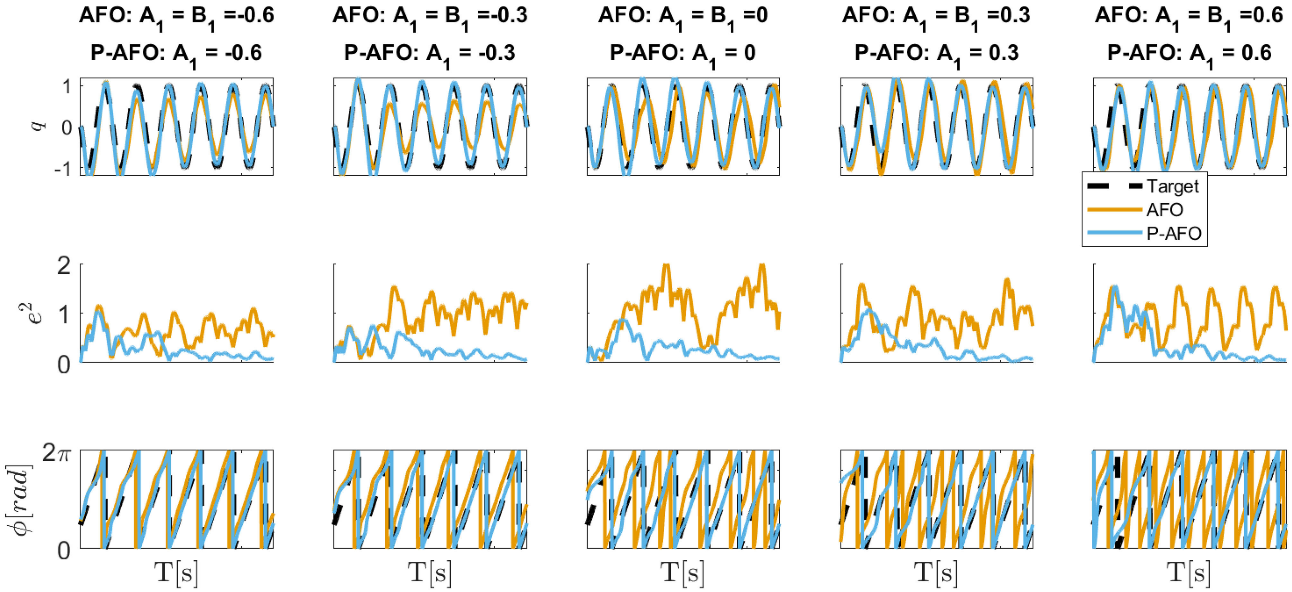


FIGURE 6 | Difference between AFO and P-AFO system, both used with P-CMPs. The top plots show the desired and actual joint movements when using previously learned P-CMPs from the example in **Figure 5**. The middle plots show the tracking error and the bottom plots show the phase synchronization of CMPs.

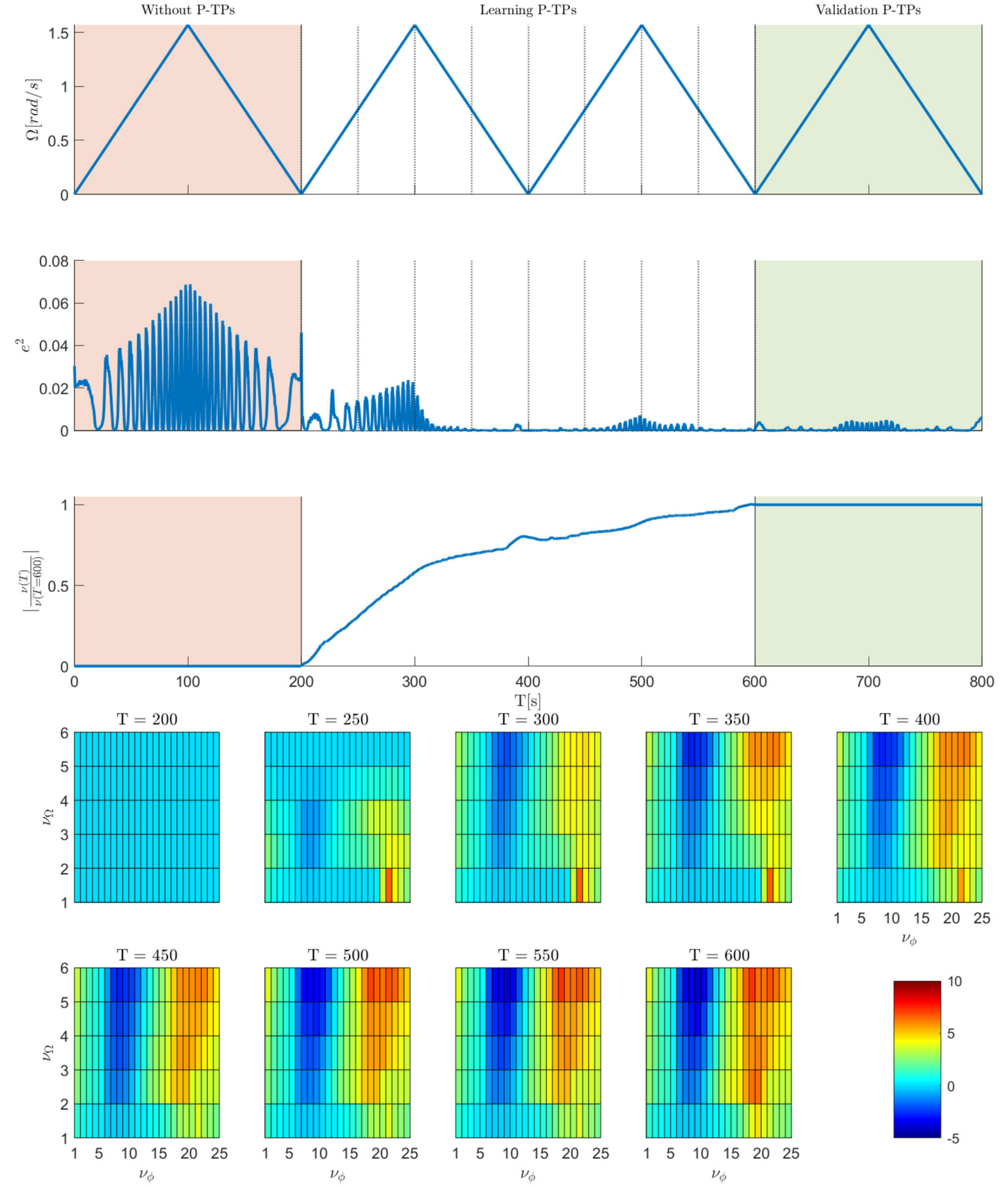


FIGURE 8 | Results of elbow stretching example. The top plot shows the desired motion frequency. The second plot shows the sum of square tracking errors. The third plot shows the relationship between current and final weight matrix. Bottom plots show the P-TPs weight matrix values for one joint at a certain time during the learning process.

from the bottom plot that tracking error is significantly reduced in a relatively short time. After about four periods of motion, the feed-forward torque signals converge to the final shape.

In **Figure 6** also we show why it is crucial to ensure that the phase ϕ is unambiguously defined considering the input signal. In this experiment we used the learned P-CMPs from the example in **Figure 5** to compare the original AFO system with the P-AFO system. Note that both AFO and P-AFO systems were used with the proposed multi-layered control system based on P-CMPs. As expected when AFO is used, the feed-forward torque primitives might be shifted due to the properties of the original AFO approach. The phase shift of the torque primitives encoded in P-TPs, clearly results in a larger error compared to the new P-AFO approach which has an unambiguously define phase. As also shown in the example on **Figure 3** the extracted phase of the P-AFO system is always clearly defined with respect to the input signal regardless of the initial conditions, while the extracted phase from the original AFO system might vary, with respect to the input signal. Note that an unambiguously defined phase is needed to ensure a reliable response of the P-CMPs.

4.3. Physical Interaction Example

In the last example, the proposed P-CMPs method was demonstrated on a task where the robot was holding a human hand model with the simulated elbow joint as shown in **Figure 7**. In this experimental setup, we used a Franka Emika Panda robot. Note that the dimensional proportions and weight are equal to an adult human arm. The physical arm model is part of the evaluation of the possibility to help rehabilitation specialists during rehabilitation procedures. Here a typical strategy would be that the rehabilitation specialist defines the desired kinematic motion for rehabilitation using imitation learning. In such a case, considering that the robot could also hold an actual human hand, it would be difficult to obtain a mathematical dynamical model. Due to human variability, it would be a very specific, complex, and time-consuming task.

Instead, we can use the proposed P-CMPs approach to learn task-specific, appropriate torques for a given kinematic trajectory. This task could be performed with the original CMPs system combined with the statistical generalization. However, this would not be most effective since it would require to learn the CMPs at the specific frequency to build the database. In contrast, the proposed P-CMPs framework allows learning at an arbitrary frequency, as the frequency dependence is built into the P-TPs system. Working with a compliantly controlled robot, i.e., low feedback gains, with the ability to produce accurate trajectory tracking makes the system also safer for the environment, operator, and user.

To show the P-CMPs performance, the kinematic motion for elbow stretching was defined by using kinesthetic teaching (Deniša et al., 2013). The robot feedback loop gains for all joints were set to 20 Nm/rad and the feed-back error learning parameter ι was set to 10. The performance of the P-CMPs framework for this example is shown in **Figure 8**, where we show in the top plot the desired frequency of motion, in the second plot the corresponding kinematic tracking error and in the third plot we show the relationship between current and final weight matrix for one joint. The experiment was divided into three parts, motion tracking without feed-forward P-TPs model, learning of P-TPs model, and validation of learned P-TPs model, respectively. The bottom plots show the evolution of P-TPs weights for one degree of freedom during the learning process. Note that P-TPs weights for one degree of freedom are a matrix ν with size of $M \times K$, where $M = 25$ and $K = 6$. Note that M is the number of phase kernels, and K is the number of frequency kernels. Here the c_i^{ϕ}, $i = 1, 2, ...M$ is equally distributed between 0 and 2π and c_j^{Ω}, $j = 1, 2, ...K$ is equally distributed between 0 and π.

The sum of square tracking errors shows that the proposed approach can significantly improve the kinematic tracking. We can see also that, by performing one sweep through the frequency space already significantly improves the tracking error. As seen in the third plot and bottom plots in **Figure 8**, at $T = 400$s the weights are already at the 80% of the final value. Note that in the second sweep through the frequency space the weights changes for less than 20% with respect to the weights value at $T = 600$s. Furthermore, the validation part shows that tracking error remains low, even after the learning process, as expected. Since the proposed P-CMPs approach remains parametric in terms of P-DMPs and P-TPs weights, all previously developed statistical methods can also be applied, allowing for further expansion of the task-specific learning of dynamics.

5. CONCLUSION

We presented a new P-CMPs framework consisting of a novel P-AFO frequency and phase synchronization systems, periodic DMPs, and a novel P-TPs system encoding task-specific primitives. The proposed P-CMPs system uses feedforward torque signals which are associated with corresponding kinematic motions. We showed, that the novel approach is able to unambiguously extract not only the frequency but also the phase from an arbitrary signal which allows anchoring the P-TPs to the P-DMPs trajectories. Furthermore, the novel extension of the P-TPs system also makes P-TPs frequency-dependent, which enables smooth frequency modulation of the P-CMPs. Integrating the feedback error learning concept in P-CMPs also improves the usability of the system. Our results indicate that the system was able to synchronize the kinematic and dynamics signals enabling compliant behavior while maintaining high tracking accuracy, without the need for developing mathematical dynamical models of the robot or the task.

The proposed P-CMPs framework is an improvement compared to the previews CMPs framework, enabling better learning performance and smooth frequency modulation abilities of periodic tasks.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

TP contributed to the design, execution, and drafting of this work.

FUNDING

This work was supported by Slovenian Research Agency grant N2-0130.

REFERENCES

Argall, B. D., Chernova, S., Veloso, M., and Browning, B. (2009). A survey of robot learning from demonstration. *Robot. Auton. Syst*. 57, 469–483. doi: 10.1016/j.robot.2008.10.024

Billard, A., Calinon, S., Dillmann, R., and Schaal, S. (2008). *Robot Programming by Demonstration*. Berlin; Heidelberg: Springer. doi: 10.1007/978-3-540-30301-5_60

Buchli, J., Stulp, F., Theodorou, E., and Schaal, S. (2011). Learning variable impedance control. *Int. J. Robot. Res*. 30, 820–833. doi: 10.1177/0278364911402527

Calandra, R., Ivaldi, S., and Deisenroth, M. P. (2015). "Learning torque control in presence of contacts using tactile sensing from robot skin," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)* (Seoul: IEEE), 690–695. doi: 10.1109/HUMANOIDS.2015.7363429

Calinon, S. (2016). A tutorial on task-parameterized movement learning and retrieval. *Intell. Service Robot*. 9, 1–29. doi: 10.1007/s11370-015-0187-9

Calinon, S., Alizadeh, T., and Caldwell, D. G. (2013). "On improving the extrapolation capability of task-parameterized movement models," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Tokyo: IEEE), 610–616. doi: 10.1109/IROS.2013.6696414

Calinon, S., Bruno, D., Malekzadeh, M. S., Nanayakkara, T., and Caldwell, D. G. (2014). Human-robot skills transfer interfaces for a flexible surgical robot. *Comput. Methods Prog. Biomed*. 116, 81–96. doi: 10.1016/j.cmpb.2013.12.015

Del Prete, A., and Mansard, N. (2016). Robustness to joint-torque-tracking errors in task-space inverse dynamics. *IEEE Trans. Robot*. 32, 1091–1105. doi: 10.1109/TRO.2016.2593027

Deniša, M., Gams, A., Ude, A., and Petrič, T. (2016). Learning Compliant Movement Primitives Through Demonstration and Statistical Generalization. *IEEE/ASME Transactions on Mechatronics*, 21, 2581–2594. doi: 10.1109/TMECH.2015.2510165

Deniša, M., Petrič, T., Asfour, T., and Ude, A. (2013). "Synthesizing compliant reaching movements by searching a database of example trajectories," in *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (Atlanta, GA: IEEE), 540–546. doi: 10.1109/HUMANOIDS.2013.7030026

Filippeschi, A., Schmitz, N., Miezal, M., Bleser, G., Ruffaldi, E., and Stricker, D. (2017). Survey of motion tracking methods based on inertial sensors: a focus on upper limb human motion. *Sensors* 17:1257. doi: 10.3390/s17061257

Forte, D., Gams, A., Morimoto, J., and Ude, A. (2012). On-line motion synthesis and adaptation using a trajectory database. *Robot. Auton. Syst*. 60, 1327–1339. doi: 10.1016/j.robot.2012.05.004

Franklin, D. W., and Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron* 72, 425–442. doi: 10.1016/j.neuron.2011.10.006

Gams, A., Denisa, M., and Ude, A. (2015). "Learning of parametric coupling terms for robot-environment interaction," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)* (Seoul: IEEE), 304–309. doi: 10.1109/HUMANOIDS.2015.7363559

Gams, A., Do, M., Ude, A., Asfour, T., and Dillmann, R. (2010). "On-line periodic movement and force-profile learning for adaptation to new surfaces," in *2010 10th IEEE-RAS International Conference on Humanoid Robots* (Nashville, TN: United States), 560–565. doi: 10.1109/ICHR.2010.5686306

Gams, A., Ijspeert, A. J., Schaal, S., and Lenarčič, J. (2009). On-line learning and modulation of periodic movements with nonlinear dynamical systems. *Auton. Robots* 27, 3–23. doi: 10.1007/s10514-009-9118-y

Gopalan, N., Deisenroth, M. P., and Peters, J. (2013). "Feedback error learning for rhythmic motor primitives," in *2013 IEEE International Conference on Robotics and Automation* (Karlsruhe: IEEE), 1317–1322. doi: 10.1109/ICRA.2013.6630741

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., and Schaal, S. (2013). Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput*. 25, 328–373. doi: 10.1162/NECO_a_00393

Kawato, M. (1990). "Feedback-error-learning neural network for supervised motor learning," in *Advanced Neural Computers*. ed R. Eckmiller (Amsterdam: Elsevier), 365–372. doi: 10.1016/B978-0-444-88400-8.50047-9

Khansari-Zadeh, S. M., and Billard, A. (2011). Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Trans. Robot*. 27, 943–957. doi: 10.1109/TRO.2011.2159412

Kober, J., Wilhelm, A., Oztop, E., and Peters, J. (2012). Reinforcement learning to adjust parametrized motor primitives to new situations. *Auton. Robots* 33, 361–379. doi: 10.1007/s10514-012-9290-3

Koropouli, V., Hirche, S., and Lee, D. (2015). Generalization of force control policies from demonstrations for constrained robotic motion tasks. *J. Intell. Robot. Syst*. 80, 133–148. doi: 10.1007/s10846-015-0218-y

Kronander, K., and Billard, A. (2013). Learning compliant manipulation through kinesthetic and tactile human-robot interaction. *IEEE Trans. Hapt*. 7, 367–380. doi: 10.1109/TOH.2013.54

Kronander, K., Khansari, M., and Billard, A. (2015). Incremental motion learning with locally modulated dynamical systems. *Robot. Auton. Syst*. 70, 52–62. doi: 10.1016/j.robot.2015.03.010

Kulvicius, T., Biehl, M., Aein, M. J., Tamosiunaite, M., and Worgotter, F. (2013). Interaction learning for dynamic movement primitives used in cooperative robotic tasks. *Robot. Auton. Syst*. 61, 1450–1459. doi: 10.1016/j.robot.2013.07.009

Lee, D., and Ott, C. (2011). Incremental kinesthetic teaching of motion primitives using the motion refinement tube. *Auton. Robots* 31, 115–131. doi: 10.1007/s10514-011-9234-3

Matsubara, T., Hyon, S.-H., and Morimoto, J. (2011). Learning parametric dynamic movement primitives from multiple demonstrations. *Neural Netw*. 24, 493–500. doi: 10.1016/j.neunet.2011.02.004

Mülling, K., Kober, J., Kroemer, O., and Peters, J. (2013). Learning to select and generalize striking movements in robot table tennis. *Int. J. Robot. Res*. 32, 263–279. doi: 10.1177/0278364912472380

Nakanishi, J., and Schaal, S. (2004). Feedback error learning and nonlinear adaptive control. *Neural Netw*. 17, 1453–1465. doi: 10.1016/j.neunet.2004.05.003

Nguyen-Tuong, D., and Peters, J. (2011). Model learning for robot control: a survey. *Cogn. Process*. 12, 319–340. doi: 10.1007/s10339-011-0404-1

Peternel, L., Petrič, T., and Babič, J. (2018). Robotic assembly solution by human-in-the-loop teaching method based on real-time stiffness modulation. *Auton. Robots*. 42, 1–17. doi: 10.1007/s10514-017-9635-z

Peternel, L., Petrič, T., Oztop, E., and Babič, J. (2014). Teaching robots to cooperate with humans in dynamic manipulation tasks based on multi-modal human-in-the-loop approach. *Auton. Robots*. 36, 123–136. doi: 10.1007/s10514-013-9361-0

Petrič, T., Curk, B., Cafuta, P., and Žlajpah, L. (2010). Modelling of the robotic powerball®: a nonholonomic, underactuated and variable structure-type system. *Math. Comput. Modell. Dynamic. Syst*. 16, 327–346. doi: 10.1080/13873954.2010.484237

Petrič, T., Gams, A., Colasanto, L., Ijspeert, A. J., and Ude, A. (2018). Accelerated sensorimotor learning of compliant movement primitives. *IEEE Trans. Robot*. 34, 1636–1642. doi: 10.1109/TRO.2018.2861921

Petrič, T., Gams, A., Ijspeert, A. J., and Žlajpah, L. (2011). On-line frequency adaptation and movement imitation for rhythmic robotic tasks. *Int. J. Robot. Res*. 30, 1775–1788. doi: 10.1177/0278364911421511

Petrič, T., Gams, A., Žlajpah, L., and Ude, A. (2014). "Online learning of task-specific dynamics for periodic tasks," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Chicago, IL: IEEE), 1790–1795. doi: 10.1109/IROS.2014.6942797

Racca, M., Pajarinen, J., Montebelli, A., and Kyrki, V. (2016). "Learning in-contact control strategies from demonstration," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Daejeon), 688–695. doi: 10.1109/IROS.2016.7759127

Rozo, L., Calinon, S., Caldwell, D. G., Jimenez, P., and Torras, C. (2016). Learning physical collaborative robot behaviors from human demonstrations. *IEEE Trans. Robot*. 32, 513–527. doi: 10.1109/TRO.2016.2540623

Schwarz, M., and Behnke, S. (2013). "Compliant robot behavior using servo actuator models identified by iterative learning control," in *Robot Soccer World Cup*, Lecture Notes in Computer Science, Vol. 8371, eds S. Behnke, M. Veloso, A. Visser, and R. Xiong (Berlin; Heidelberg: Springer), 207–218. doi: 10.1007/978-3-662-44468-9_19

Sciavicco, L., and Siciliano, B. (2012). *Modelling and Control of Robot Manipulators*. London: Springer Science & Business Media.

Steinmetz, F., Montebelli, A., and Kyrki, V. (2015). "Simultaneous kinesthetic teaching of positional and force requirements for sequential in-contact tasks," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)* (Seoul: IEEE), 202–209. doi: 10.1109/HUMANOIDS.2015.7363552

Stulp, F., Raiola, G., Hoarau, A., Ivaldi, S., and Sigaud, O. (2013). "Learning compact parameterized skills with a single regression," in *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (Atlanta, GA: IEEE), 417–422. doi: 10.1109/HUMANOIDS.2013.7030008

Ude, A., Gams, A., Asfour, T., and Morimoto, J. (2010). Task-specific generalization of discrete and periodic dynamic movement primitives. *IEEE Trans. Robot*. 26, 800–815. doi: 10.1109/TRO.2010.2065430

Wang, Y., Gao, F., and Doyle, F. J. III. (2009). Survey on iterative learning control, repetitive control, and run-to-run control. *J. Process Control* 19, 1589–1600. doi: 10.1016/j.jprocont.2009.09.006

Welschehold, T., Dornhege, C., and Burgard, W. (2016). "Learning manipulation actions from human demonstrations," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Daejeon), 3772–3777. doi: 10.1109/IROS.2016.7759555

Žlajpah, L., and Petrič, T. (2019). Unified virtual guides framework for path tracking tasks. *Robotica*. 38, 1807–1823. doi: 10.1017/S0263574719000973

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

ORIGINAL RESEARCH
published: 12 November 2020
doi: 10.3389/fnbot.2020.582385

Correlated Skin Surface and Tumor Motion Modeling for Treatment Planning in Robotic Radiosurgery

Shumei Yu[1], Pengcheng Hou[1], Rongchuan Sun[1]*, Shaolong Kuang[1], Fengfeng Zhang[1], Mingchuan Zhou[2], Jing Guo[3] and Lining Sun[1]

[1] School of Mechanical and Electrical Engineering, Soochow University, Suzhou, China, [2] Computer Aided Medical Procedures, Technical University of Munich, Munich, Germany, [3] School of Automation, Guangdong University of Technology, Guangzhou, China

In robotic radiosurgery, motion tracking is crucial for accurate treatment planning of tumor inside the thoracic or abdominal cavity. Currently, motion characterization for respiration tracking mainly focuses on markers that are placed on the surface of human chest. Nevertheless, limited markers are not capable of expressing the comprehensive motion feature of the human chest and abdomen. In this paper, we proposed a method of respiratory motion characterization based on the voxel modeling of the thoracoabdominal torso. Point cloud data from depth cameras were used to achieve three-dimensional modeling of the chest and abdomen surface during respiration, and a dimensionality reduction algorithm was proposed to extract respiratory features from the established voxel model. Finally, experimental results including the accuracy of voxel model and correlation coefficient were compared to validate the feasibility of the proposed method, which provides enhanced accuracy of target motion correlation than traditional methods that utilized external markers.

Keywords: respiratory motion characterization, voxel model, correlation model, tumor tracking, surface modeling

Edited by:
Zhan Li,
University of Electronic Science and Technology of China, China

Reviewed by:
Qiang Li,
Bielefeld University, Germany
Yongliang Yang,
Michigan State University, United States

***Correspondence:**
Rongchuan Sun
sunrongchuan@suda.edu.cn

Received: *11 July 2020*
Accepted: *25 September 2020*
Published: *12 November 2020*

Citation:
Yu S, Hou P, Sun R, Kuang S, Zhang F, Zhou M, Guo J and Sun L (2020) Correlated Skin Surface and Tumor Motion Modeling for Treatment Planning in Robotic Radiosurgery. Front. Neurorobot. 14:582385. doi: 10.3389/fnbot.2020.582385

INTRODUCTION

Robotic radiosurgery shows significant advantages in treating tumors that are not suitable to be treated by chemotherapy (Coste-Maniere et al., 2005). Since control of radioactive beam according to tumor motion tracking ensures treatment of high precision, accurate tumor respiratory motion tracking is crucial for stereotactic radiosurgical robots. Movement of tumors caused by respiration is complex to be modeled; usually movement of tumors is obtained by X-ray imaging. However, the method of locating tumors *in vivo* by frequent irradiation of X-rays causes unnecessary secondary damage to normal tissues. Therefore, correlating the skin motion with tumor motion simultaneously caused by respiration is necessary before establishing a prediction model to obtain the upcoming position of tumor.

To develop a treatment planning system considering tumor motion caused by human breath, respiratory motion characterization of the skin surface is needed first. Several methods have been developed to model the respiratory skin motion. According to surrogate variations, researchers have studied the correlation of tumor movements with motion of markers on the chest and abdomen surface, multimodal (airflow, tidal volume, acceleration, force) features, and motion measurement of the chest and abdomen surface, respectively.

Currently, placing markers on the skin surface is the mostly adopted way to study the respiration characterization and external–internal motion correlation. In the study of the characterization of respiratory motion with a single marker, the motion of one marker on the abdominal skin and tumor was fitted as a linear relationship *in vivo*, but insufficient characterization in experiments was found (Kanoulas et al., 2007). CyberKnife (Cheng and Adler, 2006) used three markers to record the motion of skin surface in respiratory tracking system. Nevertheless, if the patient performs complex breathing pattern, limited surrogates could not reflex the complicated respiration mode. The correlation models between three markers and tumor *in vivo* based on linear/quadratic fitting, artificial neural network, and fuzzy logic were compared, respectively, and the results found that the correlation coefficient of fuzzy logic algorithm is higher (Torshabi et al., 2013). The tumor tracking method proposed (Iizuka et al., 2015) is based on a pre-established 4D model, which connects internal tumor motion with external respiratory signals. Respiratory signals were collected by monitoring four infrared (IR) markers on the abdominal wall of the patient using an IR camera. The method can effectively reduce the influence of radiotherapy on normal tissues and further provides local control. However, due to insufficient characterization of respiratory signals by a limited number of IR markers, there was inevitable accumulation error of targeting in the abdominal region. The correlation factors between four surface markers and the tumor in different directions of motion were analyzed (Koch et al., 2004). Studies have shown that the association degree of skin surface markers depends on the placement of markers and the breathing pattern of patients. In the study of six markers, the centroid motion curves of all markers and the corresponding tumor motion curves were calculated; finally, tumor locations by proportional linear interpolation were obtained (Schweikard et al., 2000). Furthermore, 19 infrared LED markers on abdominal skin of swine were adopted to study the respiration feature and proved that the more markers utilized on the skin, the more respiratory movement information can be extracted (Ernst et al., 2009).

For the study of multimodal information of the chest and abdomen surface, researchers studied the linear correlation factors of tidal volume, abdominal surface displacement, and the anterior–posterior displacement of tumor during human respiration and concluded that tidal volume has higher correlation with tumor displacement (Hoisak et al., 2004). However, the correlation between tidal volume and tumor has a noticed drift with respiratory phase, and the measurement of tidal volume brings great pain to patients. In addition, multimodal information using tension bands, optical markers, acceleration sensors, airflow acceleration, and temperature sensors was collected to compare their correlation effects of inner–outer respiratory motion and found that optical marker achieves the best performance for the tumor motion tracking (Durichen et al., 2013).

For the study of correlating tumor motion with chest and abdomen surface motion, there have been some exploratory researches on the measurement of the whole thoracoabdominal surface motion. A multiradar wireless system was designed to track respiratory movements in real time. Through the two radar devices integrated on the linear accelerator, the movements of the chest and abdomen are monitored in real time (Gu et al., 2011; Gu and Li, 2015). However, the subjects need to spend much time and energy on training the breathing modes, which are required for the breathing exercise. One study printed markers as 11 circles on Lycra T-shirt (Ernst and Saß, 2015). Experiments demonstrated the advantages of depth camera over traditional optical acquisition equipment in measuring multiple moving targets, along with defects of image distortion and noise increase. Studies used RGB-D camera to collect continuous depth images of chest and abdomen regions of patients, with principal component analysis (PCA) used to create a respiratory motion model that displayed as tidal volume change (Wijenayake and Park, 2017). A method was proposed to estimate tidal volume changes by using depth camera to reconstruct three-dimensional isochronal surface of chest (Transue et al., 2016). The whole isosurface was extracted to characterize the deformation and volume changes of the patient's chest. Due to the inherent inconsistencies in the depth values provided by the depth camera, there are still problems such as depth measurement errors that need to be resolved.

The study of chest and abdomen surface measurement shows the advantages of the whole skin surface for the complete expression of motion information. Thus, to better model the motion of the whole thoracoabdominal surface and correlate, the surface information with the tumor motion is crucial. To obtain the respiratory motion characteristics of chest and abdomen surface, proper and accurate feature extraction is necessary. To the best knowledge of the authors, there are no available methods to obtain the feature extraction for the respiratory motion model comprised of voxels. However, our proposed method is inspired by previous research on feature extraction specially focusing on the classification and recognition in image processing. Studies used a convolutional network to train simple linear iterative clustering (SLIC) superpixels and obtained the feature embedding space of each superpixel, from which the superpixel feature vector was extracted (Liu et al., 2018). Some researchers used multifactor distribution as soft label and extracted supplementary information through visual input. Convolution neural network was used for visual feature learning, and the painting styles were classified through label distribution learning (Yang et al., 2018). In the research of data dimensionality reduction, locally linear embedding (LLE) and Gustafson–Kessel algorithms were used for dimension reduction in gray and RGB color images (Wang et al., 2018). A novel method introduced an end-to-end trainable and deep Siamese-like network named PersEmoN that consists of two convolutional network branches. Both networks share their bottom feature extraction module and are optimized within a multitask learning framework (Zhang et al., 2019). Experiments showed the effectiveness of PersEmoN on two apparent personality and emotion datasets.

In this paper, based on our previous work on dynamic voxel modeling of thoracic–abdominal surface (Hou et al., 2019), a characterization method of respiratory motion was proposed, as shown in **Figure 1**. A visual information acquisition system with two depth cameras was established to collect the point

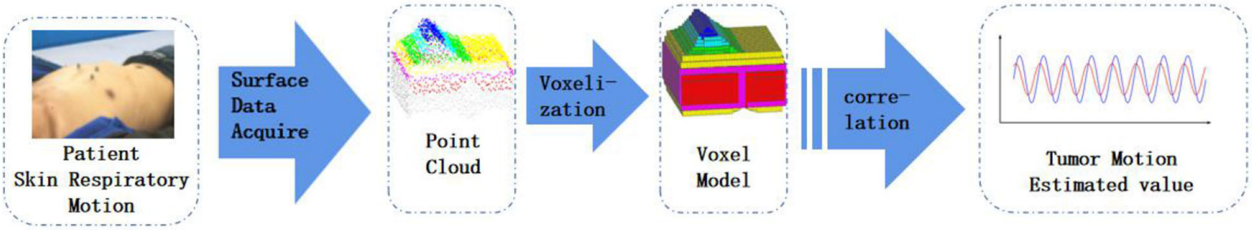


FIGURE 1 | Flowchart of the respiration tracking system.

cloud data of respiration movement. Afterwards, we constructed the point cloud surface into a watertight model using boundary interpolation. The watertight model was built into voxel model in the final step. Specifically, the established surface models based on time series are unprecedentedly to be transformed into voxel models with more three-dimensional structural information. Respiration features are extracted from voxel models by dimensionality reduction and used as a description of the whole thoracoabdominal torso. A correlation model between respiratory features and tumor motion was established. Finally, experimental results with model accuracy and correlation factor were studied to validate the proposed approach.

Compared with the solutions using limited number of infrared markers placed on the abdomen, the proposed method obtained more rich information about the patient's body surface by building 3D voxel models. A dimensionality reduction method based on 3D voxel model could extract more robust respiratory features from the established body surface model. It could also overcome the limitation of information loss problem that existed in current respiration tracking methods.

MATERIALS AND METHODS

The main framework of constructing correlation model is (1) establish dynamic thoracoabdominal surface voxel model, (2) reduce the dimension of the voxel model and extract the low-dimensional representation vector of the voxel model, and (3) establish the correlation model between the representation vector and the tumor motion state.

Dynamic Thoracoabdominal Surface Voxel Modeling

Three-dimensional modeling of dynamic human thoracoabdominal skin surface during respiration mainly includes point cloud collecting of dynamic skin surface using multiple cameras of Kinect V2, model establishment of thoracoabdominal skin surface, and surface reconstruction into voxel model.

Point Cloud Acquisition System

With modeling integrity considered, we adopted time-series-based strategy when building the model with data from multiple cameras. Simultaneous frames of different cameras were fused into one frame, and then, the fused frames were arranged in time sequences. Due to the overlap in the fusion of two frames of time asynchrony, it is necessary to consider the time synchronization problem of multicamera sampling.

Since the exposure and trigger times of multiple cameras cannot be completely unified, we considered sampling with an approximate synchronization strategy. Two computers control two cameras in a one-to-one manner.

As images from multiple cameras have their own coordinates, an algorithm based on 2D image calibration was developed to unify the separate coordinates. In the algorithm, a calibration plate coordinate was built first by identification of the corner points, and then, the camera coordinate could be converted to the universal coordinates. Transformation from the camera coordinates system at any point x_s to the calibration plate coordinate system as x_m was as follows:

$$x_m = R_s\left(x_s - t_s\right) \tag{1}$$

Among them, t_s is the center of the calibration plate in the camera coordinate system, and R_s is the rotation matrix. The position of the calibration plate in the world coordinate system is known, so any point can be converted from the coordinate system x_m to the world coordinate system x_c:

$$x_c = R_m x_m + t_m \tag{2}$$

where R_m and t_m are the rotation and translation of the calibration plate in the world coordinate system. **Figure 2A** shows the calibration setup consisting of two depth cameras and one calibration plate, and **Figure 2B** shows the point cloud data obtained through the above calibration procedures.

Modeling of Thoracoabdominal Surface With Respiratory Movement

Although the point clouds of multiple cameras have been unified into the same coordinate system, the raw data of point cloud has noises and outliers brought by cameras themselves and infrared interference between each other camera. To pre-process the raw data, we used bilateral filtering (Tomasi and Manduchi, 1998) in denoise and statistical filtering (Moore, 1978) to eliminate outliers.

Due to certain calibration errors, the registration of the adjusted multiple point clouds still has overlap, which requires precise registration. We used a classical point cloud registration

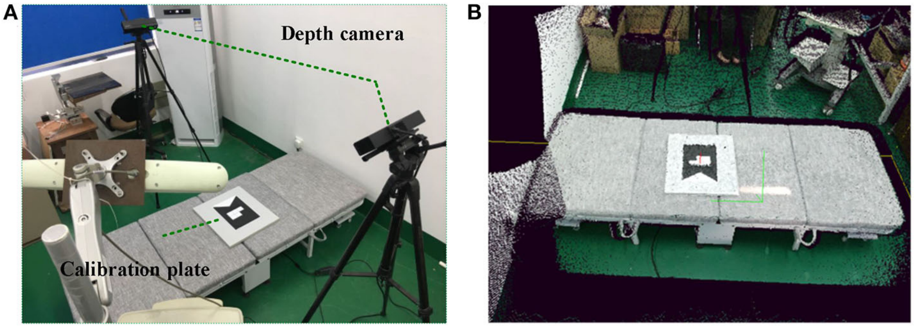


FIGURE 2 | Calibration process for two Kinect V2 cameras. **(A)** Multicamera system. **(B)** Point cloud after calibration.

algorithm iterative closest point (ICP) to register multiple point clouds. The basic principle of ICP algorithm is to find the nearest point (p_i, q_i) in the target point cloud P and the source point cloud Q according to certain constraint conditions and then calculate the optimal matching parameter R and t to minimize the error function $E(R, t)$. The error function is:

$$E\left(R,t\right)=\frac{1}{n}\sum_{i=1}^{n}\left\|q_i-\left(Rp_i+t\right)\right\|^2 \tag{3}$$

where n is the number of nearest point pairs, p_i is one point in the target point cloud P, q_i is the nearest point corresponding to p_i in the source point cloud Q, R is the rotation matrix, and t is the translation vector. **Figure 2** shows the two-frame point cloud fusion.

In order to obtain the required model of thoracoabdominal surface region and reduce the computation of point cloud processing, the existing model needs to be segmented. Due to the fixed placement of multiple cameras and the fixed scene, this paper adopted a fast and convenient segmentation algorithm with distance and color thresholds. The chest and abdomen area of the lying subject and the position of the treatment bed are constrained, and the auxiliary limit of the color threshold is applied to divide the expected chest and abdomen surface area. The segmented surface is uneven and has burrs. In order to make the thoracoabdominal surface model smooth, point clouds need to be smoothed. In this paper, moving least squares method (Breitkopf et al., 2005) was used to smooth point clouds.

Watertight Thoracoabdominal Model Establishment and Voxelization

To present more available information for respiration movement feature extraction, the surface model was transformed into a watertight model first. The procedure mainly contains boundary estimation and boundary insertion. The purpose of transforming the point cloud into watertight model is to facilitate the follow-up research to analyze the three-dimensional structure of the chest and abdomen surface based on time series and to avoid the structural problems such as unsmooth boundary and voids on the surface caused by the unclosed model.

For boundary estimation, angles between searching points p and its adjacent points $p_1,\ldots,p_k$ were adopted as criteria to detect the boundary. Threshold of the angle value was set to classify the boundary after angles of point p, and its neighbor points were calculated. Classified boundary points combined a set named as Q_b. **Figure 3A** shows the boundary point sets Q_b were divided into $bw, bs, be \in Q_b$, and **Figure 3B** shows the watertight thoracoabdominal model after point cloud interpolation.

Figure 4 shows the workflow of the watertight model construction. First, the original point cloud input was used to generate the back part of the model. During the process, the projection plane was established as a xy projection plane PL that is 5 cm lower than the xy plane bounded by bw boundary. The back part of the point cloud set $B(t)$ was built by projecting the corresponding points $q_i \in Q$ in the original point cloud Q to the plane PL. To make the model watertight, boundary insertion was executed by using a method based on projection interpolation. As has been described above in **Figure 3A**, the boundary point cloud input was divided into the shoulder, waist, and abdominal end boundaries by boundary segmentation. After that, the shoulder part $S(t)$ and the end of the abdomen $E(t)$ was obtained by interpolating point cloud equally along the z negative axis to the plane PL, respectively; the waist part $W(t)$ was obtained by interpolating point cloud equally along the waist curvature to the plane PL.

The watertight model composed of point cloud expressed the 3D information of the measured surface by independent points. To illustrate the space relationship of the entire surface, meshing was made to the point cloud model. Here, we used the classical Poisson reconstruction method (Kazhdan et al., 2006), in which normal vectors of the point clouds were calculated to display the curvature changes of the surface model. A smooth surface was built by estimation through indicator function's implicit fitting. Thus, the watertight thoracoabdominal point cloud model can be reconstructed into a model with smooth surface, as is shown in **Figure 5A**.

To study the volume changes of thoracoabdominal trunk movement, voxelization of the surface model is needed to convert the geometric representation of an object to the nearest voxel representation, which not only contains the

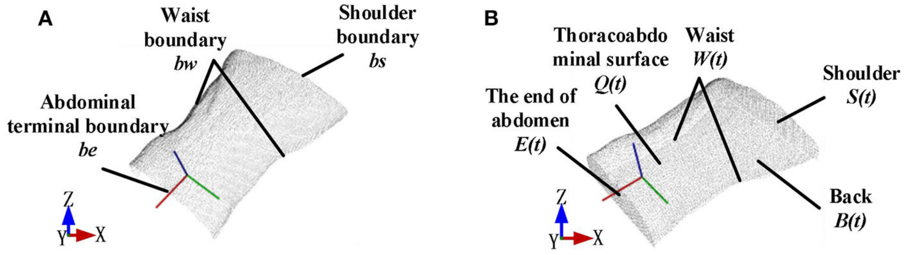


FIGURE 3 | Point cloud interpolation to form a watertight model. **(A)** Boundary division before point cloud interpolation. **(B)** Watertight model after point cloud interpolation.

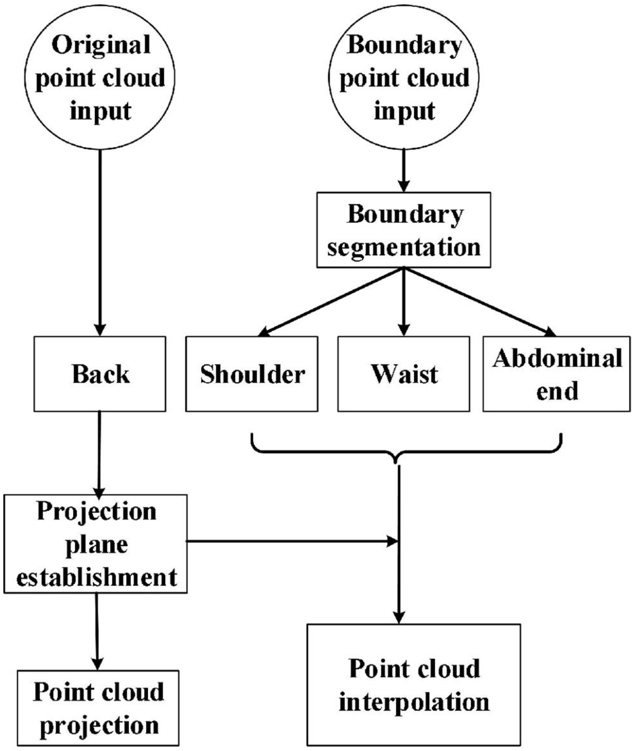


FIGURE 4 | Flowchart of watertight model construction.

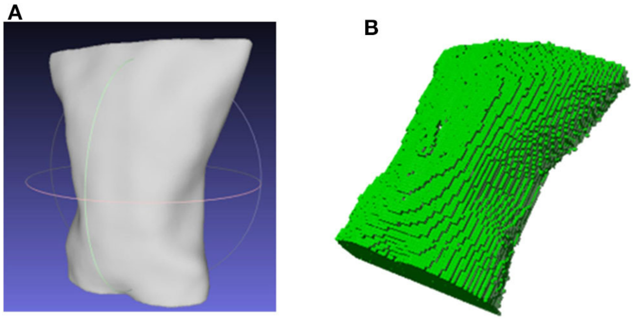


FIGURE 5 | 3D model of thoracoabdominal region. **(A)** Surface reconstruction. **(B)** Voxel model.

surface information of the model but also describes the internal properties of the model. We used Octomap library (Hornung et al., 2013) to transform the point cloud into a voxel model. The voxelization of the model is shown in **Figure 5B**.

Feature Vector of Dynamic Voxel Model

Due to the huge dimension of voxel model (arranging each voxel unit in columns up to several 100,000 dimensions), it is unrealistic to directly use these data to correlate with tumor motion because of a lot of noise and redundant information in huge data. In order to more effectively correlate with tumor motion and fully reflect surface motion characteristics, we first extract its physical significance features, then perform data analysis on the voxel model to extract respiratory features.

Physical Characteristics

The movement of breathing is accompanied by the heaving of the chest and abdomen. Intuitively speaking, the expansion and contraction of the chest and abdomen must cause the changes of the entire volume V and surface area S of thoracoabdominal trunk.

The entire voxel model's voxel number variation with time series reflects the volume change of the thoracoabdominal trunk. To study the respiration characteristic displayed by volume changes, the voxel units of voxel model of each frame were traversed, and then, the volume of voxel units was accumulated; therefore, the volume change is reflected in the time series. The area change of voxel model is reflected in the area of voxel's outer layer in the similar way.

Intrinsic Data Characteristics

Intrinsic data characteristics refer to the information that reveals the motional information of the body surface obtained by the voxel model. It is obtained by reducing the dimensionality of the voxel model. It is necessary for the construction of the respiration tracking model.

Just as point cloud generation came with time stamps, the state of voxel model varied over time series. Each frame of the voxel model stood for a respiratory state, which was expressed by probability of voxel occupancy. The overview of the proposed method was presented in Algorithm 1.

Algorithm 1: | Construct Voxel-Bounding Box

```
Input : $M(t)$ − $n$-frame voxel model
Output : $B'(t)$ − $n$-frame modified-bounding box
for $M_i \in M(t)$ do
    Set the probability of each voxel in $M_i$ to 1
    Calculate the length, width, and height of the bounding box of $M_i$ as: $L_{Mi} \in L$, $W_{Mi} \in W$, $H_{Mi} \in H$
end for
Search for the maximum value of length, width and height respectively as $L_{\max}, W_{\max}, H_{\max}$, to form a template-bounding box $B$
Divide $B$ into space units that decided by the resolution of each voxel
for $M_i \in M(t)$ do
    Compare the template-bounding box $B$ and $M_i$ to obtain their intersections. In template-bounding box $B$, set the probability of the intersection voxels to 1 and set those of the rest voxels in to 0, to obtain the modified bounding box $B_i'$
end for
```

Depending on the voxels' occupation and idleness, the state of voxels in the template-bounding box were marked with probability 1, 0, respectively. Thus, the voxel model state varied by time was expressed by the marker value of each voxel in the template-bounding box, as shown in the left diagram in **Figure 6**. Since the bounding box contained a large number of voxels, which could be illustrated as columns of a superhigh-dimensional vector, it would cost huge calculation if the original vectors were used to build the correlation model. Therefore, the vector Υ with superhigh dimension was transformed into a low-dimensional vector ψ that remained the characteristics of the voxel model changes. To accomplish dimension reduction, an algorithm based on LLE (Roweis, 2000) is shown in Algorithm 2.

Through combining physical variables V, S and essential parameters $\psi = [\psi_1, \psi_2, .., \psi_m] (m \leq d+1)$ after data analysis, characteristic variables Γ that can express time-varied states of the voxel model can be obtained as:

$$\Gamma = [V, S \ldots \psi_1, \psi_2, .., \psi_m]^T \tag{4}$$

Correlation Between Characterization Vector and Tumor Motion

After feature extraction of voxel model in previous section, we obtained the representation vector of the whole thoracoabdominal surface motion. The establishment of respiratory correlation model is a key part of tumor tracking. The correlation function is to take the breathing surrogate signal (the characteristic signal of skin surface motion that is highly correlated with tumor motion in the body) as the input data and realize the motion estimation of the internal tumor by correlating the surrogate with the movement of the internal tumor. Therefore, we used the extracted representation vector as external surface motion surrogate to establish a correlation model with internal tumor motion.

Algorithm 2: | Respiration Representation Based on LLE

```
Input : $B'(t)$ − $n$-frame modified-bounding box
Output : $\psi$ −Low-dimensional vector
for $B_i' \in B'(t)$ do
    Define a column vector $x_i$ by the probability arrangement in a certain spatial order of each voxel in bounding box $B_i'$, $x_i \in X$
end for
Obtain a sample space $X = \{x_1, x_2, \ldots, x_n\}$ as input for dimensionality reduction
for $x_i \in X$ do
    Search for the $K$ nearest neighbors of each sample $x_i$
    Construct the loss function of the mean square error of $x_i$ and the linear representation of its $K$ nearest neighbors and normalize the weight coefficients
    Solve the optimal solution of the loss function as the weight coefficient vector $W_i$ of each sample $x_i$
end for
Weight coefficient matrix $W$ is composed of obtained $W_i$
Calculate $M = (I - W)(I - W)^T$ and solve the eigenvalues and eigenvectors of $M$
Choose the optimal dimension $d$ considering the contribution rate of eigenvalue and obtain the sample matrix with lower dimension as $\psi = \{\psi_2, \psi_3, \ldots, \psi_{d+1}\}$
```

Tumor movement is complex since it is caused by various factors such as respiration, heartbeat, and abdominal pressure. During inhalation, the volume of air inhaled by the lungs becomes larger because of the thoracoabdominal cavity and the diaphragm deform. During the exhalation process, the volume of gas remaining in the lungs becomes smaller; at the same time, the thoracoabdominal cavity and diaphragm are restored to their original state. During normal breathing, at the same transpulmonary pressure, the expiratory volume is greater than the inspiratory volume. The so-called hysteresis phenomenon (the phase lag between body surface characteristic respiratory motion and tumor respiratory motion) is attributed to the complex respiratory pressure–volume relationship of the lungs and chest and abdomen.

Therefore, exhalation and inhalation are two irreversible motion states. Before establishing the correlation model, we need to divide the entire breathing process into exhalation part and inhalation part and model these two processes, respectively. This paper divides the respiratory phase according to the amplitude of the external motion and uses the peak and valley values of the motion as the dividing points of exhalation and inspiration.

Due to the non-linear relationship between internal (the motion information of tumor *in vivo*) and external motion (the motion information of body surface), we adopt a polynomial model (Peressutti et al., 2012). That is, the trajectory of the tumor *in vivo* is approximated as a linear combination of multiple power terms of the external signal. In this paper, different polynomial functions are used to model the breathing movement during the exhalation and inhalation phases:

$$X_{T_i} = \begin{cases} \sum\limits_{j=0}^{N} A_j^+ r_i^j, r \geq k_i \\ \sum\limits_{j=0}^{N} A_j^- r_i^j, r \leq k_i \end{cases} \tag{5}$$

in which A_j is a polynomial coefficient, N is the highest power, and k_i is the dividing points of each expiratory and inspiratory process. The highest power of the polynomial model is preferably 2 or 3. Higher powers are prone to overfitting and reduce the generalization ability of the model.

EXPERIMENTS AND RESULTS

Experiments have been carried out to verify the feasibility of the respiratory motion characterization method based on thoracoabdominal surface modeling. Our experiments mainly focused on two aspects as validity of thoracoabdominal surface described by point cloud and validity of feature extraction on surface–tumor correlation. *Position Comparison of Multiple Markers* and *Comparison of Marker Position Obtained by Point Cloud and NDI Method* validated that whether the representation method of skin surface modeling has the same effect as the traditional representation method of finite markers; *Optimization of Dimension Reduction* and *Correlation Coefficient of Dimension-Reduced Feature Vector and NDI Markers With Tumor Motion* validated whether this method has more comprehensive characterization capability than the traditional finite marker method; in *Correlation Model of Dimension-Reduced Feature and Tumor*, correlation model between dimension-reduced features of thoracoabdominal voxel data and tumor motion was built and compared with the traditional method.

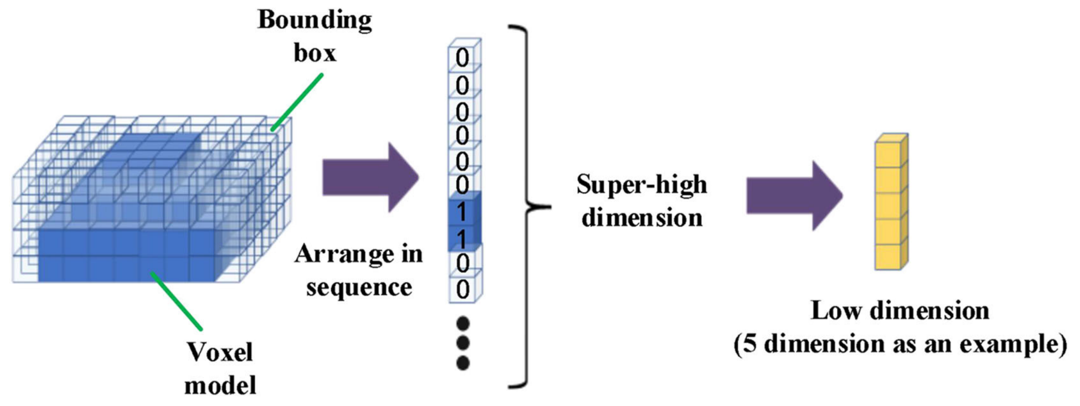


FIGURE 6 | Flowchart of voxel model dimensionality reduction.

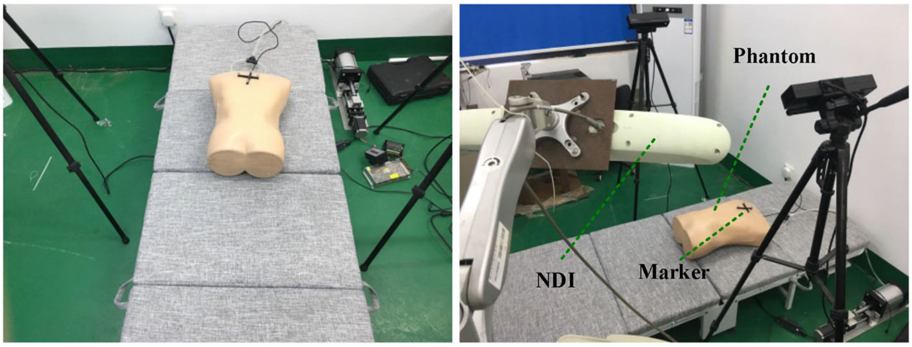


FIGURE 7 | Scene of NDI marker experimental data acquisition.

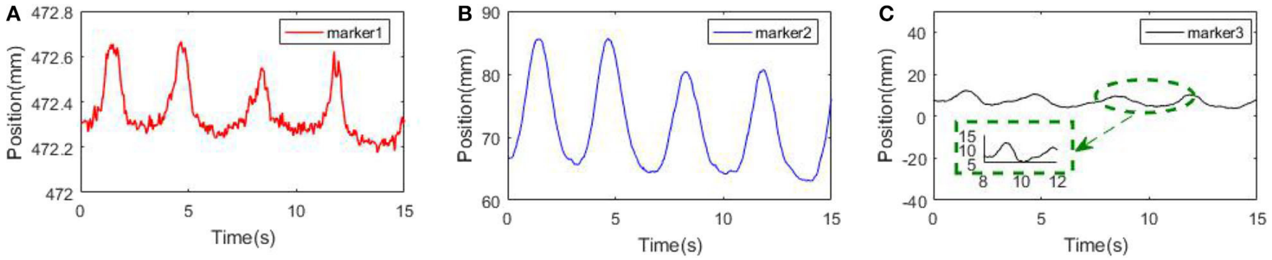


FIGURE 8 | Principal component motion of surface markers. **(A)** Marker 1. **(B)** Marker 2. **(C)** Marker 3.

We used two Kinect V2 (Microsoft Co.) depth cameras placed at both sides of the experimental bed to collect data. Because the view of a single camera is limited, blind areas cannot be observed on the abdomen exit and will prevent a successful 3D modeling of thoracic–abdominal surface. However, using two Kinect V2 cameras has a complete view of the patient's body and is conducive to building the whole thoracic–abdominal model. The experimental subject is a phantom developed in our lab (Hou et al., 2018) for simulating respiratory movement of thoracoabdominal cavity. The experimental scene is shown in **Figure 7**. In order to compare the skin surface characterization method with traditional markers, we used NDI Polaris Spectra (Northern Digital Inc.), an optical tracking system (accuracy, 0.25 mm RMS) to record the skin markers' movement. Before experiments, we need to calibrate the coordinate systems of NDI Polaris Spectra and Kinect V2 depth cameras. To unify the two coordinate systems of point cloud and NDI system, spherical center of NDI optical markers in point cloud data was fitted to match with the marker's coordinates obtained by NDI Polaris Spectra.

Position Comparison of Multiple Markers

In order to display the differences among multiple markers in characterizing the respiratory movement, the position of the markers on the chest and abdomen surface is changed, and the NDI sensor is utilized to track the motion change. **Figure 8** shows the principal movement of the markers analyzed by PCA at different positions. Although the frequency and phase of the three

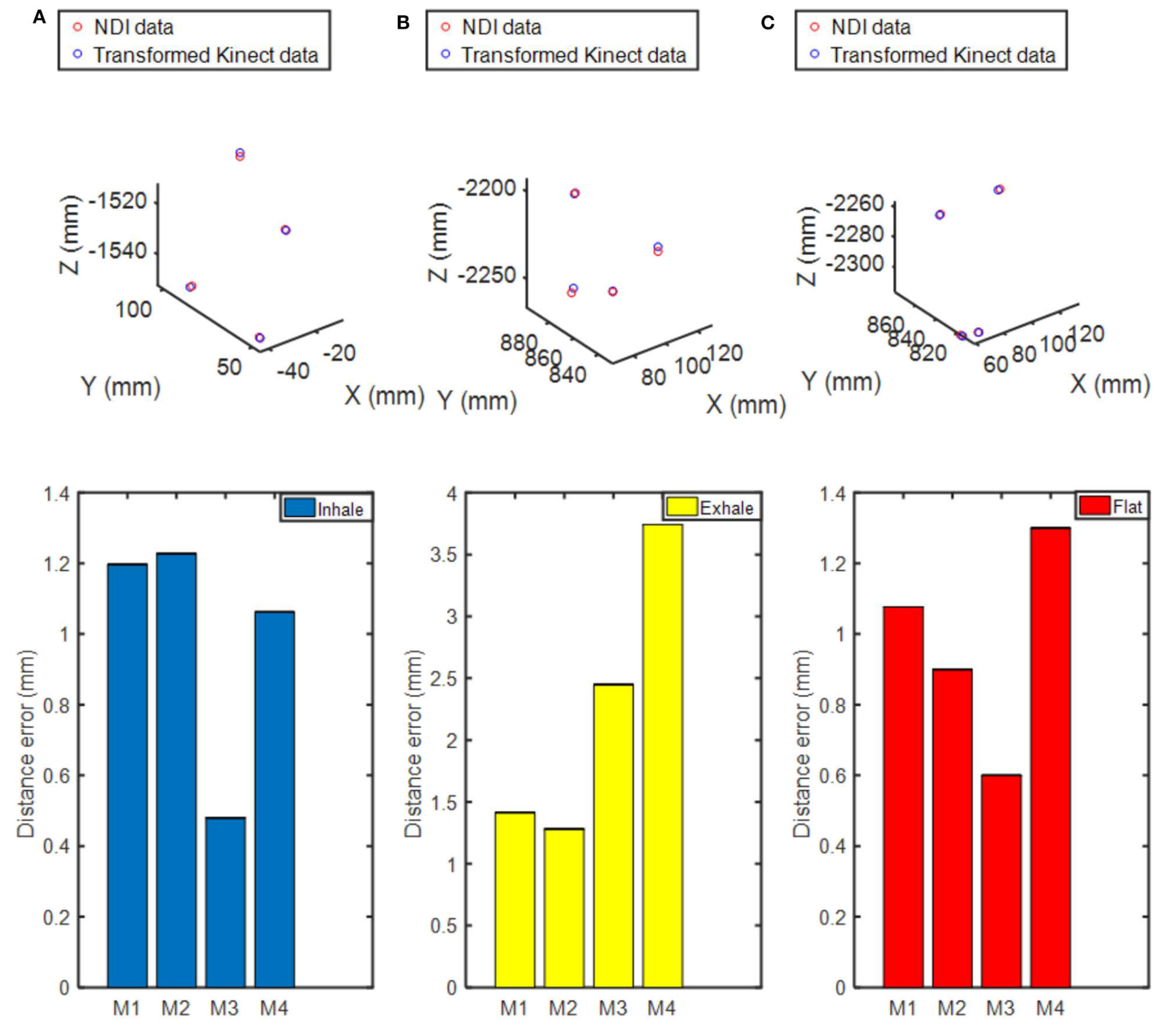


FIGURE 9 | Comparison of the spatial position of the markers in the depth camera coordinate system and the NDI system coordinate system. **(A)** Inhalation. **(B)** Exhalation. **(C)** Transition between inhalation and exhalation. The uppermost figure describes the spatial position of the NDI markers and the transformed ones from point cloud data in the NDI system coordinate. The figure below shows the distance error between the transformed position and the actual NDI position.

markers are the same, the maximum motion range of marker 1 is about 5 mm, the maximum motion range of marker 2 is about 2 cm, and the maximum motion range of marker 3 is about 1 cm. The markers at different positions reflect different movements. It can be inferred that limited skin surface markers are not complete in motion representation of the whole thoracoabdominal surface.

Comparison of Marker Position Obtained by Point Cloud and NDI Method

To clarify whether our characterization method based on thoracoabdominal surface modeling contains the information provided by traditional finite markers, we made comparisons of the same marker position obtained by Kinect V2 and NDI optical tracking system. Before the comparison experiments, we need to unify the point cloud data into the NDI coordinates system. Therefore, calibration of NDI coordinates system and Kinect V2 coordinate system was carried out as follows.

Suppose the transformation equation of the two coordinate systems is $A = RB + T$, where A is the position of one point in coordinates of NDI, R is the rotation matrix, B is the coordinates of the same point in the point cloud coordinate system, and T is the translation matrix. To solve R and T, three markers were placed on the chest of the phantom for five times, each time with different positions, respectively. NDI optical tracking system and Kinect V2 recorded data simultaneously at the five experiments. After collecting the data, NDI data were used to calculate the corresponding transformation matrix R, T. Then, the point cloud obtained by Kinect V2 system was compared with the position of the marker under the NDI sensor.

Center positions of four NDI markers' by fitting of point cloud obtained by Kinect V2 system were transformed and compared with the position of NDI markers under NDI sensor at three breathing fractions. As shown in **Figure 9**, there are three respiratory states: inhalation, exhalation, and transition between

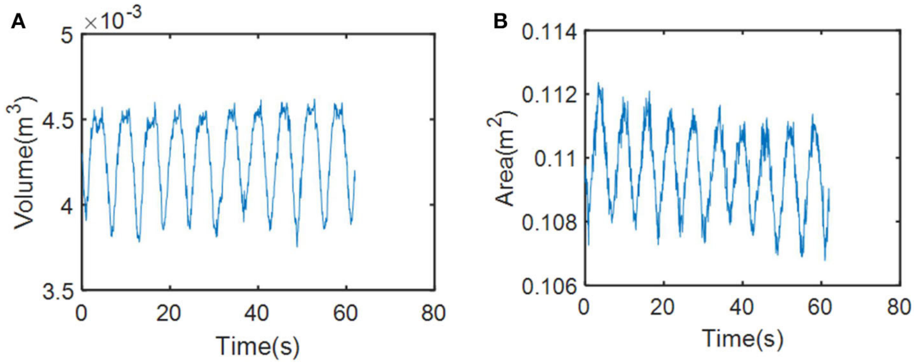


FIGURE 10 | Changes in respiratory motion of physical characteristics. **(A)** Changes in respiratory motion of volume feature. **(B)** Changes in respiratory motion of area feature.

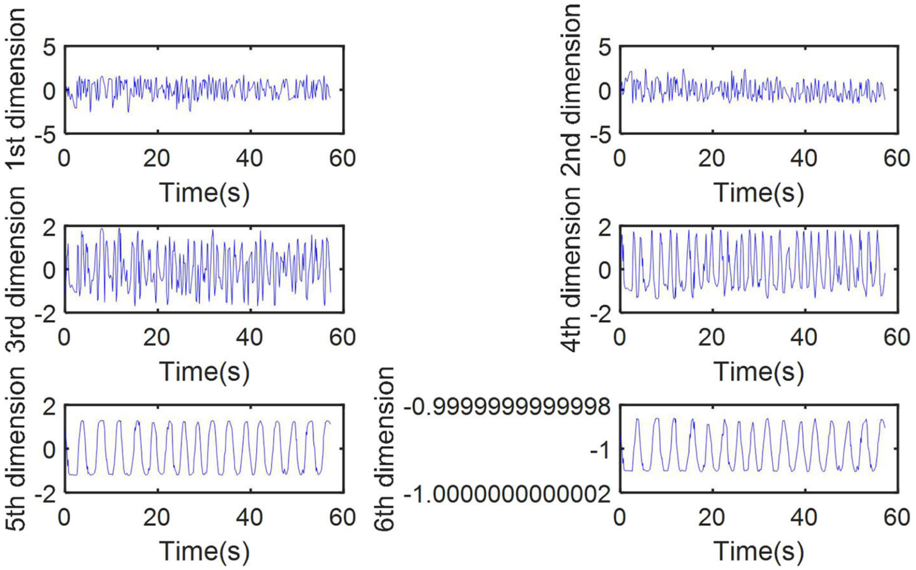


FIGURE 11 | Respiratory motion features of thoracoabdominal model in six dimensions after reduction.

TABLE 1 | The correlation coefficient between each dimension and tumor motion.

	Pearson correlation coefficient									
Phantom data	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**
P1	0.0427	0.0135	0.0912	0.0363	0.0119	0.0177	0.1188	0.0353	0.0141	0.9230
P2	0.0262	0.0160	0.0521	0.0023	0.0940	0.0355	0.1030	0.1926	0.0039	0.9050
P3	0.0872	0.0393	0.0311	0.0690	0.1734	0.1176	0.0280	0.0215	0.0476	0.9491

The Pearson correlation coefficient between each dimension and tumor motion is calculated in the table.

TABLE 2 | The correlation coefficient between the dimensionality reduction variable and NDI marker.

	Pearson correlation coefficient	
Phantom data	**NDI marker to tumor**	**Dimension-reduced feature to tumor**
P1	0.8553	0.9141
P2	0.8369	0.9039
P3	0.8419	0.9476

The Pearson correlation coefficient between each parameter and tumor motion is calculated in the table.

inhalation and exhalation. The uppermost figure represents the position relationship between the two coordinates in the three-dimensional space after conversion. By comparing the Euclidean distance between the two coordinates, it can be seen from the figure that the maximum distance error between the two coordinates is within 4 mm. The coordinates of the two can basically correspond to each other through conversion. It can be concluded that the characterization method based on thoracoabdominal surface modeling has the same effect as the traditional characterization method of finite markers.

Optimization of Dimension Reduction

Unlike the traditional method of representing several isolated points, the advantage of our method lies in characterizing the entire surface feature. We collected respiratory motion data from our phantom for 1 min and obtained voxel models based on time series through the modeling method described in *Dynamic Thoracoabdominal Surface Voxel Modeling*. **Figure 10** shows the volume and area features, respectively, extracted by using the method mentioned above. We can see that volume and area characteristics display periodicity and rhythmicity and conform to the law of respiratory movement. Then, we used dimension reduction analysis of voxel model proposed in *Feature Vector of Dynamic Voxel Model*. The bounding box contains 172,530 voxel units with a resolution of 5 mm and constitutes a 172,530-dimensional column vector, which can be reduced to any dimension. For example, we reduced the high-dimensional data to a characteristic vector with six dimensions. The respiratory characteristics in six dimensions are shown in **Figure 11**.

Whether the results of dimensionality reduction can accurately reflect the respiratory movement of the surface needs to be validated. In addition, optimal number of the characteristic vector after dimension reduction needs to be ascertained. To solve these problems, Pearson correlation method (Person, 1895) was used for correlation analysis in this paper. The formula is as follows:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{6}$$

where the Pearson correlation coefficient $\rho_{X,Y}$ of two continuous variables (X, Y) is equal to the covariance $cov\,(X, Y)$ between them divided by the product of their respective standard deviations (σ_X, σ_Y). The value of the coefficient is always between -1.0 and 1.0. Variables close to 0 are regarded as uncorrelated, while variables close to 1 or -1 are regarded as strongly correlated.

Three sets of phantom data were, respectively, used to build voxel models and were reduced into characteristic vectors with 10 dimensions first. Then, correlation analysis between each dimension and tumor movement *in vivo* is carried out using Equation (6). As shown in **Table 1**, there is only one dimension with a correlation coefficient >0.5 with tumor motion *in vivo*. The correlation coefficient between this only dimension and tumor *in vivo* reached more than 0.9. Therefore, we can conclude that reducing the superhigh-dimensional voxel model to 1 dimension is the optimal operation.

Correlation Coefficient of Dimension-Reduced Feature Vector and NDI Markers With Tumor Motion

In order to verify the correlation between each variable and tumor movement, we also calculated the correlation coefficient between tumor motion, dimension-reduced data, and NDI markers, respectively, by using Pearson method as Equation (6). The results are in **Table 2**. The correlation coefficient between the dimension-reduced feature vector and tumor motion is higher than that of skin surface markers and tumor motion. To sum up, the movement of chest and abdomen surface can be featured by one-dimensional feature vector. Besides, fusion of multidimensional voxel model data represents movement of chest and abdomen surface more accurately than markers.

Correlation Model of Dimension-Reduced Feature and Tumor

As the correlation model has been introduced in the previous section, one-dimensional data were obtained by the dimension reduction method to establish the correlation model with tumor motion *in vivo*.

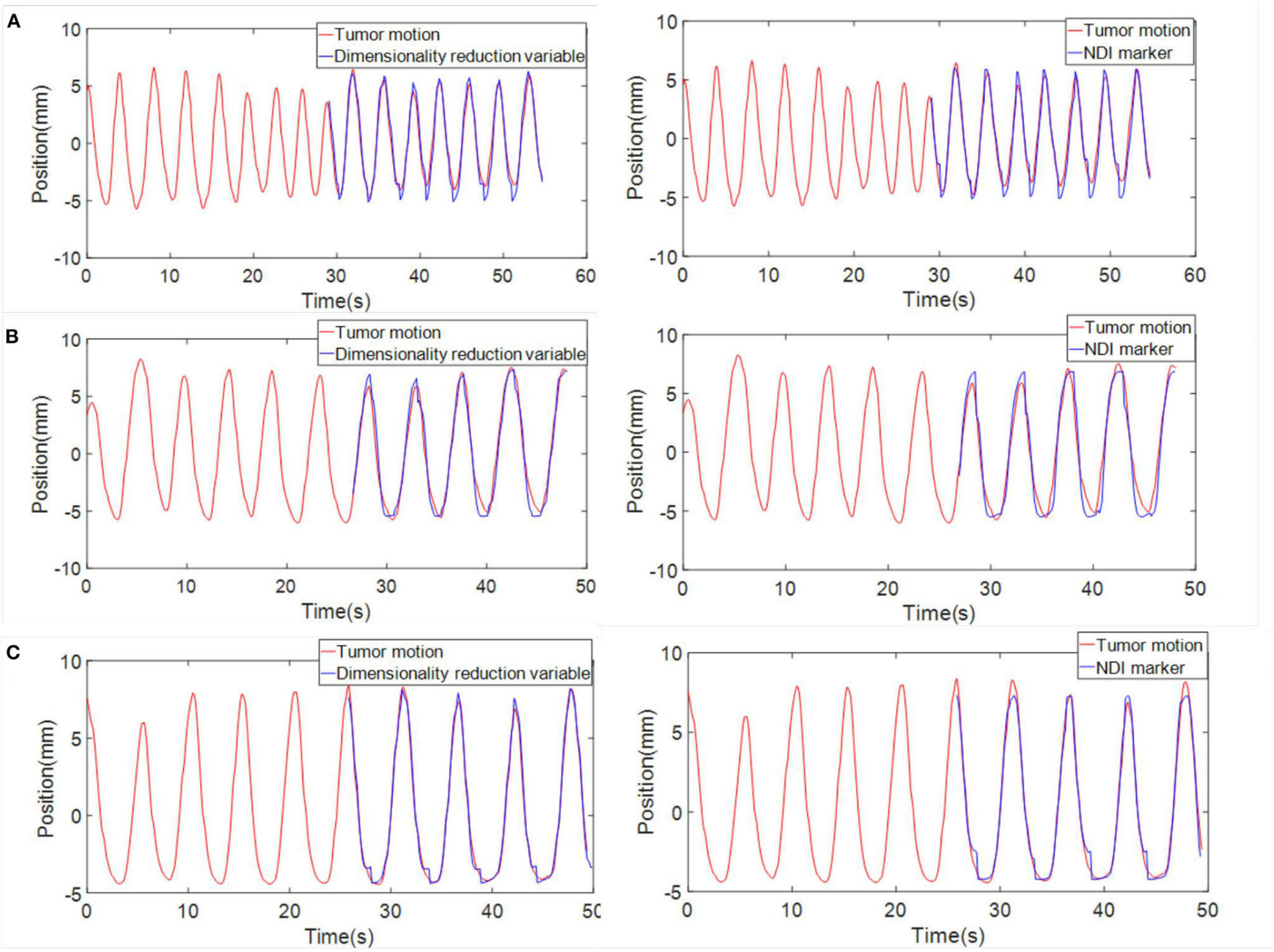


FIGURE 12 | Comparison between the tumor motion estimation and their actual value (red, actual value; blue, estimated value). **(A)** Phantom data 1. **(B)** Phantom data 2. **(C)** Phantom data 3. Three sets of data are displayed. Curves in the left column show the estimation of tumor motion obtained through the correlation model built with dimensionality reduction results. Curves in the right column shows the estimation of tumor motion when NDI marker position was chosen as the input of the correlation model.

TABLE 3 | The mean squared error (MSE) error between the estimated and actual values.

Phantom data	MSE error(mm)	
	NDI Marker	Dimensionality reduction variable
1	0.0136	0.0117
2	0.0082	0.0055
3	0.0051	0.0037

The obtained data (last for 1 min) is divide into two parts: one part was used to establish the correlation model, and the other one was used to verify the accuracy of the correlation model. That is, the correlation model was used to estimate the tumor motion *in vivo* using skin surface motion, while the feasibility and accuracy of the method were proved by comparing the error between the actual tumor motion and the estimated value.

Then, we established the correlation model with tumor motion by using the dimensionality reduction results and NDI marker positions, respectively. After that, correlation models were used to estimate tumor motion. Three sets of phantom data with point cloud collection and NDI maker positions were used to test and compare the two correlation models. **Figure 12** shows the comparison results. The mean squared error (MSE) between the estimated and actual values was also calculated and presented in **Table 3**. From **Table 3**, it can be seen that the accuracy of the correlation model obtained by dimension reduction results is much higher than the one of the correlation model obtained by NDI marker, which also indicates that the proposed approach in this work is feasible to extract respiratory features based on horacoabdominal surface modeling, and it is capable to establish an enhanced respiratory correlation model than skin surface marker method.

DISCUSSION AND CONCLUSION

In this paper, a novel characterization method of human external respiratory motion for tumor correlation was proposed. Compared with previous methods of limited markers and multiple modality surrogates, our method accomplished better characterization by establishing a 3D voxel model containing rich structural information of thoracic and abdominal torso during the respiratory motion. Dimensions of the voxel model were reduced by a dimensionality reduction method. Through these procedures, features that can reflect changes in external respiratory motion were obtained. This paper builds the model of chest and abdomen torso's respiratory motion, from multicamera system construction, point cloud processing, and watertight model construction to model voxelization. A dimensionality reduction method was proposed to extract respiratory features from the established voxel model. The feasibility of the characterization method based on chest and abdomen surface modeling was verified and compared with the traditional finite surface marker representation method. The proposed method was proved to be more accurate than the traditional one with limited external markers. Furthermore, it breaks through the limitation of information loss in current respiration correlation methods, which describe the respiratory motion by using sparse sensing data. In future work, to form a sound solution of respiration tracking for radiosurgical robots, we will optimize the process of thoracoabdominal skin surface modeling, study the feature extraction of voxel model, and apply this method to the prediction of tumor motion under the clinical application of tumor tracking.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

SY contributed to the manuscript writing, data analysis, and experiment design. PH contributed to the manuscript writing and data analysis. RS conceived the presented idea and contributed to the algorithm design. MZ and JG contributed to the manuscript revision and system design. FZ, SK, and LS contributed to the support of experimental equipment. All authors discussed the results and contributed to the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China Study on three-dimensional reconstruction of body surface and tumor motion tracking for stereo radiotherapy robot based on point cloud data fusion (project number 61773273). This work was supported by grants from the National Natural Science Foundation of China (nos. 61673288 and 61773273).

ACKNOWLEDGMENTS

The authors would like to acknowledge the contribution of multiple volunteers to the experimental data and Professor Zhang toward the experimental equipment. The authors would also like to acknowledge Jiateng Wang for manuscript revision and the work of experiments.

REFERENCES

Breitkopf, P., Naceur, H., Rassineux, A., and Villon, P. (2005). Moving least squares response surface approximation: formulation and metal forming applications. *Comput. Struct.* 83, 1411–1428. doi: 10.1016/j.compstruc.2004.07.011

Cheng, W., and Adler, J. (2006). An overview of cyberknife radiosurgery. *Chin. J. Clin. Oncol.* 3, 229–243. doi: 10.1007/s11805-006-0049-5

Coste-Maniere, E., Olender, D., and Kilby, W., Schulz, R. A. (2005). Robotic whole body stereotactic radiosurgery: clinical advantages of the cyberknife integrated system. *Int. J. Med. Rob. Comput. Assist. Surg.* 1, 28–39. doi: 10.1002/rcs.39

Durichen, R., Davenport, L., Bruder, R., Wissel, T., Schweikard, A., and Ernst, F. (2013). "Evaluation of the potential of multi-modal sensors for respiratory motion prediction and correlation," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka), 5678–5681. doi: 10.1109/EMBC.2013.6610839

Ernst, F., and Saß, P. (2015). Respiratory motion tracking using microsoft's kinect v2 camera. *Curr. Dir. Biomed. Eng.* 1, 192–195. doi: 10.1515/cdbme-2015-0048

Ernst, F., Schlaefer, A., and Schweikard, A. (2009). Smoothing of respiratory motion traces for motion-compensated radiotherapy. *Med. Phys.* 37, 282–294. doi: 10.1118/1.3271684

Gu, C., and Li, C. (2015). Assessment of human respiration patterns via noncontact sensing using doppler multi-radar system. *Sensors* 15, 6383–6398. doi: 10.3390/s150306383

Gu, C., Li, R., Jiang, S. B., and Li, C. (2011). "A multi-radar wireless system for respiratory gating and accurate tumor tracking in lung cancer radiotherapy," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Boston, MA), 417–420. doi: 10.1109/IEMBS.2011.6090054

Hoisak, J. D. P., Sixel, K. E., Tirona, R., Cheung, P. C. F., and Pignol, J.-P. (2004). Correlation of lung tumor motion with external surrogate indicators of respiration. *Int. J. Radiat. Oncol. Biol. Phys.* 60, 1298–1306. doi: 10.1016/j.ijrobp.2004.07.681

Hornung, A., Wurm, K. M., Bennewitz, M., and Stachniss, C., Burgard, W. (2013). OctoMap: an efficient probabilistic 3D mapping framework based on octrees. *Auton. Rob.* 34, 189–206. doi: 10.1007/s10514-012-9321-0

Hou, P., Sun, R., and Yu, S. (2019). "Feasibility of respiratory motion characterization based on skin-tumor correlated thoracoabdominal surface modeling," in *Proc. IEEE Int. Conf. Intelligent Robots and Systems* (Macau).

Hou, P., Sun, R., Yu, S., and Sun, L. (2018). "Design of a holistic thoracoabdominal phantom for respiration tracking test in robotic radiosurgery," in *3rd*

International Conference on Advanced Robotics and Mechatronics (Singapore), 318–322. doi: 10.1109/ICARM.2018.8610803

Iizuka, Y., Matsuo, Y., Ishihara, Y., Akimoto, M., Tanabe, H., Takayama, K., et al. (2015). Dynamic tumor-tracking radiotherapy with real-time monitoring for liver tumors using a gimbal mounted linac. *Radiother. Oncol.* 117, 496–500. doi: 10.1016/j.radonc.2015.08.033

Kanoulas, E., Aslam, J. A., Sharp, G. C., Berbeco, R. I., Nishioka, S., Shirato, H., et al. (2007). Derivation of the tumor position from external respiratory surrogates with periodical updating of the internal/external correlation. *Phys. Med. Biol.* 52, 5443–5456. doi: 10.1088/0031-9155/52/17/023

Kazhdan, M., Bolitho, M., and Hoppe, H. (2006). "Poisson surface reconstruction," in *Proceedings of Symposium on Geometry Processing* (Cagliari), 61–70.

Koch, N., Liu, H. H., Starkschall, G., Jacobson, M., Forster, K., Liao, Z., et al. (2004). Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: part I—correlating internal lung motion with skin fiducial motion. *Int. J. Radiat. Oncol. Biol. Phys.* 60, 1459–1472. doi: 10.1016/j.ijrobp.2004.05.055

Liu, Y., Jiang, P., Petrosyan, V., Li, S., Bian, J., Zhang, L., et al. (2018). "DEL: deep embedding learning for efficient image segmentation," in *27th Proceedings of International Conference on Artificial Intelligence Main track* (Stockholm), 864–870. doi: 10.24963/ijcai.2018/120

Moore, J. B. (1978). Statistical filtering. *Math. Control Theory* 680, 152–169. doi: 10.1007/BFb0065316

Peressutti, D., Rijkhorst, E. J., Barratt, D. C., Penney, G. P., and King, A. P. (2012). "Estimating and resolving uncertainty in cardiac respiratory motion modelling," in *9th IEEE International Conference Symposium on Biomedical Imaging* (Barcelona), 262–265. doi: 10.1109/ISBI.2012.6235534

Person, K. (1895). Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* 58, 240–224. doi: 10.1098/rspl.1895.0041

Roweis, S. T. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science* 290, 2323–2326. doi: 10.1126/science.290.5500.2323

Schweikard, A., Glosser, G., Bodduluri, M., Murphy, M. J., and Adler, J. R. (2000). Robotic motion compensation for respiratory movement during radiosurgery. *Comput. Aided Surg.* 5, 263–277. doi: 10.3109/10929080009148894

Tomasi, C., and Manduchi, R. (1998). "Bilateral filtering for gray and color images," in *1998 IEEE International Conference on Computer Vision* (Bombay), 839–846. doi: 10.1109/ICCV.1998.710815

Torshabi, A. E., Riboldi, M., Fooladi, A. A. I., Mosalla, S. M. M., and Baroni, G. (2013). An adaptive fuzzy prediction model for real time tumor tracking in radiotherapy via external surrogates. *J. Appl. Clin. Med. Phys.* 14, 102–114. doi: 10.1120/jacmp.v14i1.4008

Transue, S., Nguyen, P., Vu, T., and Choi, M.-H. (2016). "Real-time tidal volume estimation using iso-surface reconstruction," in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (Washington, DC), 209–218. doi: 10.1109/CHASE.2016.72

Wang, X., Xie, Q., Ma, T., and Zhu, J. (2018). Feature extraction based on dimension reduction and clustering for maize leaf spot images. *Int. J. Pattern Recognit. Artif. Intell.* 32:1854029. doi: 10.1142/S0218001418540290

Wijenayake, U., and Park, S.-Y. (2017). Real-time external respiratory motion measuring technique using an RGB-D camera and principal component analysis. *Sensors* 17:1840. doi: 10.3390/s17081840

Yang, J., Chen, L., Zhang, L., Sun, X., She, D., Lu, S.-P., et al. (2018). "Historical context-based style classification of painting images via label distribution learning," in *2018 ACM Multimedia Conference on Multimedia Conference* (Amsterdam), 1154–1162. doi: 10.1145/3240508.3240593

Zhang, L., Peng, S., and Winkler, S. (2019). PersEmoN: a deep network for joint analysis of apparent personality, emotion and their relationship. *IEEE Trans. Affect. Comput.* 1–1. doi: 10.1109/TAFFC.2019.2951656

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

METHODS
published: 16 November 2020
doi: 10.3389/fnbot.2020.579338

A New Feature Extraction and Recognition Method for Microexpression Based on Local Non-negative Matrix Factorization

Junli Gao [1], Huajun Chen [1], Xiaohua Zhang [2]*, Jing Guo [1] and Wenyu Liang [3]

[1] School of Automation, Guangdong University of Technology, Guangzhou, China, [2] College of Automation, Zhongkai University of Agriculture and Engineering, Guangzhou, China, [3] Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

Microexpression is usually characterized by short duration and small action range, and the existing general expression recognition algorithms do not work well for microexpression. As a feature extraction method, non-negative matrix factorization can decompose the original data into different components, which has been successfully applied to facial recognition. In this paper, local non-negative matrix factorization is explored to decompose microexpression into some facial muscle actions, and extract features for recognition based on apex frame. However, the existing microexpression datasets fall short of samples to train a classifier with good generalization. The macro-to-micro algorithm based on singular value decomposition can augment the number of microexpressions, but it cannot meet non-negative properties of feature vectors. To address these problems, we propose an improved macro-to-micro algorithm to augment microexpression samples by manipulating the macroexpression data based on local non-negative matrix factorization. Finally, several experiments are conducted to verify the effectiveness of the proposed scheme, which results show that it has a higher recognition accuracy for microexpression compared with the related algorithms based on CK+/CASME2/SAMM datasets.

Keywords: macro-expression, micro-expression, macro-to-micro transformation, feature extraction, non-negative matrix factorization, CK+/CASME2/SAMM datasets

Edited by:
Zhan Li,
University of Electronic Science and Technology of China, China

Reviewed by:
Yangsong Zhang,
Southwest University of Science and Technology, China
Wellington Pinheiro dos Santos,
Federal University of Pernambuco, Brazil

****Correspondence:***
Xiaohua Zhang
xzhang69@163.com

Received: *02 July 2020*
Accepted: *29 September 2020*
Published: *16 November 2020*

Citation:
Gao J, Chen H, Zhang X, Guo J and Liang W (2020) A New Feature Extraction and Recognition Method for Microexpression Based on Local Non-negative Matrix Factorization. Front. Neurorobot. 14:579338. doi: 10.3389/fnbot.2020.579338

1. INTRODUCTION

Expression is one of the important ways for human to communicate emotion. In 1970s, American psychologist Paul Ekman defined six basic expressions of human, namely, happiness, anger, surprise, fear, disgust, and sadness. Facial expression recognition is to extract the specific states from given images or video, then identify the psychological emotions of the recognized object and understand its facial expressions. Expression recognition has many applications in psychology, intelligent monitoring, robotics, etc. Moreover, sometimes people may disguise their emotion and expression for various purposes. However, people cannot completely suppress their emotions under external strong emotional stimulus. There are some subtle and fast facial actions, which were first discovered and named "micro-momentary" expressions by Haggard and Isaacs (1966). Then, Ekman and Friesen formally named them microexpressions (Ekman and Friesen, 1969). It is an uncontrolled expression, which can be classified into the six basic emotions (Wu and Fu, 2010).

Different from the general expression, the microexpression is only reflected in a few facial action units, and the duration is about $\frac{1}{12}$ to $\frac{1}{2}$ s, which is difficult to detect. In addition, microexpression usually appears when people try to cover up their emotions. It is one kind of subconscious, inside-to-outside, uncontrolled, and undetectable behaviors with naked eyes. Because microexpression cannot be hidden, people can exploit this kind of weak, partial, short-term behaviors to acquire the hidden real emotions. Notably, microexpression recognition has many valuable applications in psychology, clinical diagnosis, business negotiation, interrogation, human–robot interaction, and so on, but it needs special training to master relevant recognition skills. In 2002, Ekman developed microexpression training tool (METT) (Ekman, 2006) that can train the recognition skills of the six basic emotions and other kinds of expressions, such as contempt, pain, and so on. At first, the recognition abilities of observers are tested through METT, then related knowledge of microexpression recognition is taught. After repeated training and consolidation, the recognition accuracy of observers can be improved by 40%. Nevertheless, the accuracy may be affected by various subjective factors, such as mood or preconceived thinking of the observers.

Microexpression is weak, short term, and difficult to detect, so the traditional expression recognition algorithms do not work well at all for this task. Generally, microexpression recognition can be divided into detection and classification. The former is to determine whether there are microexpressions in an image sequences, and detect the start/apex/end frames of a microexpression. The latter includes feature extraction and classification, which is similar to the general tasks of pattern classification. Significantly, the feature extraction is to acquire the abstract information from the data, which usually is some vectors obtained by image processing. The related algorithms can be used for extracting features, which can reflect the microexpression action information and distinguish various kinds of emotions. The feature classification is to train a classifier on the obtained vectors, directly related to the recognition accuracy, to distinguish the types of microexpression.

The main contributions of this paper are summarized as follows: (i) A local non-negative matrix factorization (LNMF) is developed to extract the features of apex frame on microexpression, which exploits local properties of LNMF to reflect the features of local action on microexpression. (ii) An improved macro-to-micro (MtM) transformation algorithm is proposed to augment the samples of microexpressions from macroexpression data based on LNMF. (iii) The performance of the proposed scheme is verified on CK+, CASME2, and SAMM datasets, which can benefit this work on human–robot interaction.

The rest of the paper is organized as follows. Related works are discussed in section 2. In section 3, the overall scheme, including theoretical derivation on LNMF and MtM algorithm design, is presented. Section 4 provides the experimental process and result analysis. Finally, we conclude this paper in section 5.

2. RELATED WORK

Local binary pattern (LBP) is a commonly used method for extracting texture feature of images. LBP from three orthogonal planes (LBP-TOP) is an extension of LBP in video data. Ojala et al. (2002) and Zhao and Pietikainen (2007) acquired the feature vectors of the whole video by extracting the XY, XT, YT plane features of video. Yan et al. (2014) used LBP-TOP to extract the features of cropped face video in CASME2, and take support vector machine (SVM) as the classifier to recognize five categories of expressions with an accuracy of 63.41%. To reduce information redundancy and computational complexity of LBP-TOP, Wang et al. proposed LBP six intersection points (LBP-SIP) feature extraction algorithm (Wang Y. et al., 2014). Ben et al. (2018) proposed the second-order descriptor hot wheel patterns from TOP (HWP-TOP). It adopts 16 points on the inner and outer circles for calculation to extract more abundant feature information instead of eight points around the center pixel used by LBP.

Optical flow method aims to quantify facial muscle actions by calculating the motion speed of each pixel in the video. On this basis, the optical strain that reflects the distortion caused by small area motion can be further calculated. If the speed of a pixel in the image is higher than that of the surrounding pixels, its optical strain value will be higher, which can be used to detect the fast and micromovement of muscles in microexpression recognition. Liong et al. (2014) used the optical strain feature weights to highlight the features of the moving area. Liu et al. (2016) proposed the main directional mean optical flow (MDMO), which takes into account the local facial spatial position, statistical motion features, and has lower feature dimension to improve the calculation efficiency. Liu et al. (2018) also proposed the sparse MDMO to solve the problems that average operation in MDMO may lose manifold structures in feature space. Moreover, Wang et al. (2014a) took microexpression video as a fourth-order tensor, and proposed the tensor-independent color space (TICS) algorithm. They also extracted microexpression, and exclude irrelevant images to conduct the recognition through low-rank decomposition of samples (Wang et al., 2014b).

To determine the facial range of feature extraction, Liong et al. (2018b) counted the action units corresponding to all kinds of microexpressions. They found that the actions are only concentrated in a few facial areas, especially in eyes and mouth. If only the features of these three regions are extracted, irrelevant facial information can be filtered out and detection accuracy can be improved effectively. Therefore, this paper determines the region of interests (RoIs) of feature extraction through the distance between inner eyes and mouth corners. As the objects of feature extraction, most of the work directly calculate features of the whole video segment (Chen et al., 2019; Cao et al., 2021), while the apex frame contains the main information of microexpression (Li et al., 2018; Liong et al., 2018a). The apex frame refers to the moment when the movement amplitudes of the facial action units reach peak value in the duration of microexpression. Obviously, only extracting the features of apex frame can dramatically decrease calculating and eliminate the

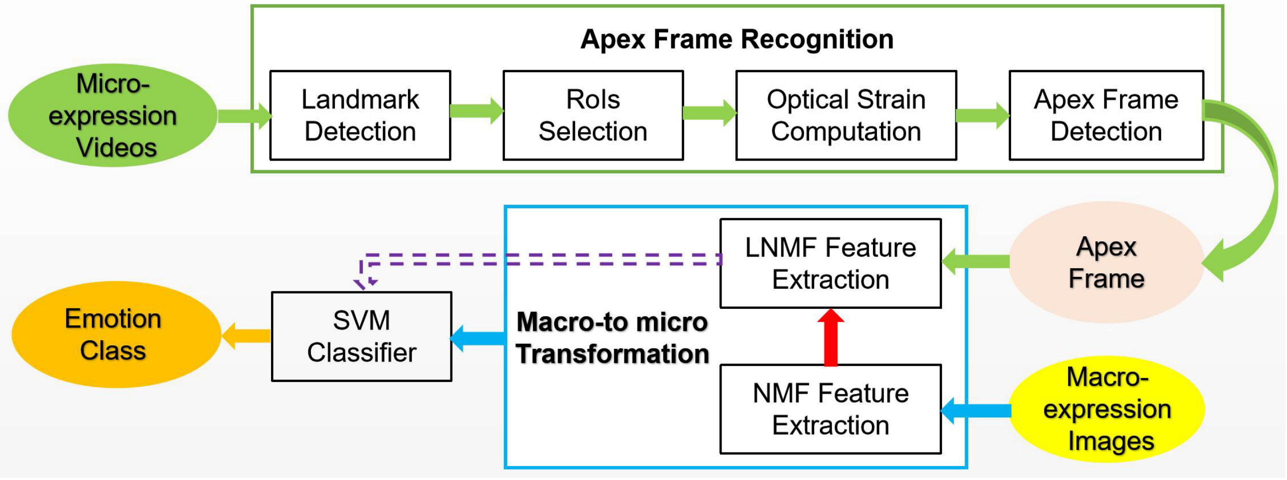


FIGURE 1 | The overall block diagram of the proposed scheme.

interference caused by irrelevant information in the video, which is also the basis of this paper.

Matrix factorization is popular in dimension-reduction fields, which has good physical significance. The original data are expressed as the weighted sum of several bases, which is transformed into a feature vector including weight coefficients to realize perception of the whole from local parts. Principal component analysis (PCA) and singular value decomposition (SVD) are the classic matrix factorization methods. However, the bases and coefficients calculated by these algorithms contain negative elements, which make the decomposition results not well-interpreted. For example, it is not practical to decompose face images into basic sub-images with negative components. To solve this problem, Lee and Seung (2000) proposed the non-negative matrix factorization (NMF) based on non-negative constraints of matrix elements. Li et al. (2001) pointed out that the bases calculated by NMF are redundant and not independent. Hence, the local constraints were added during calculation, that is, LNMF was proposed. The local action of microexpression can be reflected by the local features of LNMF, which is also the reason why we adopt LNMF to extract features of microexpression.

Nowadays, CASME (Yan et al., 2013), CASME2 (Yan et al., 2014), CAS(ME)2 (Qu et al., 2016), and SAMM (Davison et al., 2018a,b) are the widely used datasets for microexpression recognition and classification. However, each dataset has only hundreds of samples, and the number of different microexpressions is seriously unbalanced, which is not sufficient to train a classifier with better generalization ability, especially for deep neural network (DNN). Naturally, researchers hope to train microexpression classifiers by means of numerous macroexpression datasets (Wang et al., 2018; Peng et al., 2019; Zhi et al., 2019). Jia et al. (2018) proposed an MtM algorithm, which uses macroexpression data to generate microexpression samples by constructing corresponding relationship between them. The samples generated by this algorithm are closer to the truth, so it can yield better generalization. However, this algorithm is not suitable for the non-negative features. In this paper, we propose an improved MtM transformation algorithm, which can meet non-negative properties.

3. LNMF AND MTM TRANSFORMATION

The overall scheme is shown in **Figure 1**. The first row (from left-to-right): the key points of human face are located to cropped eyes and mouth as RoIs, then the optical flow features of RoIs are calculated to detect the apex frame. The second row (from right-to-left): the features of apex frame is extracted from the microexpression videos using LNMF, whereas the NMF is used for extracting the features of macroexpression images. Combined with these two, the proposed MtM transformation is used to increase the samples of microexpression considering the corresponding relationship between macro and micro features. Finally, the classifier based on SVM is trained with all the augmented microexpression samples. In the following, we will discuss the key problems of the proposed scheme about RoIs selection, apex frame detection, LNMF principle, and MtM transformation.

3.1. RoIs Selection

To determine the RoIs of eyes and mouth regions, we use open source machine learning toolkit DLIB (King, 2009; Ren et al., 2014) to locate the key points of the first frame in the video including microexpression. It should be noted that the position of key points can be assumed unchanged because the face displacement is very small in video. Therefore, we only detect the key points on the first frame. As shown in **Figure 2**, we use the distance of inner eyes/mouse corners D_{eye}/D_{mouth}, respectively, to determine the RoIs. The distance between the left and right of bounding box of the eyes is $D_{eye}/4$, the downside is $D_{eye}/5$ away from the lowest point of the eye, and the topside is located on the highest point of the eyebrow. The left and right of the bounding box of the mouth are $D_{mouth}/5$ away from the mouth corners, the

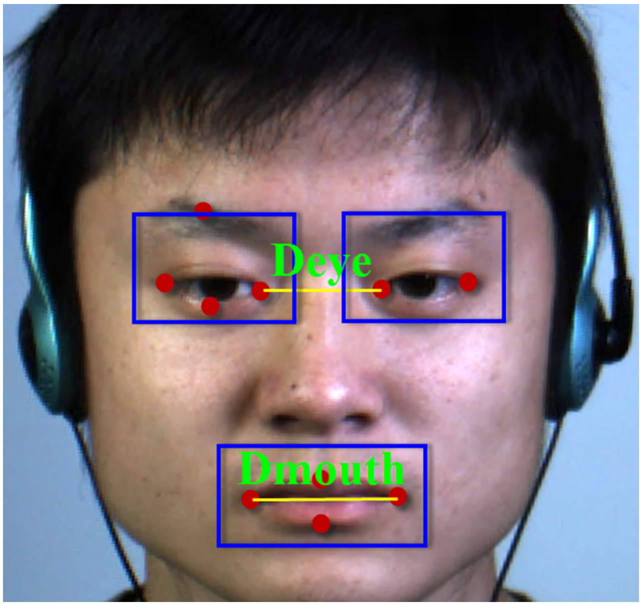


FIGURE 2 | The region of interests (RoIs) of facial expression (Yan et al., 2014).

top is $D_{mouth}/4$, and the bottom is $D_{mouth}/7$ from the highest and lowest points of the mouth, respectively.

3.2. Apex Frame Detection

The optical flow of a pixel refers to its displacement between two frames, which includes both the horizontal and vertical displacement. The optical strain is calculated as the difference of optical-flow values between pixels, which reflects the deformation degree of a non-rigid body during the motion. The microexpression is the micro movement of facial muscles, and the distortion caused by the movement is reflected by the higher optical strain value of this region.

Let v_x and v_y be the optical flow in horizontal and vertical directions, and the definition of optical strain is expressed as follows:

$$\varepsilon_m = \left[\begin{matrix} \varepsilon_{xx} = \frac{\partial v_x}{\partial x} & \varepsilon_{yx} = \frac{1}{2}(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}) \\ \varepsilon_{xy} = \frac{1}{2}(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}) & \varepsilon_{yy} = \frac{\partial v_y}{\partial y}) \end{matrix}\right] \tag{1}$$

$$\varepsilon = \sqrt{\varepsilon_{xx}^2 + \varepsilon_{yx}^2 + \varepsilon_{xy}^2 + \varepsilon_{yy}^2} \tag{2}$$

where ε_m contains the normal and tangential strain of the pixel, and ε is the optical strain value of the pixel.

The pseudo codes for the binary search algorithm to detect the apex frame (Liong et al., 2015) are shown in Algorithm 1. First, we calculate the sum of the optical strains of the pixels in each RoI of all frames and take them as the apex frame range in *Iteration*1. Then, we separate the candidate frames into two average parts, and compare the sum of the optical strains values. The larger one will be the candidate range for the next iteration. Afterward, we repeat the calculation until one frame is converged, that is, the detected apex frame.

Algorithm 1: Binary Apex Frame Detection.

```
input: {$\varepsilon_f$ | f = 1,...N} : Optical strain of every frame
output: n : Apex frame No.
function
    $result \leftarrow 0$
    $lo \leftarrow 0$
    $hi \leftarrow N$
    while $lo < hi$ do
        $mid \leftarrow (lo + hi)/2$
        $sum1 \leftarrow sum(\varepsilon_{lo}, \dots, \varepsilon_{mid})$
        $sum2 \leftarrow sum(\varepsilon_{mid}, \dots, \varepsilon_{hi})$
        if $sum1 \leq sum2$ then
            $lo \leftarrow mid$
        else
            $hi \leftarrow mid$
        end if
    end while
    return $lo$
end function
```

3.3. Local Non-negative Matrix Factorization

The definition of NMF is expressed as Equation (3):

$$D = WH \tag{3}$$

where $D \in R^{m \times n}$ is data matrix; $H \in R^{r \times n}$ is coefficients matrix, in which each column is one sample; and $W \in R^{m \times r}$ is base matrix, in which each column is a base. Define $Y = WH$. NMF takes KL divergence as loss function to measure the effect of factorization as follows:

$$D(X||Y) = \sum_{i,j}(x_{ij} \log \frac{x_{ij}}{y_{ij}} - x_{ij} + y_{ij}) \tag{4}$$

Here, NMF aims to solve the following optimization problem:

$$\begin{aligned} &\min_{W,H} D(X||WH) \\ &s.t. W, H > 0, \sum_i w_{ij} = 1 \forall j \end{aligned} \tag{5}$$

In the optimization process, only non-negative constraints are imposed without local constraint to W. The learned bases are redundant, and the samples cannot be decomposed into individual components. Nevertheless, the LNMF can solve this problem, where three constraints are added: (i) The base should be indivisible, so the sum of squares of elements in the base should be as small as possible. (ii) The bases should be as orthogonal as possible to reduce redundant information. (iii) It is hoped that the most important information in the original data will be retained in W, and the sum of squares of H column elements will be as large as possible. Define $U = W^T W$ and $V = HH^T$, then the optimization function of LNMF is expressed as follows:

$$\min_{W,H} D((X||WH) + \alpha \sum_{i,j} u_{ij} - \beta \sum_i v_{ii}) \tag{6}$$

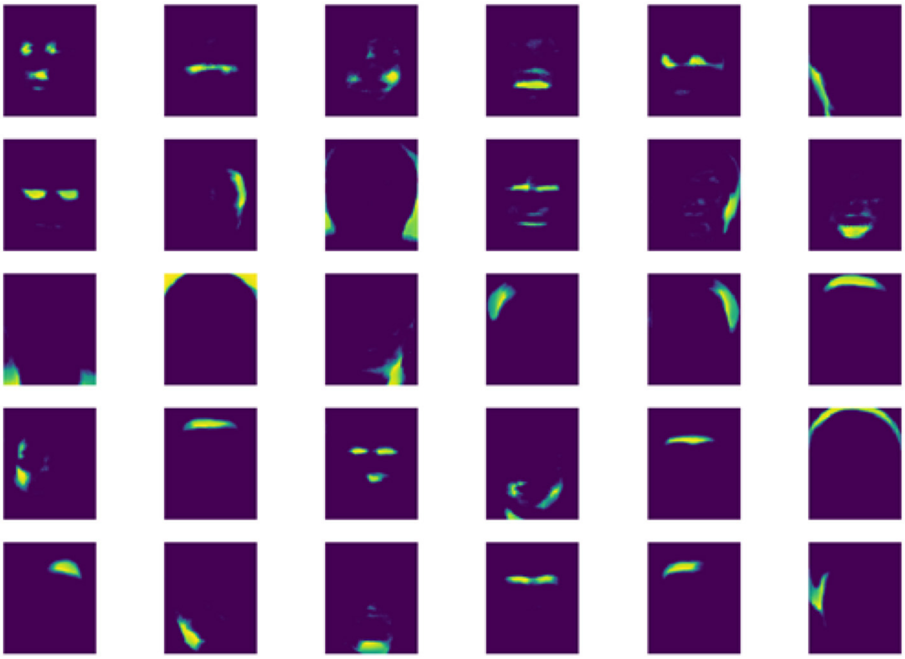


FIGURE 3 | Base matrix of local non-negative matrix factorization (LNMF) for apex frame.

where α and β are constants >0. LNMF iteratively solves Equations (7–9).

$$H = \sqrt{H * (W^T \frac{D}{WH})} \tag{7}$$

$$w_{kl} = \frac{(W * (\frac{D}{WH} H^T))_{kl}}{\sum_j h_{lj}} \tag{8}$$

$$w_{kl} = \frac{w_{kl}}{\sum_k w_{kl}} \tag{9}$$

where "product" means Hadamad product and "division" means matrix division calculation element by element.

The base matrix W obtained from LNMF is shown in **Figure 3**. It can be seen that each base is sparse, only representing the small area of the face. The microexpression is composed of these areas, which further verifies that LNMF is suitable for microexpression feature extraction. The columns of H are the extracted features. We use LNMF for the three RoIs of human face, that is, left/right eyes and mouth, respectively. Finally, we concatenate the three H as the features of samples. The advantage is that we can choose different feature dimensions for eyes and mouth to extract more sufficient features for face action. Obviously, more complex movement pattern of eyes corresponds to a higher dimension.

3.4. Macro-to-Micro Transformation

The fewer samples in the existing microexpression datasets are usually insufficient to train a classifier with good generalization. Jia et al. (2018) proposed an MtM transformation algorithm, which uses macroexpression data to generate new microexpression samples. The basic principle is described as follows:

$$M = [\begin{matrix} X_{ref}^1 & \cdots & X_{ref}^m \\ Y^1 & \cdots & Y^m \end{matrix}]$$

where X/Y represents macro-/microexpression feature sets, respectively; X can be decomposed into X_{ref} and X_{probe}; and X_{ref}^i and Y^i represent the same type of microexpression emotions. The SVD of M is given as:

$$M = USV^T \tag{10}$$

where U can be expressed as:

$$U = [\begin{matrix} R_x \\ R_y \end{matrix}] \tag{11}$$

where R_x/R_y corresponds to macro-/microexpressions in M, respectively. It is used to calculate the weighted sum of column vectors of R_x to get X_{ref}^i (the ith sample in X_{ref}), and the sum of column vectors of R_y with same weights to get Y_i (the ith sample in Y_i), respectively. That is, if we use R_x to get a macroexpression feature, we can also use R_y to get a microexpression feature. So we have:

$$X_{probe} = R_x H \tag{12}$$

$$Y_{new} = R_y H \tag{13}$$

where H is weight matrix, Y_{new} is new microexpression feature samples, and the microexpression emotion is same as each column of X_{probe}.

Because the new feature samples generated by this algorithm do not have non-negative properties, they cannot be used for feature extraction based on LNMF. The reasons include that U is an orthogonal matrix, and in order to meet the requirements of orthogonality, it is impossible that every element is a non-negative number, namely H must have negative elements. In addition, the method deriving H involves the calculation of inverse matrix. When the determinant of matrix is close to 0, the result is not accurate. To acquire the non-negative features,R_x, R_y, and H must be non-negative. Let R_x be the NMF features of macroexpression, and R_y be the LNMF features of microexpression. H is derived by NMF method, so we can get an improved MtM algorithm. The pseudo codes of the proposed algorithm are shown in Algorithm 2.

Algorithm 2: Macro-to-Micro Transformation.

```
input: X: features of macroexpression; Y: features of microexpression
output: RES: new features of microexpression
function
    $RES \leftarrow \{\}$
    for all emo do
        $\mathbf{X}_{emo} \leftarrow$ extract features of emo from $\mathbf{X}$;
        $\mathbf{Y}_{emo} \leftarrow$ extract features of emo from $\mathbf{Y}$;
        split $\mathbf{X}_{emo}$ into $\mathbf{X}_{emo,ref}$ and $\mathbf{X}_{emo,probe}$
        calculate $\mathbf{H}_{emo}$ using Equation (16), iteratively
        $\mathbf{Y}_{new} \leftarrow \mathbf{Y}_{emo}\mathbf{H}_{emo}$
        $RES \leftarrow RES \bigcup \mathbf{Y}_{new}$
    end for;
    return RES
end function
```

Let X_{emo} represents the macroexpression NMF feature set of *emo* emotion, which is deposed into $X_{emo,ref}$ and $X_{emo,probe}$. Let Y_{emo} represents the LNMF feature sample set of microexpression, and the columns number is same as $X_{emo,ref}$. Then we use $X_{emo,ref}$ to derive the linear representation of $X_{emo,probe}$:

$$X_{emo,probe} = X_{emo,ref} H_{emo} \tag{14}$$

$$Y_{emo,new} = Y_{emo} H_{emo} \tag{15}$$

Equation (16) solves H_{emo} from Equation (14) with NMF formula of fixed W:

$$H_{emo} = H_{emo} * \frac{[X^{macro}_{emo,ref}]^T X^{macro}_{emo,probe}}{X^{macro}_{emo,ref} X^{macro}_{emo,probe} H_{emo}} \tag{16}$$

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we will evaluate the proposed scheme, including experiment overview, SVM classifier selection, dimension optimization on LNMF, experiments on CK+/CASME2/SAMM datasets, and result analysis.

4.1. Experiment Overview

In general, researchers often take the predicted emotion classes of microexpressions as recognition objects (Jia et al., 2018), such as disgust, happy, sadness, surprise, and so on. We also adopt this approach. However, it is worth noting that Davison et al. (2018a) and Guo et al. (2019) classify microexpressions using facial action units, instead of predicted emotions to remove the potential bias of human reporting.

Next, we will validate the proposed scheme based on CK+ macroexpression dataset (Kanade et al., 2000; Lucey et al., 2010), CASME2 (Yan et al., 2014), and SAMM (Davison et al., 2018a,b) microexpression datasets. In the experiments, the SVM classifier is used for classifying microexpressions. The optimized dimension on LNMF can contribute to the recognition accuracy of microexpression. The pretest on CK+ is to verify that macroexpression features extracted by NMF are suitable for MtM transformation. The tests on CK+/CASME2 and CK+/SAMM validate that the proposed MtM transformation can improve the recognition accuracy and generalization. The algorithm evaluation compares the performance of the proposed MtM algorithm with original MtM/MDMO/TICS on CASME2, and SA-AT/ATNet/OFF-ApexNet on SAMM.

4.2. SVM Classifier

We adopt the SVM classifier from the Sklearn toolbox based on LIBSVM (Chang and Lin, 2011) to test the macro- and microexpression recognition accuracy. In training phase, the leave-one-sample-out (LOSO) cross-validation is adopted. For each fold, all samples from one subject are used as a testing set and the rest is for training. The final recognition accuracy is the average of five test runs. Microexpression recognition is a multiclassification problem. We use one-vs-one trajectory based on SVM binary classifier to train one SVM between any two classes. If the sample includes n classes, then we have $n * (n - 1)/2$ SVM. The classification result is determined by all the SVM voting together. To classify the linearly inseparable microexpression, we adopt poly after evaluating sigmoid, radial basis function (RBF), and poly kernel functions. Its definition is expressed as Equation (17).

$$K(x_i, x_j) = (\gamma x_i^T x_j + \alpha)^d \tag{17}$$

where x_i, x_j are the feature vectors, and γ, α, d are preset hyperparameters.

Remark 1. We acquire the optimized parameters about SVM empirically, which are $\gamma = 4, \alpha = 0, d = 4$ for CASME2, and $\gamma = 4, \alpha = 0, d = 1$ for SAMM. We just compare the final recognition accuracy with the related references, where the detailed parameter values about SVM cannot be found.

4.3. Dimension on LNMF

If the dimension is too small, microexpression features cannot be decomposed into various detailed components based on LNMF. While the dimension is too large, the features will be too scattered. We determine the optimized value through prior

TABLE 1 | Dimensions and recognition accuracy on local non-negative matrix factorization (LNMF).

		Dimension on eyes			
		40	50	60	70
Dimension on mouth	60	0.694	0.701	0.692	0.720
	70	0.708	0.698	0.702	0.702
	80	**0.726**	0.706	0.707	0.697
	90	0.700	0.690	0.700	0.680
	100	0.719	0.696	0.707	0.706

Optimal value is bold.

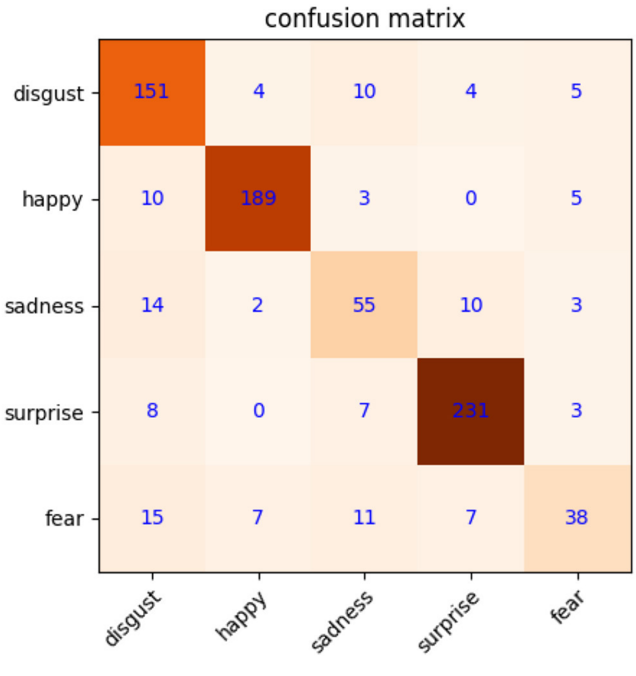


FIGURE 4 | Confusion matrix of macroexpression recognition on CK+.

testing and comparing with different dimension setting of eyes and mouth. As shown in **Table 1**, the optimized dimension is 40/40/80 (left/right eyes and mouth) based on CASME2 with a recognition accuracy of 72.6%. Adopt this same approach, we get the optimized dimension 120/120/110 based on SAMM with a recognition accuracy of 74.68%.

4.4. CK+-Based Pretest

The precondition for MtM transformation is that macroexpression features have better distinguishability. To validate this, we first calculate the weight coefficients of macroexpression, and then use them to extract the features of macroexpressions based on NMF. The image resolution of CK+ is 48 × 48. We use NMF of 200 dimensions to acquire the features directly. The confusion matrix about macroexpression recognition is shown in **Figure 4**, with a high accuracy of 83.8%. It shows that macroexpression features extracted by NMF are suitable for MtM transformation to augment microexpression samples.

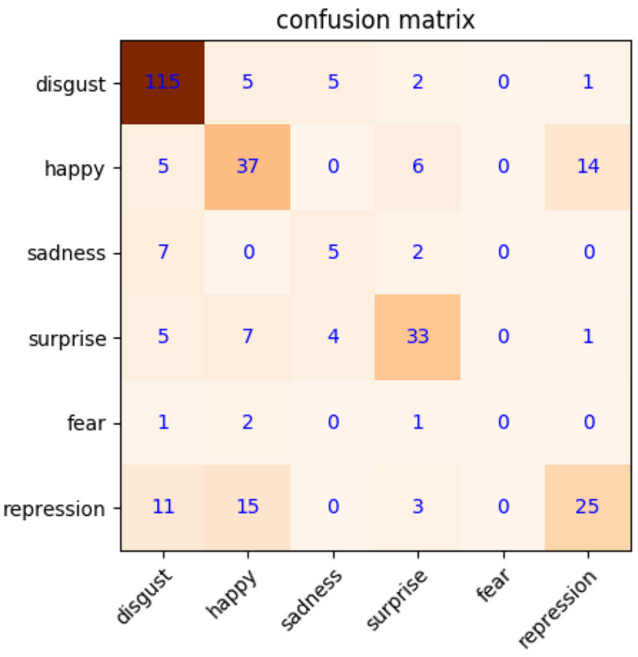


FIGURE 5 | Confusion matrix of microexpression recognition on CASME2 (without new samples).

4.5. CK+/CASME2-Based Test

First, the basic test only focuses on apex frame recognition, LNMF feature extraction, and SVM classifier on CASME2. The RoIs of microexpression are determined according to the distance between inner eyes and mouth corners. It is necessary to normalize the size of eyes to 80 × 90 and mouth to 70 × 150. The 40/40/80 dimension on LNMF is applied to two eyes and mouth regions of samples in CASME2. Three types of features are concatenated in series as the features of CASME2 samples, so the final dimension is 160. Then, the classifier based on SVM is used to test the recognition accuracy by LOSO cross-validation. As shown in **Figure 5**, the confusion matrix of microexpression with a recognition accuracy of 68.9% (without new samples from our MtM transformation). It will be our baseline compared with the next optimized test.

Second, the optimized test is carried out with the proposed MtM transformation base on the aforementioned basic test. Considering that CK+ contains anger, contempt, disgust, fear, happy, sadness, and surprise expressions, CASME2 includes disgust, happy, sadness, surprise, fear, repression, and so on. To compare with (Jia et al., 2018) under equivalent conditions, we adopt the same emotions, that is, disgust, happy, sadness, surprise, fear, and repression. In CK+, there are 792 samples labeled as disgust, happy, sadness, surprise, and fear expressions. Moreover, half of the expression samples are separated as $X_{emo,ref}$ and $X_{emo,probe}$ for subsequent MtM transformation. There are only 156 samples in CASME2, so we double them to 312 through mirroring. For one-to-one correspondence between microexpression in CASME2 and macroexpression in CK+, we use the samples in CASME2 repeatedly to match macroexpression samples in CK+. By this way, we can acquire 312 original samples and 396 new samples (total of 708) for microexpression recognition. After MtM transformation, we get more microexpress samples, including original, mirrored, and new from MtM transformation. It can contribute to train a better SVM classifier. As shown in **Figure 6**, the recognition accuracy improves about 3.8–72.7%, compared with **Figure 5**. It can be seen that the recognition accuracy of happy, sadness, and fear has increased significantly, while surprise and repression unimproved evidently. The reasons actually lie in some similarities, such as eyebrow raising movements (Guo et al., 2020) for surprise, disgust, happy, and sadness expression, and similar cheek movements for repression, disgust, and happy. These similarities cannot be distinguished only through the LNMF features of apex frame, which is a limitation to our algorithm at present, but it can be one of our future working.

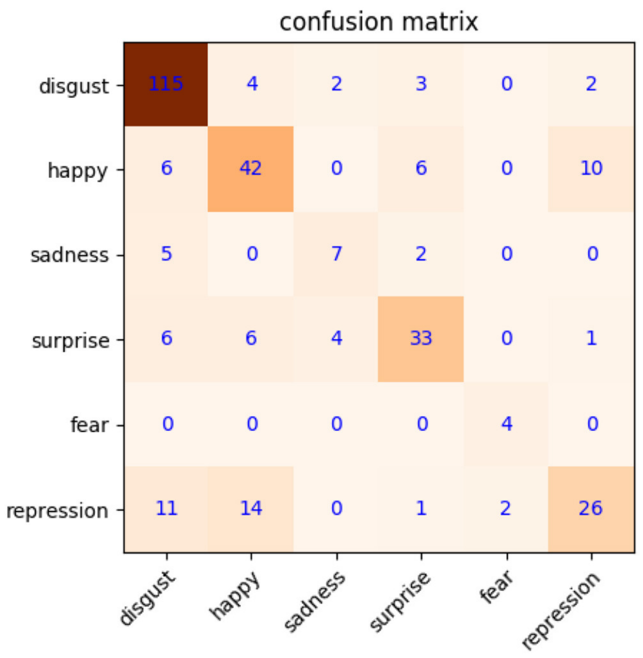


FIGURE 6 | Confusion matrix of microexpression recognition on CK+/CASME2 (with new samples).

However, we only select the original 312 samples in CASME2 for testing, instead of the newly augmented samples (only for training), to avoid distorting the recognition accuracy. Although larger number of new samples can increase the final recognition accuracy, it is not consistent with the fact. We double the samples through mirroring only in training set. When using LOSO cross-validation, it is necessary to exclude the mirrored samples for testing to prevent the false high accuracy caused by two similar samples.

4.6. CK+/SAMM-Based Test

There are totally 159 samples in SAMM dataset (Davison et al., 2018b), which includes seven types of emotions, such

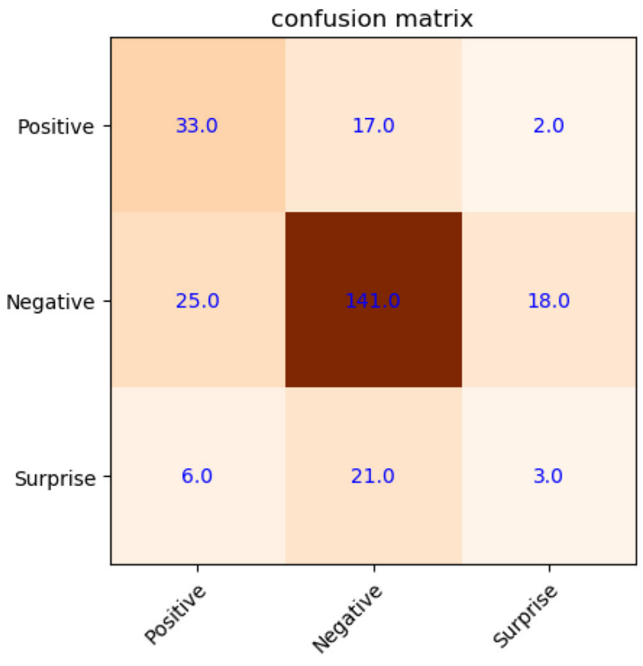


FIGURE 7 | Confusion matrix of microexpression recognition on SAMM (without new samples).

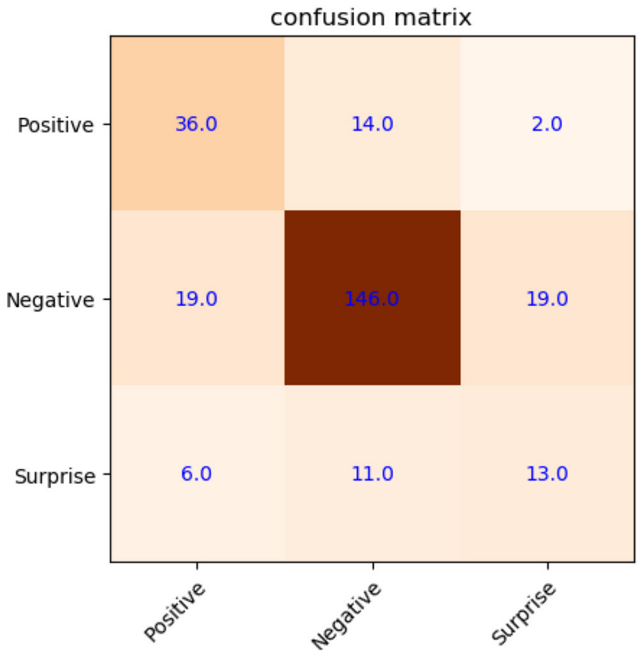


FIGURE 8 | Confusion matrix of micro-expression recognition on CK+/SAMM (with new samples).

TABLE 2 | Recognition accuracy of different algorithms on CK+/CASME2.

Our MtM	Original MtM	MDMO	TICS
0.726	0.655	0.572	0.618

Optimal value is bold.

TABLE 3 | Recognition accuracy of different algorithms on CK+/SAMM.

Our MtM	SA-AT	ATNet	OFF-ApexNet
0.733	0.549	0.701	0.682

Optimal value is bold.

as anger (57), sadness (6), fear (8), others (26), surprise (15), disgust (9), contempt (12), and happiness (26). To compare performance under equal conditions, we divide the emotions into positive (happiness), negative (anger, sadness, fear, disgust, and contempt) and surprise as the same as Liong et al. (2019), Peng et al. (2019), and Zhou et al. (2019).

Figure 7 shows the confusion matrix of microexpression with a recognition accuracy of 66.54% (without new samples). Here, only the features on the original and mirrored samples are extracted from SAMM based on LNMF directly, then SVM classifier with LOSO cross-validation is used to classify the microexpressions.

Compared with **Figure 7**, it shows a recognition accuracy of 73.3%, and increased by 6.76%, as shown in **Figure 8**. Especially in the surprise emotion, its recognition accuracy improved quite a lot with dramatically augmented samples. Here, similar with the experiment on CASME2, three types of samples are used to train the SVM with better generalization in this experiment, including original, mirrored, and new from MtM transformation.

4.7. Algorithm Evaluation

We evaluate the proposed MtM algorithm by comparing it with the original MtM (Jia et al., 2018), MDMO (Liu et al., 2016), and TICS (Wang et al., 2014a) based on CASME2, respectively. As shown in **Table 2**, the proposed MtM algorithm has better performance with a recognition accuracy of 72.6%. Therefore, LNMF can extract more accurate features, and MtM transformation can expand the training samples significantly to prompt the SVM classifier to have better generalization.

As for SAMM, we evaluate the proposed MtM algorithm by comparing it with SA-AT (Zhou et al., 2019), ATNet (Peng et al., 2019), and OFF-ApexNet (Liong et al., 2019), respectively. As shown in **Table 3**, the proposed MtM algorithm also has better performance with a recognition accuracy of 73.3%.

5. CONCLUSION

A new microexpression recognition scheme is proposed, which includes feature extracting and sample expanding. We first determine RoIs with optimized dimensions by facial feature points, then the apex frame is obtained from microexpression video by optical flow method. Afterward, LNMF is developed for each RoI, the results of which are concatenated in series as features of microexpression. Furthermore, the MtM transformation based on LNMF is used, which can increase microexpression samples significantly. A classifier based on SVM is trained with microexpression features and yields better generalization. Finally, the proposed MtM algorithm shows better performance in comparison with other algorithms.

However, the proposed algorithm cannot distinguish some expressions with similar motion features at present. There are obvious recognition confusion on similar eyebrow rising motion, such as surprise, disgust, and happy expression. Therefore, our future work will focus on better feature extraction algorithm to address this issue. Moreover, we will also consider deep forest (Ma et al., 2020; Zhang et al., 2020) as classifier and deep neural networks (Wang et al., 2018) for microexpression recognition in the future.

DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found at CK+: http://www.jeffcohn.net/Resources/, CASME2: http://fu.psych.ac.cn/CASME/casme2-en.php, and SAMM: http://www2.docm.mmu.ac.uk/STAFF/m.yap/dataset.php.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

JGa, HC, and XZ: conceptualization, methodology, and validation. JGa and HC: software. JGa, HC, XZ, and WL: writing and original draft preparation. JGa, HC, XZ, JGu, and WL: writing–review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was supported by National Natural Science Foundation (NNSF) of China under Grant 61803103 and China Scholarship Council (CSC) under Grant 201908440537.

REFERENCES

Ben, X., Jia, X., Yan, R., Zhang, X., and Meng, W. (2018). Learning effective binary descriptors for micro-expression recognition transferred by macro-information. *Pattern Recognit. Lett.* 107, 50–58. doi: 10.1016/j.patrec.2017.07.010

Cao, Z., Liao, T., Song, W., Chen, Z., and Li, C. (2021). Detecting the shuttlecock for a badminton robot: a YOLO based approach. *Expert Syst. Appl.* 164:113833. doi: 10.1016/j.eswa.2020.113833

Chang, C.-C., and Lin, C. (2011). Libsvm: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 27:1–27:27. doi: 10.1145/1961189.1961199

Chen, W., Liao, T., Li, Z., Lin, H., Xue, H., Zhang, L., et al. (2019). Using ftoc to track shuttlecock for the badminton robot. *Neurocomputing* 334, 182–196. doi: 10.1016/j.neucom.2019.01.023

Davison, A. K., Lansley, C., Costen, N., Tan, K., and Yap, M. H. (2018b). SAMM: a spontaneous micro-facial movement dataset. *IEEE Trans. Affect. Comput.* 9, 116–129. doi: 10.1109/TAFFC.2016.2573832

Davison, A. K., Merghani, W., and Yap, M. H. (2018a). Objective classes for micro-facial expression recognition. *J. Imaging* 4:119. doi: 10.3390/jimaging4100119

Ekman, P. (2006). *METT: Micro Expression Training Tool; SETT: Subtle Expression Training Tool*. San Francisco, CA: Paul Ekman.

Ekman, P., and Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry* 32, 88–106. doi: 10.1080/00332747.1969.11023575

Guo, C., Liang, J., Zhan, G., Liu, Z., Pietikäinen, M., and Liu, L. (2019). Extended local binary patterns for efficient and robust spontaneous facial micro-expression recognition. *IEEE Access* 7, 174517–174530. doi: 10.1109/ACCESS.2019.2942358

Guo, J., Liu, Y., Qiu, Q., Huang, J., Liu, C., Cao, Z., et al. (2020). A novel robotic guidance system with eye gaze tracking control for needle based interventions. *IEEE Trans. Cogn. Dev. Syst*. doi: 10.1109/TCDS.2019.2959071

Haggard, E. A., and Isaacs K. S. (1966). "Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy," in *Methods of Research in Psychotherapy*. The Century Psychology Series (Boston, MA: Springer). doi: 10.1007/978-1-4684-6045-2_14

Jia, X., Ben, X., Yuan, H., Kpalma, K., and Meng, W. (2018). Macro-to-micro transformation model for micro-expression recognition. *J. Comput. Sci.* 25, 289–297. doi: 10.1016/j.jocs.2017.03.016

Kanade, T., Cohn, J. F., and Tian, Y. (2000). "Comprehensive database for facial expression analysis," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)* (Grenoble), 46–53. doi: 10.1109/AFGR.2000.840611

King, D. (2009). Dlib-ml: a machine learning toolkit. *J. Mach. Learn. Res*. 10, 1755–1758. Available online at: https://www.semanticscholar.org/paper/Dlib-ml%3A-A-Machine-Learning-Toolkit-King/2ea6a93199c9227fa0c1c7de13725f918c9be3a4

Lee, D., and Seung, H. S. (2000). "Algorithms for non-negative matrix factorization," in *NIPS* (Denver, CO), 535–541.

Li, S. Z., Hou, X. W., Zhang, H. J., and Cheng, Q. S. (2001). "Learning spatially localized, parts-based representation," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Vol. 1 (Kauai, HI). doi: 10.1109/CVPR.2001.990477

Li, Y., Huang, X., and Zhao, G. (2018). "Can micro-expression be recognized based on single apex frame?" in *2018 25th IEEE International Conference on Image Processing (ICIP)* (Athens), 3094–3098. doi: 10.1109/ICIP.2018.8451376

Liong, S., See, J., Wong, K., Le Ngo, A. C., Oh, Y., and Phan, R. (2015). "Automatic apex frame spotting in micro-expression database," in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)* (Kuala Lumpur), 665–669. doi: 10.1109/ACPR.2015.7486586

Liong, S.-T., Gan, Y. S., Yau, W.-C., Huang, Y., and Ken, T. (2019). Off-apexnet on micro-expression recognition system. *arXiv* abs/1805.08699.

Liong, S.-T., See, J., Phan, R. C.-W., Ngo, A. C. L., Oh, Y.-H., and Wong, K. (2014). "Subtle expression recognition using optical strain weighted features," in *ACCV Workshops* (Singapore), 644–657. doi: 10.1007/978-3-319-16631-5_47

Liong, S.-T., See, J., Phan, R. C.-W., and Wong, K. (2018a). Less is more: micro-expression recognition from video using apex frame. *arXiv* abs/1606.01721. doi: 10.1016/j.image.2017.11.006

Liong, S.-T., See, J., Phan, R. C.-W., Wong, K., and Tan, S.-W. (2018b). Hybrid facial regions extraction for micro-expression recognition system. *J. Signal Process. Syst*. 90, 601–617. doi: 10.1007/s11265-017-1276-0

Liu, Y., Zhang, J., Yan, W., Wang, S., Zhao, G., and Fu, X. (2016). A main directional mean optical flow feature for spontaneous micro-expression recognition. *IEEE Trans. Affect. Comput*. 7, 299–310. doi: 10.1109/TAFFC.2015.2485205

Liu, Y.-J., Li, B.-J., and Lai, Y.-K. (2018). Sparse MDMO: learning a discriminative feature for spontaneous micro-expression recognition. *IEEE Trans. Affect. Comput*. doi: 10.1109/TAFFC.2018.2854166

Lucey, P., Cohn, J., Kanade, T., Saragih, J. M., Ambadar, Z., and Matthews, I. (2010). "The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition–Workshops* (San Francisco, CA), 94–101. doi: 10.1109/CVPRW.2010.5543262

Ma, C., Liu, Z., Cao, Z., Song, W., Zhang, J., and Zeng, W. (2020). Cost-sensitive deep forest for price prediction. *Pattern Recognit*. 107:107499. doi: 10.1016/j.patcog.2020.107499

Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell*. 24, 971–987. doi: 10.1109/TPAMI.2002.1017623

Peng, M., Wang, C., Bi, T., Chen, T., Zhou, X., and Shi, Y. (2019). "A novel apex-time network for cross-dataset micro-expression recognition," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (Cambridge, UK), 1–6. doi: 10.1109/ACII.2019.8925525

Qu, F., Wang, S., Yan, W.-J., and Fu, X. (2016). "Cas(me)2: a database of spontaneous macro-expressions and micro-expressions," in *HCI* (Toronto, ON), 48–59. doi: 10.1007/978-3-319-39513-5_5

Ren, S., Cao, X., Wei, Y., and Sun, J. (2014). "Face alignment at 3000 fps via regressing local binary features," in *2014 IEEE Conference on Computer Vision and Pattern Recognition* (Columbus, OH), 1685–1692. doi: 10.1109/CVPR.2014.218

Wang, S., Li, B., Liu, Y., Yan, W., Ou, X., Huang, X., et al. (2018). Micro-expression recognition with small sample size by transferring long-term convolutional neural network. *Neurocomputing* 312, 251–262. doi: 10.1016/j.neucom.2018.05.107

Wang, S., Yan, W., Li, X., Zhao, G., and Fu, X. (2014a). "Micro-expression recognition using dynamic textures on tensor independent color space," in *2014 22nd International Conference on Pattern Recognition* (Stockholm), 4678–4683. doi: 10.1109/ICPR.2014.800

Wang, S., Yan, W.-J., Zhao, G., Fu, X., and Zhou, C. (2014b). "Micro-expression recognition using robust principal component analysis and local spatiotemporal directional features," in *ECCV Workshops* (Zurich), 325–338. doi: 10.1007/978-3-319-16178-5_23

Wang, Y., See, J., Phan, R. C.-W., and Oh, Y.-H. (2014). "LBP with six intersection points: reducing redundant information in LBP-top for micro-expression recognition," in *ACCV* (Singapore), 525–537. doi: 10.1007/978-3-319-16865-4_34

Wu, Q., and Fu, X. (2010). Micro-expression and its applications. *Adv. Psychol. Sci.* 18, 1359–1368. Available online at: https://www.semanticscholar.org/paper/Micro-expression-and-Its-Applications-Qi-Xun-Bing/2397917d52e2c52698843285b9dd9929ea28e2e6

Yan, W., Li, X., Wang, S.-J., Zhao, G., Liu, Y.-J., Chen, Y., et al. (2014). CASME II: an improved spontaneous micro-expression database and the baseline evaluation. *PLoS ONE* 9:e86041. doi: 10.1371/journal.pone.0086041

Yan, W. J., Wu, Q., Liu, Y. J., Wang, S. J., and Fu, X. (2013). "CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces," in *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (Shanghai), 1–7. doi: 10.1109/FG.2013.6553799

Zhang, L., Chen, Z., Cui, W., Li, B., Chen, C.-Y., Cao, Z., et al. (2020). Wifi-based indoor robot positioning using deep fuzzy forests. *IEEE Internet Things J*. doi: 10.1109/JIOT.2020.2986685

Zhao, G., and Pietikainen, M. (2007). Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell*. 29, 915–928. doi: 10.1109/TPAMI.2007.1110

Zhi, R., Xu, H., Wan, M., and Li, T. (2019). Combining 3D convolutional neural networks with transfer learning by supervised pre-training for facial micro-expression recognition. *IEICE Trans. Inf. Syst*. 102-D, 1054–1064. doi: 10.1587/transinf.2018EDP7153

Zhou, L., Mao, Q., and Xue, L. (2019). "Cross-database micro-expression recognition: a style aggregated and attention transfer approach," in *2019 IEEE International Conference on Multimedia Expo Workshops (ICMEW)* (Shanghai), 102–107. doi: 10.1109/ICMEW.2019.00025

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fnbot.2020.565702

Reinforcement Q-Learning Control With Reward Shaping Function for Swing Phase Control in a Semi-active Prosthetic Knee

Yonatan Hutabarat[1]*, Kittipong Ekkachai[2], Mitsuhiro Hayashibe[1,3] and Waree Kongprawechnon[4]

[1] Neuro-Robotics Laboratory, Graduate School of Biomedical Engineering, Tohoku University, Sendai, Japan, [2] Smart Machine and Mixed Reality (SMR) Laboratory, National Electronics and Computer Technology Center (NECTEC), Pathum Thani, Thailand, [3] Department of Robotics, Graduate School of Engineering, Tohoku University, Sendai, Japan, [4] School of Information Computer and Communication Technology (ICT), Sirindhorn International Institute of Technology (SIIT), Thammasat University, Pathum Thani, Thailand

In this study, we investigated a control algorithm for a semi-active prosthetic knee based on reinforcement learning (RL). Model-free reinforcement Q-learning control with a reward shaping function was proposed as the voltage controller of a magnetorheological damper based on the prosthetic knee. The reward function was designed as a function of the performance index that accounts for the trajectory of the subject-specific knee angle. We compared our proposed reward function to a conventional single reward function under the same random initialization of a *Q*-matrix. We trained this control algorithm to adapt to several walking speed datasets under one control policy and subsequently compared its performance with that of other control algorithms. The results showed that our proposed reward function performed better than the conventional single reward function in terms of the normalized root mean squared error and also showed a faster convergence trend. Furthermore, our control strategy converged within our desired performance index and could adapt to several walking speeds. Our proposed control structure has also an overall better performance compared to user-adaptive control, while some of its walking speeds performed better than the neural network predictive control from existing studies.

Keywords: reinforcement learning, reward shaping, Q-learning, semi-active prosthetic knee, magnetorhelogical damper

Edited by:
Yan Wu,
Institute for Infocomm Research
(A*STAR), Singapore

Reviewed by:
Gian Carlo Cardarilli,
University of Rome Tor Vergata, Italy
Noman Naseer,
Air University, Pakistan

***Correspondence:**
Yonatan Hutabarat
y.hutabarat@neuro.mech.tohoku.ac.jp

Received: 26 May 2020
Accepted: 31 August 2020
Published: 26 November 2020

Citation:
Hutabarat Y, Ekkachai K, Hayashibe M and Kongprawechnon W (2020) Reinforcement Q-Learning Control With Reward Shaping Function for Swing Phase Control in a Semi-active Prosthetic Knee.
Front. Neurorobot. 14:565702.
doi: 10.3389/fnbot.2020.565702

1. INTRODUCTION

The knee joint enables one to perform basic movements, such as walking. The loss of this function such as in the case of transfemoral amputation could severely restrict movements. The lower limb prosthetic system, which comprises either the prosthetic knee, leg, or foot, could replace the function of the biological knee. Generally, the prosthetic knee is divided into two categories, that is, a mechanical-based control and microprocessor controlled. Reportedly, using the microprocessor-controlled prosthetic knee can improve the lower extremity joint kinetics symmetry, gait, and balance, as well as reduce the frequency of stumbling and falling, compared to using the mechanical or passive knee (Hafner et al., 2007; Kaufman et al., 2007, 2012; Sawers and Hafner, 2013).

Generally, the actuator in a microprocessor-controlled prosthetic knee can be divided into two categories: semi-active and active mechanisms. An active mechanism can generate a net positive force. Several institutions have been developing the active knee for research and development purposes (Hoover et al., 2012; Lawson et al., 2014; Flynn et al., 2015). However, owing to the high requirements of the actuation unit as well as the control system in terms of design and cost (Windrich et al., 2016), there has been only a few of the commercialized product in this category, such as the Power Knee (Össur, Iceland)[1].

On the contrary, a semi-active mechanism or also called a variable-damping mechanism could only manipulate damping force. Magnetorheological (MR) damper is one of the examples that utilize this function by manipulating the strength of the magnetic field, which is applied to magnetic particles in a carrier fluid. The advantages of using this system are the rapid response and low power consumption, among others (Şahin et al., 2010). Therefore, from the cost-effective and functionality point of view, a semi-active prosthetic knee is still more favorable for the end user compared to the active mechanism. Consequently, in this study we focused on the control of prosthetic knee devices with a semi-active mechanism in a swing phase of the gait cycle.

Many studies on the prosthetic knee control algorithm have been conducted. The user-adaptive control as investigated in Herr and Wilkenfeld (2003) is an example of an adaptive control that applied the MR damper-based prosthetic knee. The underlying principle of this controller is to change the necessary damping required in each state if the biological knee trajectory deviated based on the information of the local sensing device. A finite state machine-based controller is often found in the powered knee (Wen et al., 2017). This controller is programmed to provide a control output of the current state machine obtained from specific rules based on varying sensing information. There has been an attempt to unify the prosthetic controller through discrete Fourier transform virtual constraints (Quintero et al., 2017). Furthermore, EMG-based control has been investigated in several studies, such as in Hoover et al. (2012). While this control has promising results, its application is limited to those who still have intact muscle function on the amputation site.

Several studies have tried to apply machine learning algorithm to control prosthetic (Ekkachai and Nilkhamhang, 2016; Wen et al., 2017, 2019). Neural network predictive control (NNPC) was employed as a control structure for the swing phase in the prosthetic knee (Ekkachai and Nilkhamhang, 2016). The swing phase model was constructed following a feed-forward neural network structure in which the input and the output were the knee angle, control voltage, and prediction of future knee angle. However, it requires an off-line training process to find weight and bias of neural network. Thus, when neural network has been trained, it will not have a mechanism to adapt the model. This raises a need of online learning model that could adapt if users change walking pattern due to weight change or using different costume.

An adaptive dynamic programming was employed in each state of walking for automatic tuning of the knee joint impedance parameter (Wen et al., 2017) and further improved into an online reinforcement learning (RL)-based control to tune a total of 12 impedance parameters of robotic knee prosthesis (Wen et al., 2019). Although it has shown potential outcome for human-prosthesis control tuning in a real time setting, the proposed algorithm is needed to tune a total 12 impedance parameters for 4 phases of walking. This is understandable since it was applied to powered prosthetic knee (Wen et al., 2019).

In this study, we investigated a model-free Q-learning control algorithm with a reward shaping function as the swing phase control in the MR damper-based prosthetic knee. A model-free algorithm could simplify the need for prior information, thus it could be implemented to different subjects effectively. We found that our proposed reward shaping function leads to better performance in terms of normalized root mean squared error and also showed a faster convergence trend compared to a conventional single reward function. Our proposed approach was also compared to user-adaptive control and NNPC from existing studies, which resulted in overall better performance across tested walking speeds.

The rest of this paper is organized as follows. Section 2 describes the specific MR damper system, double pendulum model as the environment, and the dataset that we used, as well as the details on Q-learning control. Section 3 presents the simulation and results. Finally, Section 4 discusses the algorithm comparison, the limitations, and the future works of this study.

2. MATERIALS AND METHODS

In this section, we introduce the system, the environment model, and the RL algorithm we designed in this study. MR damper is defined as the system, that is, the main actuator to be controlled. Meanwhile, the environment is defined as the application where the system was used; in this case, a simple double pendulum model was used as the simulated environment to perform swing phase on a gait cycle. Section 2.1 covers a brief descriptions on the system and environment as well as dataset used in this study. Further, Q-learning algorithm designed for this study is discussed in detail in section 2.2.

2.1. System, Environment Model, Dataset

2.1.1. System Description

In this study, prosthetic knee is actuated by MR damper having non-linear characteristics such as hysteresis and dynamic response that are difficult to control. To capture these behaviors of MR damper, the elementary hysteresis model (EHM) based feed-forward neural network (FNN) model is used in our simulation. It was proposed in Ekkachai et al. (2012) and modified in Ekkachai and Nilkhamhang (2016). The model consists of two FNNs. Here, one FNN coupled with EHM acted as a hysteresis model, and the output of this network was fed to the other FNN that acted as the gain function. Voltage is filtered by the first-order lag filter. Piston velocity and acceleration are used as inputs to estimate MR damper force. The MR damper model is shown in **Figure 1A**. The model was trained by using

---

[1] Available online at: https://www.ossur.com/en-us/prosthetics/knees/power-knee (accessed July 17, 2020).

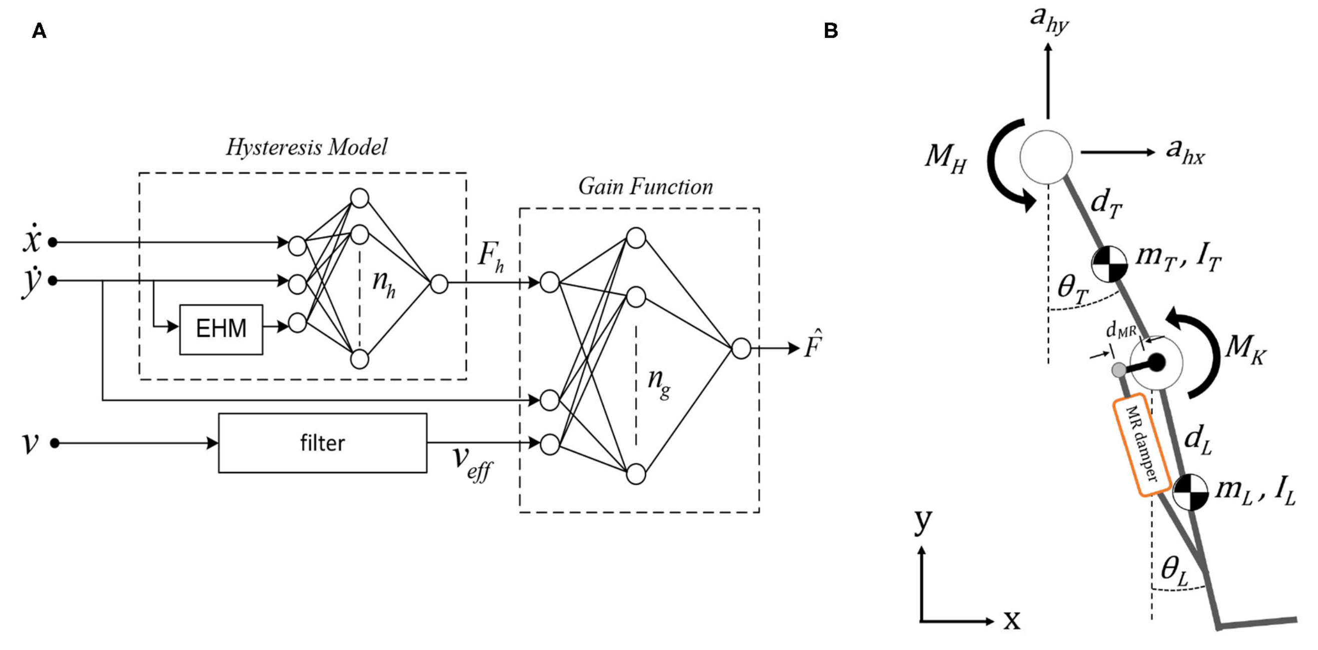


FIGURE 1 | (A) Control structure of magnetorheological (MR) damper (Ekkachai et al., 2013). **(B)** Double pendulum model to simulate swing phase with MR damper attachment with distance d_{MR} from the knee joint.

data from the experimental system of an actual MR damper, Lord RD-8040-1, described in Ekkachai et al. (2013).

$$M_K = d_{MR} \cdot |\hat{F}| \cos \theta_K \tag{1}$$

The MR damper is attached at a distance, d_{MR}, away from the knee joint. Based on this distance, the torque generated at knee joint by the MR damper is calculated by Equation (1), where $\hat{F}$ is the force generated by MR damper (**Figure 1A**) and θ_K is the knee angle. θ_K is calculated by $\theta_K = \theta_T - \theta_L$, where subscripts T and L denote thigh and leg segment, respectively, as shown in **Figure 1B**.

2.1.2. Environment Model

The double pendulum model is proposed as the environment model for the swing phase (Putnam, 1991). The model consists of two links, that is, thigh and a lumped shank, as well as a foot segment, as depicted in **Figure 1B**. There are two actuated joints with a total of four degrees of freedom, where the hip joint has one rotational degree of freedom on the z-axis and two translation degrees of freedom on the x and y-axes; meanwhile, the knee joint has one rotational degree of freedom on the z-axis.

$$\begin{aligned} M_K &= I_L\alpha_L + m_L d_L(L_T\alpha_T \cos(\theta_L - \theta_T) + \omega_T L_T \sin(\theta_L - \theta_T) \\ &+ a_{hx} \cos\theta_L + (a_{hy} + g) \sin\theta_L) \end{aligned} \tag{2}$$

$$\begin{aligned} M_H &= M_K + (m_L L^2{}_T + I_T)\alpha_T + m_L d_L L_T(\alpha_L \cos(\theta_L - \theta_T) \\ &- \omega^2{}_L \sin(\theta_L - \theta_T)) + (m_L L_T + m_T d_T)(a_{hx} \cos\theta_T \\ &+ (a_{hy} + g) \sin\theta_T) \end{aligned} \tag{3}$$

This model was simulated in MATLAB (Mathworks Inc., Natick, MA, USA) SimMechanics environment. The torque generated by each joint, derived from Lagrange equation, are governed by Equations (2) and (3), where M_K and M_H are the torques at knee and hip, respectively. $m, I, d, and\ L$ are segment mass, moment of inertia at segment's center of mass, length measured from the proximal end of the segment to the center of mass, and segment length, respectively. The subscripts L and T denote the leg segment and thigh segment, respectively, while a_{hx} and a_{hy} are the linear acceleration at hip joint along the x and y axes. Further, $\theta, \omega, \alpha, and\ g$ are the angle, angular velocity, angular acceleration, and gravitational constant at 9.8 m/s^2, respectively.

2.1.3. Dataset

The gait data used in this study are also normal gait data collected from Ekkachai and Nilkhamhang (2016) for convenience in comparison study of the controller. In this manner, the proposed controller performance can be compared to the previous method with same dataset. It allows us to analyze the difference from the previous work result keeping the same experimental condition. A male subject with 83 kg of weight and 1.75 m height at the time of the experiment were asked to walk on a treadmill at various speed, where in this study walking speed was set at 2.4, 3.6, and 5.4 km/h (Ekkachai and Nilkhamhang, 2016). A high-speed camera was used to capture joints coordinate and later converted to relative joint angles. To capture the respective joints coordinate, reflective markers were placed at hip, knee, and ankle joints. In this study, as only the control in the swing phase is discussed, the gait data used will be constrained into the swing phase only. Since we proposed a RL-based algorithm, all the recorded knee angle data with a total of 200 sets per walking speed will be used. The average knee angle data at the swing phase used in this study are depicted in **Figure 2A**.

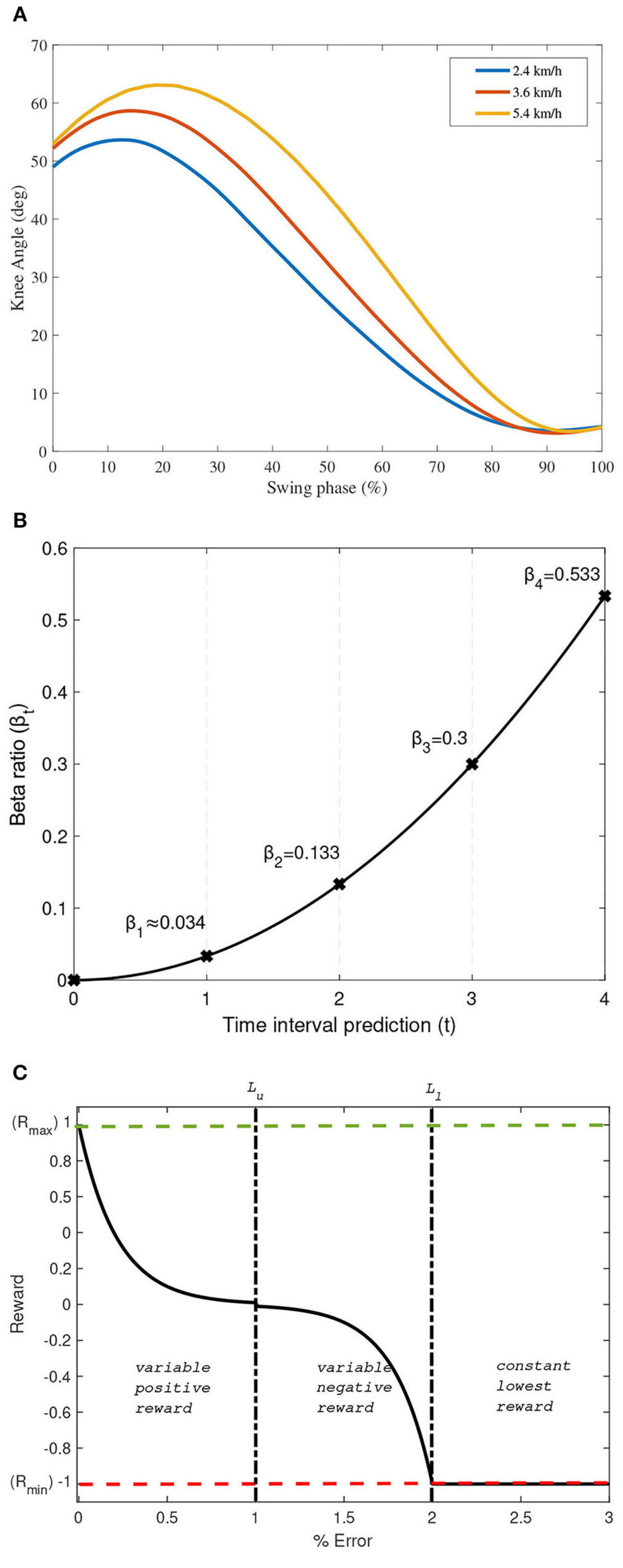


FIGURE 2 | (A) Average knee angle data used in this study. **(B)** β_t as an exponential function with $n = 4$. **(C)** Proposed reward shaping function as a function of E_t.

2.2. Q-Learning Control

Here, the proposed Q-learning control is discussed. Q-learning belongs to the tabular RL group in the machine learning algorithm. Generally, RL learns the control policies within a specified environment where the performance and training information are provided in terms of whether the applied control policy is a success or failure (Sutton and Barto, 2018). Success or failure in this case is determined by a certain performance index depending on the system and environment involved.

2.2.1. Q-Learning Structure

The general structure of RL is consisted of an agent and a system/environment. An agent executes an action, a_t, to the system and environment. Based on the given action, the system will react to another state, s_t, while also gives a reward, R_t, based on the performance index calculated from the current state. In this study, the agent is the Q-function with a mathematical description, as shown in Equation (4).

$$Q_{(s_t,a_t)} \leftarrow Q_{(s_t,a_t)} + \alpha\left[R_{(s_t,a_t)} + \gamma \max Q_{(s_{t+1},a_t)} - Q_{(s_t,a_t)}\right] \quad (4)$$

In Equation (4), Q and R are the action-value and reward functions, respectively. Further, $s, a, \alpha, and\ \gamma$ are the state, action, learning rate, and discounted rate, respectively, while subscript t denotes the time. Learning rate and discounted rate are dimensionless variables between 0 and 1. Higher learning rate, which if sets closer to 1, indicates that the Q-function is updated quickly per iteration, while the Q-function is never be updated if it is set to 0. The discounted factor is a variable that determines how the Q-function acts toward the reward. If it is set closer to 0 means, it will only consider the instantaneous reward, while if it is set closer to 1, it strives more into the long-term higher rewards (Sutton and Barto, 2018).

$$\begin{aligned} Q_{(\theta_{K(t)},\dot{\theta}_{K(t)},a_t)} \leftarrow\ & Q_{(\theta_{K(t)},\dot{\theta}_{K(t)},a_t)} + \alpha\big[R_t + \gamma \max Q_{(\theta_{K(t+1)},\dot{\theta}_{K(t+1)},a_t)} \\ & -Q_{(\theta_{K(t)},\dot{\theta}_{K(t)},a_t)}\big] \end{aligned} \quad (5)$$

In this study, Q-learning is proposed to be used as a controller of a dynamics system of the MR damper in the prosthetic knee in a double pendulum-simulated environment. The state is the parameter extracted from the environment that contains necessary information to be used to evaluate the control policies. In most cases, Q-function with multistate is used to better learn the environment (Fernandez-Gauna et al., 2013; Sadhu and Konar, 2017; Chai and Hayashibe, 2020). Particularly, this paper (Chai and Hayashibe, 2020) has explored deep RL for motion generation in a simulated environment. In this study, θ_K and derivative of knee angle, $\dot{\theta}_K$, are used as states, while the command voltage, v, is used as the action. Thus, the update rule of the Q-function can be written as in Equation (5). As Q-learning is following an off-policy method, actions were selected based on the maximum value of the Q-function on the current states, $\max Q_{(s_{1(t)},s_{2(t)})}$. Meanwhile at the initialization stage of learning, action selection follows a greedy policy to explore the Q-function

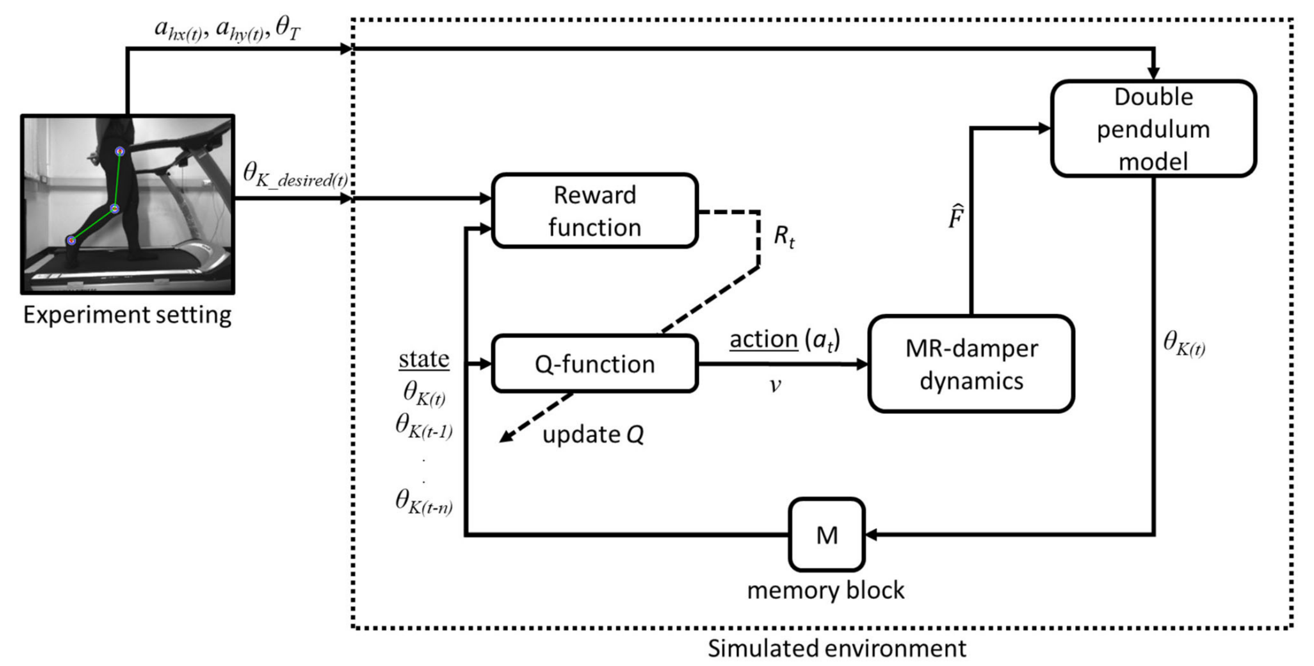


FIGURE 3 | Block diagram of the proposed Q-learning control in a simulated environment with external experiment data.

for possible solutions.

$$Q_{(\theta_{K(t)},\dot{\theta}_{K(t)},a_t)} \leftarrow Q_{(\theta_{K(t)},\dot{\theta}_{K(t)},a_t)} + \alpha \left[\sum_{t=1}^{n} \beta_t R_t + \gamma \max Q_{(\theta_{K(t+1)},\dot{\theta}_{K(t+1)},a_t)} - Q_{(\theta_{K(t)},\dot{\theta}_{K(t)},a_t)} \right] \quad (6)$$

2.2.2. Reward Shaping Function

The structure of the reward mechanism in the Q-learning algorithm used in this study is modified into a rationed multiple rewards as a function of time. This structure enables the learning process to provide more reward to latter horizon events due to the response time required by the MR damper to generate the necessary damping mechanism. The mathematical descriptions of this multiple reward mechanism are expressed in Equation (6), where $\beta_t = ct^2$ and $\sum_{t=1}^{n} \beta_t = 1$.

In Equation (6), β_t is the specifically designed ratio of reward priority, n is the number of prediction horizon, and c is a constant that depends on n. In this study, n is set to 4; thus, $c = 0.033$ to be conveniently compared to the NNPC algorithm studied in Ekkachai and Nilkhamhang (2016) that set the prediction horizon to 4. Further, the reward priority given at the specified prediction horizon is an exponential function, as depicted in **Figure 2B**.

As the controller aims to mimic the biological knee trajectory in the swing phase, the reward will be given according to whether the prosthetic knee can follow the biological knee trajectory. In this study, the reward is designed as a function of a performance index (*PI*). A simple absolute error, e_t, is selected as the performance index and evaluated per interval time. The reward function is also designed to have a continuous value over a specified boundary and follow a decaying exponential function. The mathematical descriptions of the proposed designed reward functions are expressed in Equations (7)–(11).

$$R_t = f(PI) \quad (7)$$

$$PI = e_t = \left| \frac{\theta_K - \theta_{K(val)}}{\theta_{K(val)}} \right| \quad (8)$$

$$R_t = \begin{cases} R_{max}\delta^{E_t} & ;\ 0 < E_t < L_u \quad (9) \\ R_{min}\delta^{|L_t - E_t|} & ;\ L_u < E_t < L_l \quad (10) \\ R_{min} & ;\ E_t > L_l \quad (11) \end{cases}$$

In Equations (7)–(11), $\theta_{K(val)}$ is the validation of knee angle at time t, R_{max}, and R_{min} are the maximum reward and minimum reward set to 1 and −1, respectively. E_t is the percentage of e_t, which can be written as $E_t = 100e_t$. Further, δ, L_u, *and* L_l are the reward constants set arbitrarily to 0.01, performance limit to obtain the positive reward, and performance limit to obtain the lowest reward, respectively. In this study, *PI* is aimed to be within 0.01, indicating that the error should be under 1%. Thus, L_u is set to be 1, and L_l could be set to any number larger than L_u to provide a variable negative reward. In this case, L_l is set to be twice the value of L_u.

The graphical description of this reward design is depicted in **Figure 2C**. Note that δ, L_u, L_l, R_{max}, and R_{min} can be defined accordingly for other applications depending on the system being evaluated. The reward shaping function is preferred to follow a decayed exponential function rather than a linear function to better train the *Q*-function to reach the state with the largest reward value, which can lead to faster convergence.

3. SIMULATION AND RESULTS

In this section, a simulation of swing phase control using the proposed controller is discussed along with a comparison study. The simulation was computed using Intel® Core^{TM} i7 6th Generation 3.5 GHz processor with 8 GB RAM. The overall diagram of our study is depicted in **Figure 3**. The figure shows an experiment setting that provide kinematics data of the subject and a simulated environment where our proposed framework is tested. On the simulated environment, we have a *Q*-function block with input of multistate of knee angle from double pendulum model and updated by the reward function. The input of the reward function are the actual knee angle $\theta_K(t)$ and the desired knee angle $\theta_{K(\text{desired})}(t)$ from experimental data. The output of *Q*-function is an action (a_t) in the form of control voltage (v) that is passed on to MR-damper dynamics block. The voltage is converted into $\hat{F}$ following **Figure 1A** and passed on to the double pendulum model for swing phase simulation.

There are several parameters in Q-learning control that must be defined and optimized. First, as this control approach is a tabular RL using the Q-learning method, each value of the *Q*-function is stored in a Q-matrix. The size of the Q-matrix depends on the number of states and actions. In this simulation, the structure of the Q-matrix is a three-dimensional matrix consisting of l rows of state $\theta_{K(t)}$, m columns of state $\dot{\theta}_{K(t)}$, and n layers of action v. Q-matrix must cover all the states and actions available on the system. Based on the data used, the state $\theta_{K(t)}$ is within the range of 0 and 70° with a predefined step size of 0.5°, resulting with 141 rows. State $\dot{\theta}_{K(t)}$ is set from -7 to 7° per unit of time with predefined 0.05 step size, thus resulting with 281 columns. The range of command voltage is set from 0 to 5 V with 0.1 resolution, resulting with 51 layers of action.

Second, learning rate α need to be defined. In this simulation, several values of learning rate are simulated to determine its effect on the number of iteration required to achieve best performance. The performance index used to evaluate this simulation is the normalized root mean squared error (*NRMSE*) as expressed in Equation (12), where n_s is the number of samples in dataset.

$$NRMSE = \frac{\sqrt{\frac{1}{n_s}\sum_{t=1}^{n_s}[\theta_{K(\text{desired})}(t) - \theta_K(t)]^2}}{\max(\theta_{K(\text{desired})}) - \min(\theta_{K(\text{desired})})} \qquad (12)$$

On the first simulation, we compared our reward shaping function as formulated in Equations (7)–(11) to a single reward mechanism expressed in Equation (4). We used 2.4, 3.6, and 5.4 km/h walking speed dataset, simulated separately with same value of randomized Q-matrix initialization. We then measured the moving average of *NRMSE* parameter with a constrained maximum iterations of 3000 and a fixed learning rate of 0.1. The results of this simulation are depicted in **Figure 4A**. It can be concluded from this simulation that the reward shaping function performed better over time in terms of *NRMSE*, compared to a single reward function.

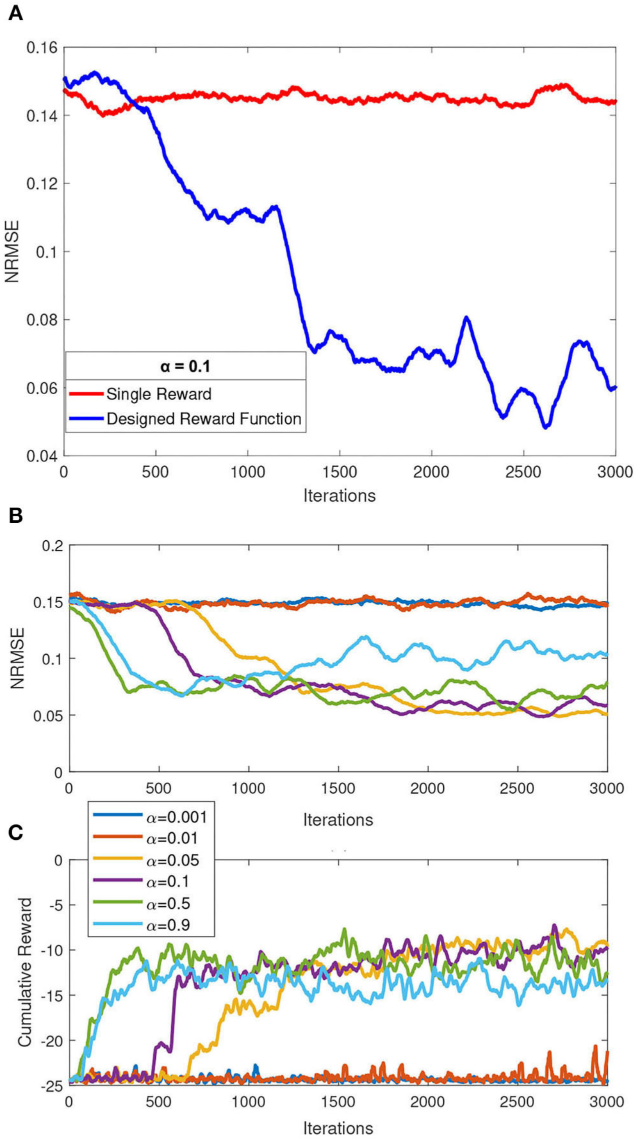


FIGURE 4 | Summary of simulation results over a constrained iteration of 3000. **(A)** Comparison of single reward mechanism and our proposed reward shaping function. **(B)** Effect of various learning rates to the overall performance (normalized root mean squared error, NRMSE). **(C)** Comparison of cumulative reward over iteration by each of the simulated learning rates.

In the second simulation, several values of learning rate $\alpha = [0.001, 0.01, 0.05, 0.1, 0.5, 0.9]$ are picked a priori to be simulated with a maximum 3000 iteration in a single speed simulation (mid speed of 3.6 km/h). For each learning rate, simulation was performed three times and average *NRMSE* for each learning rate were recorded. The effect of these learning rate to *NRMSE* is shown in **Figures 4B,C**. We concluded that the two lowest learning rate ($\alpha = 0.001$ and $\alpha = 0.01$) simulated with a

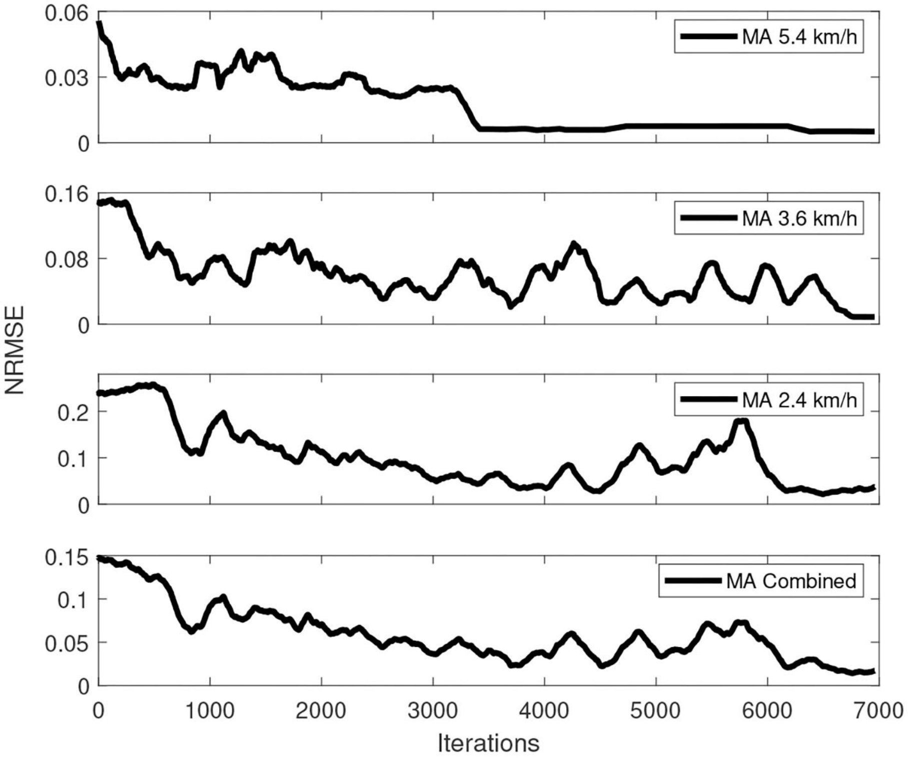


FIGURE 5 | Overall training process of multispeed of walking under one control policy simulation.

constrained iteration of 3,000 performed the worst among other learning rates. Those two learning rates did also not show any significant performance changes over the constrained iteration. As also observed, a higher learning rate does not guarantee better performance, as inspected from $\alpha = 0.9$, compared to $\alpha = [0.05, 0.1, 0.5]$. For the next simulation, we picked learning rate $\alpha = 0.5$ based on this simulation and considering faster exploration of Q-matrix that could potentially lead to finding better local minimum as solution.

There are many approaches to train the *Q*-function in this study. Training one *Q*-function for a specific case of a single walking speed is easy, while training multispeed at once under one *Q*-function is challenging. In this simulation, training multispeed under one control policy is proposed. Slowest, mid, and fast walking speeds of 2.4, 3.6, and 5.4 km/h, respectively, are used for training. In this simulation, the time interval is set to 20 ms; thus, the action or command voltage to the prosthetic knee is updated every 20 ms. The dataset of 2.4, 5.4, and 3.6 km/h is selected randomly for every iteration of the simulation. There are two conditions for the simulation to stop: first is if all the *NRMSE* of all trained speed falls under the defined *PI* criterion, and second is if all the trained speed converges into one final value of *NRMSE* for at least after 10 further iterations.

The best training process of this simulation over a total of 10 training processes is depicted in **Figure 5**. As shown in this figure, the fastest convergence was achieved by the fastest walking speed, which converges at around 3,300 iterations, followed by the walking speed of 3.6 km/h, which converges at around 6,700 iterations, and the latest is the slowest walking speed, which converges at around 6,900 iterations. This occurrence happened because a faster walking speed generally indicates a short time in the gait cycle, resulting in a less swing-phase time. The lesser time in the swing phase with a fixed control interval of 20 ms indicates that the *Q*-function calculates fewer actions than the slower walking speed dataset.

4. DISCUSSION

In this study, we investigated our proposed control algorithm for the swing phase controller in the MR-damper-based prosthetic knee. The proposed controller was designed with the structure of a tabular reinforcement Q-learning algorithm, a subset in machine learning algorithms. The Q-learning control comprised a *Q*-function that stores its value in a Q-matrix and a reward function following the reward shaping function proposed in this study. The advantages of using this control structure are that it can be trained online, and also it is a model-free control algorithm that does not require prior knowledge of the system to be controlled. A variable reward as a function of *PI* associating a decayed function, which is proposed as a reward function herein, has led to a better reward mechanism. We have shown that our proposed reward function demonstrated a trend of faster convergence compared to a single reward mechanism as depicted in **Figure 4A**.

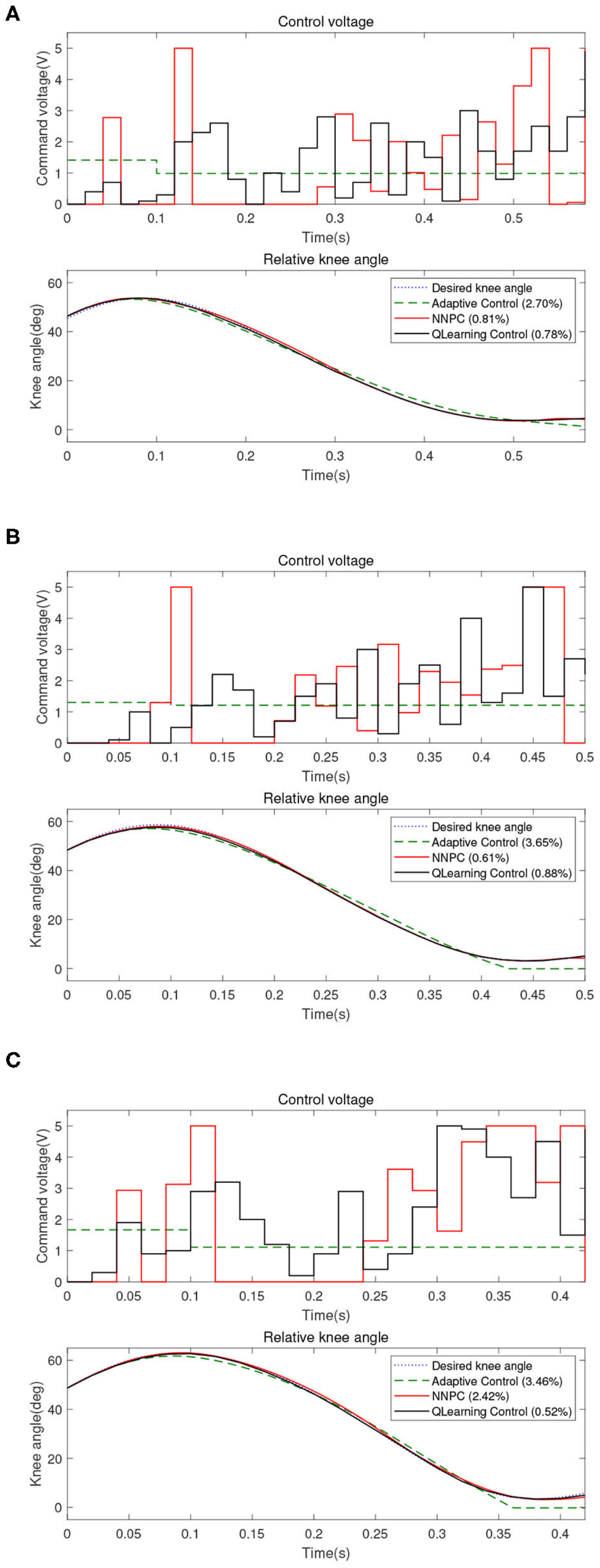


FIGURE 6 | Comparison between user-adaptive control (green dashed line), neural network predictive control (NNPC) (red line), and Q-learning control (black line) for different walking speeds: **(A)** 2.4 km/h, **(B)** 3.6 km/h, and **(C)** 5.4 km/h.

TABLE 1 | Comparison between user adaptive, neural network predictive control (NNPC), and Q-learning control.

Walking speed (km/h)	NRMSE(%)		
	User-adaptive	NNPC	Q-learning
2.4	2.70	0.81	0.78*
3.6	3.65	0.61*	0.88
5.4	3.46	2.42	0.52*
Average	3.27	1.28	0.73*

Best performance.

The proposed controller is then compared to the user-adaptive controller (Herr and Wilkenfeld, 2003) and the NNPC algorithm (Ekkachai and Nilkhamhang, 2016). The comparison of 2.4, 3.6, and 5.4 km/h walking speed are depicted in **Figure 6** and **Table 1**. The table depicts that for the walking speed of 2.4 km/h, Q-learning method performed the best with 0.78 of *NRMSE*, compared to NNPC (0.81) and user-adaptive control (2.70). Further, for the walking speed of 3.6 km/h, the best performance was achieved by NNPC with 0.61 of *NRMSE*, compared with Q-learning (0.88) and user-adaptive control (3.65). Lastly, for the walking speed of 5.4 km/h, Q-learning performed the best with the lowest *NRMSE* of 0.52, compared with NNPC (2.42) and user-adaptive control (3.46). Overall, Q-learning method perform within 1% of *NRMSE*, which followed the designed common reward function for different walking speed.

This control structure also shows adaptability to various walking speeds. Moreover, we have successfully trained a unified control policy for every simulated walking speed. *PI* verified with the experimental result indicates that this control structure performs better than the user-adaptive control. Moreover, in some of the walking speeds, this control structure performs better than the NNPC algorithm. The total performance over different walking speeds showed promising results by using the proposed approach.

In terms of cost function, knee trajectory is only one of the parameters to be optimized among other correlated systems, such as ankle and foot prostheses, to achieve better gait symmetry and reduce metabolic costs. Although there has not been a detailed study about the acceptable criterion in terms of the *NRMSE* performance index of the knee trajectory in a prosthetic knee, this study aims to mimic the biological knee trajectory, which is shown by *PI*.

On the applicability point of view, our proposed Q-learning control had no prior knowledge of the structure and characteristics of MR-damper. Signals observed by Q-learning control were the states of knee angle and its derivatives, as well as the reward signal R_t that was given based on the performance of the controller to shape the control policy. Based on this facts, our proposed Q-learning control can potentially be used for other structure of MR-damper or even other impedance-based machine for semi-active prosthetic.

Although we cannot provide detailed comparison of our proposed method with another RL-based method in Wen et al.

(2019), a brief comparison is discussed as follows. The ADP-based RL algorithm resulted in 2.5° of *RMSE* on the robotic knee kinematics. The average performance of our proposed method was 0.73 of *NRMSE* or was 1.59° if converted to average *RMSE*. Conversely, in this study, we employed the RL algorithm to control the output of the control voltage for the MR damper, resulting in only one simple output variable. Meanwhile, this existing study (Wen et al., 2019) used the RL algorithm to tune a total of 12 impedance parameters of the robotic knee; thus, the output variables are 12. We also treated the swing phase as one state, while in Wen et al. (2019), the swing phase was divided into swing flexion and swing extension where the ADP tuner would tune the impedance parameters accordingly with respect to each state.

In this study, we focused on developing a unique control that can adapt and accommodate a range of subject-specific walking speed. Unique means that it can only be valid for the subject. The reason was, like any other prosthetic, it is tuned personally to the wearer. In this study, the control policy that we train is valid only for the subject whose data we used. However, the idea of our proposed control framework and algorithm can be applied to other subjects.

While it has shown a promising result, we also identified some of the limitations of our study. Using the computational hardware mentioned at the previous section and source code implemented in MATLAB, the overall calculation and online update *Q*-function process consumed approximately 40.4 ms, while each evaluation of NNPC with pretrained swing phase model consumed approximately 13.2 ms (Ekkachai and Nilkhamhang, 2016). Changing the source code implementation in C language and using dedicated processing hardware could shorten the calculation time to be within the proposed control interval of 20 ms.

There are several areas that can be explored for future works. First, another training strategy can be explored further to shorten the calculation time. Second, this study proposed a tabular-discretized *Q*-function stored in a Q-matrix. A continuous *Q*-function could also be explored to better cover all the states and actions. Third is to test our proposed control strategy to other subjects and possibly to test a transfer learning approach from control policy that was learnt in this study for dataset from other subjects.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: datasets analyzed in this article are available upon request. Requests to access these datasets should be directed to kittipong.ekkachai@nectec.or.th.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YH contributed to algorithm design and development, data analysis and interpretation, and writing the first draft. KE supported the development of the system and environment model, collecting datasets, and data analysis. MH provided critical review and contributed additional texts to the draft. WK contributed to study conception and design, provided critical review, and supervised the overall study. All authors read, reviewed, and approved the final manuscript.

FUNDING

This work was partly funded by the G-7 Scholarship Foundation and was also supported by the Data Sciences Program II (DSP II) of Tohoku University.

REFERENCES

Chai, J., and Hayashibe, M. (2020). Motor synergy development in high-performing deep reinforcement learning algorithms. *IEEE Robot. Autom. Lett.* 5, 1271–1278. doi: 10.1109/LRA.2020.2968067

Ekkachai, K., and Nilkhamhang, I. (2016). Swing phase control of semi-active prosthetic knee using neural network predictive control with particle swarm optimization. *IEEE Trans. Neural Syst. Rehabil. Eng.* 24, 1169–1178. doi: 10.1109/TNSRE.2016.2521686

Ekkachai, K., Tungpimolrut, K., and Nilkhamhang, I. (2012). A novel approach to model magneto-rheological dampers using EHM with a feed-forward neural network. *Science Asia* 38, 386–393. doi: 10.2306/scienceasia1513-1874.2012.38.386

Ekkachai, K., Tungpimolrut, K., and Nilkhamhang, I. (2013). Force control of a magnetorheological damper using an elementary hysteresis model-based feedforward neural network. *Smart Mater. Struct.* 22:5030. doi: 10.1088/0964-1726/22/11/115030

Fernandez-Gauna, B., Marques, I., and Graña, M. (2013). Undesired state-action prediction in multi-agent reinforcement learning for linked multi-component robotic system control. *Inform. Sci.* 232, 309–324. doi: 10.1016/j.ins.2012.12.021

Flynn, L., Geeroms, J., Jimenez-Fabian, R., Vanderborght, B., and Lefeber, D. (2015). CYBERLEGS Beta-Prosthesis active knee system. *IEEE Int. Conf. Rehabil. Robot.* 2015, 410–415. doi: 10.1109/ICORR.2015.7281234

Hafner, B. J., Willingham, L. L., Buell, N. C., Allyn, K. J., and Smith, D. G. (2007). Evaluation of function, performance, and preference as transfemoral amputees transition from mechanical to microprocessor control of the prosthetic knee. *Arch. Phys. Med. Rehabil.* 88, 207–217. doi: 10.1016/j.apmr.2006.10.030

Herr, H., and Wilkenfeld, A. (2003). User-adaptive control of a magnetorheological prosthetic knee. *Indus. Robot* 30, 42–55. doi 10.1108/01439910310457706

Hoover, C. D., Fulk, G. D., and Fite, K. B. (2012). The design and initial experimental validation of an active myoelectric transfemoral prosthesis. *J. Med. Devices Trans.* 6:011005. doi: 10.1115/1.4005784

Kaufman, K. R., Frittoli, S., and Frigo, C. A. (2012). Gait asymmetry of transfemoral amputees using mechanical and microprocessor-controlled prosthetic knees. *Clin. Biomech.* 27, 460–465. doi: 10.1016/j.clinbiomech.2011.11.011

Kaufman, K. R., Levine, J. A., Brey, R. H., Iverson, B. K., McCrady, S. K., Padgett, D. J., et al. (2007). Gait and balance of transfemoral amputees using passive mechanical and microprocessor-controlled prosthetic knees. *Gait Post.* 26, 489–493. doi: 10.1016/j.gaitpost.2007.07.011

Lawson, B. E., Mitchell, J., Truex, D., Shultz, A., Ledoux, E., and Goldfarb, M. (2014). A robotic leg prosthesis: design, control, and implementation. *IEEE Robot. Autom. Mag.* 21, 70–81. doi: 10.1109/MRA.2014.2360303

Putnam, C. A. (1991). A segment interaction analysis of proximal-to-distal sequential segment motion. *Med. Sci. Sports Exerc.* 23, 130–144.

Quintero, D., Martin, A. E., and Gregg, R. D. (2017). Toward unified control of a powered prosthetic leg: a simulation study. *IEEE Trans. Control Syst. Technol.* 26, 305–312. doi: 10.1109/TCST.2016.2643566

Sadhu, A. K., and Konar, A. (2017). Improving the speed of convergence of multi-agent Q-learning for cooperative task-planning by a robot-team. *Robot. Auton. Syst.* 92, 66–80. doi: 10.1016/j.robot.2017.03.003

Şahin, I., Engin, T., and Cemeci, S. (2010). Comparison of some existing parametric models for magnetorheological fluid dampers. *Smart Mater. Struct.* 19:035012. doi: 10.1088/0964-1726/19/3/035012

Sawers, A. B., and Hafner, B. J. (2013). Outcomes associated with the use of microprocessor-controlled prosthetic knees among individuals with unilateral transfemoral limb loss: a systematic review. *J. Rehabil. Res. Dev.* 50, 273–314. doi: 10.1682/jrrd.2011.10.0187

Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Wen, Y., Si, J., Brandt, A., Gao, X., and Huang, H. (2019). Online reinforcement learning control for the personalization of a robotic knee prosthesis. *IEEE Trans. Cybern.* 50, 2346–2356. doi: 10.1109/TCYB.2019.2890974

Wen, Y., Si, J., Gao, X., Huang, S., and Huang, H. H. (2017). A new powered lower limb prosthesis control framework based on adaptive dynamic programming. *IEEE Trans. Neural Netw. Learn. Syst.* 28, 2215–2220. doi: 10.1109/TNNLS.2016.2584559

Windrich, M., Grimmer, M., Christ, O., Rinderknecht, S., and Beckerle, P. (2016). Active lower limb prosthetics: a systematic review of design issues and solutions. *Biomed. Eng. Online* 15:140. doi: 10.1186/s12938-016-0284-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurorobotics

ORIGINAL RESEARCH
published: 06 January 2021
doi: 10.3389/fnbot.2020.617293

Matrix Inequalities Based Robust Model Predictive Control for Vehicle Considering Model Uncertainties, External Disturbances, and Time-Varying Delay

Wenjun Liu[1], *Guang Chen*[2,1]* *and Alois Knoll*[1]

[1] *Department of Informatics, Technical University of Munich, Munich, Germany,* [2] *School of Automotive Studies, Tongji University, Shanghai, China*

In this paper, we design a robust model predictive control (MPC) controller for vehicle subjected to bounded model uncertainties, norm-bounded external disturbances and bounded time-varying delay. A Lyapunov-Razumikhin function (LRF) is adopted to ensure that the vehicle system state enters in a robust positively invariant (RPI) set under the control law. A quadratic cost function is selected as the stage cost function, which yields the upper bound of the infinite horizon cost function. A Lyapunov-Krasovskii function (LKF) candidate related to time-varying delay is designed to obtain the upper bound of the infinite horizon cost function and minimize it at each step by using matrix inequalities technology. Then the robust MPC state feedback control law is obtained at each step. Simulation results show that the proposed vehicle dynamic controller can steer vehicle states into a very small region near the reference tracking signal even in the presence of external disturbances, model uncertainties and time-varying delay. The source code can be downloaded on https://github.com/wenjunliu999.

Keywords: robust model predictive control, vehicle dynamic control, matrix inequities, robust positively invariant, model uncertainties, external disturbances, time-varying delay

Edited by:
Yan Wu,
Institute for Infocomm Research
*(A*STAR), Singapore*

Reviewed by:
Wenyu Liang,
National University of Singapore,
Singapore
Yu Cao,
Huazhong University of Science and
Technology, China

Correspondence:
Guang Chen
guangchen@tongji.edu.cn

Received: *14 October 2020*
Accepted: *04 December 2020*
Published: *06 January 2021*

Citation:
Liu W, Chen G and Knoll A (2021) Matrix Inequalities Based Robust Model Predictive Control for Vehicle Considering Model Uncertainties, External Disturbances, and Time-Varying Delay. Front. Neurorobot. 14:617293. doi: 10.3389/fnbot.2020.617293

1. INTRODUCTION

Dynamic control is one of the most crucial tasks for autonomous driving vehicle (Chen et al., 2020). H-infinity output feedback controller (Hu et al., 2016), sliding mode controller (Jiang and Wu, 2018), model predictive control (MPC) (Sun et al., 2019), etc. have been designed for vehicle control. Because MPC has the ability to incorporate soft and hard constraints into the online optimizations in a multivariable control framework, it is widely applied in vehicle control field (Hu et al., 2019).

Vehicle dynamic control performance may be seriously affected by external disturbances resulting from unpredictable environment (Yu et al., 2019). Besides, it is impossible to establish a precise vehicle model when the vehicle is moving (Liu et al., 2019). Therefore, robust model predictive control for vehicle is researched by scholars. Robust feedback MPC (Shamaghdari et al., 2015), tube-based robust MPC (Sakhdari and Azad, 2018), min-max robust MPC (Wang X. et al., 2016), feedback min-max MPC (Liu et al., 2014) and linear matrix inequities (LMIs) or bilinear matrix inequities (BMIs) based robust MPC (Cheng et al., 2020) have been proposed to deal with vehicle control where the vehicle has model uncertainties or external disturbances or both model uncertainties and external disturbances. The optimization problems involving LMIs or BMIs are convex and hence have solvable global optimal solutions and can be solved numerically efficiently and reliably (Duan and Yu, 2013). So, LMIs or BMIs based robust MPC control is widely applied to vehicle control. However, most of existing LMIs or BMIs based robust MPC vehicle control papers only consider the model uncertainties. Few matrix inequalities based robust MPC papers consider both model uncertainties and external disturbances of vehicle.

Moreover, time delay can be frequently seen in vehicle system (Ren et al., 2019) and affects the vehicle dynamic control performance and even threatens the stability and safety of the vehicle system. Some literatures concerned MPC controller for time-delay vehicle system have been addressed (Liu and Li, 2019; Wang et al., 2019). However, most existing papers often assume that the delay is known and fixed (Yu et al., 2018), or do not consider the model uncertainties (Xu et al., 2020) or the external disturbances (Bououden et al., 2016; Nahidi et al., 2019). Therefore, the research on robust MPC controller for vehicle with time-varying delay has not been completely investigated and hence several problems still remain unsolved.

Due to the fact that the influence of the delayed states can cause a violation of the monotonic decrease condition that a standard Lyapunov function obeys, systems affected by delays can not apply the classical Lyapunov theory directly (Gielen et al., 2012a). Generally, there are two types of approaches to deal with time-delay systems, Lyapunov-Razumikhin function (LRF) and Lyapunov-Krasovskii function (LKF) (Teng, 2018). In our paper, we focus on discrete vehicle dynamic control. For discrete-time systems, LKF makes use of an augmentation of the state vector with all delayed states, which yields the applications of classical Lyapunov methods to an augmented system without delay (Teng et al., 2017). LRF is constructed for time-delay system based on a type of small-gain condition (Teel, 1998) and can be considered as a special case of LKF. Compared to LKF, LRF is conservative but its computational complexity is lower than LKF (Gielen et al., 2012b).

In this paper, the aim of dynamic control is to guarantee the vehicle dynamic state tracking performance, where the vehicle dynamic state tracking reference signal is determined by the upper kinematic control and assumed to be known in this paper. To suppress the influence of model uncertainties, external disturbances, and time-varying delay on vehicle dynamic state tracking performance, we design a matrix inequalities based robust MPC controller. It is known that robust positively invariant (RPI) set plays an crucial role in robust MPC, the control law involved in RPI can ensure that when the system state enters the RPI set, it never goes out (Yang and Feng, 2013). A LRF candidate is adopted to guarantee this. Since the infinite horizon cost function can not be optimized online, a cost function with finite terms is usually considered instead. To compute the bound of the infinite horizon cost function, a LKF candidate related to time-varying delay is designed in this paper. Then, the robust MPC state feedback control law is obtained by minimizing the upper bound at each step using matrix inequalities technology.

Different from existing researches, there are two main contributions of this paper. Firstly, the vehicle dynamic model simultaneously considers model uncertainties, external disturbances and the time-varying delay of the vehicle state, which is more in line with actual vehicle operating conditions. Then a robust MPC controller is designed to steer vehicle states into a very small region near the reference tracking signal even in the presence of external disturbances, model uncertainties and time-varying delay. Secondly, compared to the robust controller considering fixed time delay, a delay-range-dependent LKF is designed by using the information of the upper and lower bounds of the time-varying delay and relaxed technique, which also simultaneously takes model uncertainties, external disturbances and the time-varying delay into account.

The rest of this paper is structured as follows. In section 2, two auxiliary lemmas and vehicle model are introduced. In section 3, the proposed robust MPC controller for vehicle subjected to bounded model uncertainties, norm-bounded external disturbances and bounded time-varying delay is designed. In section 4, simulation examples are illustrated to verify the effectiveness of the proposed method. Finally, we conclude in section 5.

Notations: $\boldsymbol{I}$ is an identity matrix with appropriate dimension. diag$\{\cdots\}$ denotes a block-diagonal matrix, $\|\cdot\|$ denotes 2-norm. $\boldsymbol{X} > 0$ and $\boldsymbol{X} \geq 0$ denotes the matrix is a positive definite matrix and positive semi-definite matrix, respectively. The symbol $*$ induces a symmetric structure or a transpose item, e.g., when $\boldsymbol{H}$ and $\boldsymbol{R}$ are symmetric matirces, then

$$\begin{bmatrix} \boldsymbol{H} & * \\ \boldsymbol{S} & \boldsymbol{R} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H} & \boldsymbol{S}^{\mathrm{T}} \\ \boldsymbol{S} & \boldsymbol{R} \end{bmatrix}$$

when the expression has the format $\boldsymbol{Q} + \boldsymbol{S} + \boldsymbol{S}^{\mathrm{T}}$, we simplify it to $\boldsymbol{Q} + \boldsymbol{S} + *$.

2. PRELIMINARIES

In this section, we first introduce two necessary lemmas and then derive the dynamic model of the vehicle with model uncertainties, external disturbances, and time-varying delay.

2.1. Auxiliary Lemmas

Lemma 1: (Duan and Yu, 2013) Let $\boldsymbol{X} \in \boldsymbol{R}^{m\times n}$, $\boldsymbol{Y} \in \boldsymbol{R}^{n\times m}$, $\boldsymbol{Q} \in \boldsymbol{R}^{m\times m}$. Then

$$\boldsymbol{Q} + \boldsymbol{X}\boldsymbol{F}\boldsymbol{Y} + \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}} < 0 \tag{1}$$

holds for all $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} \leq \boldsymbol{I}$ if and only if there exits a positive scalar δ, such that

$$\boldsymbol{Q} + \delta\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} + \delta^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y} < 0 \tag{2}$$

Lemma 2: (Jiang and Wang, 2001) If an input-to-state (ISS) Lyapunov function can be found for system $\boldsymbol{x}(k+1) = f(\boldsymbol{x}(k), \boldsymbol{p}(k))$, then the system is ISS.

The definition of ISS-Lyapunov function is as follows.

If a continuous function $V : \boldsymbol{R}^n \rightarrow \boldsymbol{R}_{\geq 0}$ satisfy the following inequalities, then it is an ISS-Lyapunov function for system $\boldsymbol{x}(k+1) = f(\boldsymbol{x}(k), \boldsymbol{p}(k))$ (Yang and Feng, 2013).

$$\beta_1(\|\boldsymbol{x}(k)\|) \leq V(\boldsymbol{x}(k)) \leq \beta_2(\|\boldsymbol{x}(k)\|) \tag{3}$$

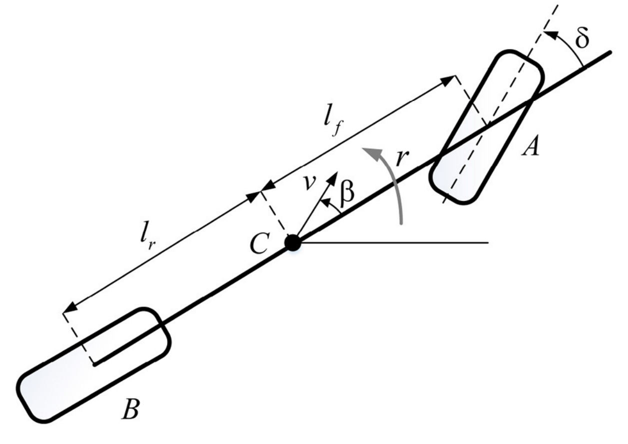


FIGURE 1 | Two-degree-of-freedom bicycle model of vehicle.

$$V(\boldsymbol{x}(k+1)) - V(\boldsymbol{x}(k)) \leq -\beta_3(\|\boldsymbol{x}(k)\|) + \phi(\|\boldsymbol{p}(k)\|) \tag{4}$$

where β_1, β_2, and β_3 are K_∞-function, ϕ is K-function, $\boldsymbol{x}$ denote system state vector and $\boldsymbol{p}$ denote external disturbances.

2.2. Vehicle Dynamic Model

An extensively used two-degree-of-freedom bicycle model of vehicle dynamics is adopted in this paper and is shown in **Figure 1**. This model is based on the assumption of small slip angles, no road gradient or bank angles, no load transfer and no rolling or pitching moment. Only the front wheel is steerable in this case (Elbanhawi et al., 2018). The two front wheels and the two rear wheels are respectively represented by one front wheel A and one rear wheel B. The center of gravity (c.g.) of the vehicle is at point C. The distances of A and B from the c.g. of the vehicle are l_f and l_r respectively. v is the velocity at the c.g. of the vehicle, β is the sideslip angle, δ is the steering angle and r is the yaw rate. The dynamic model with regard of yaw rate and sideslip angle is adopted in this paper, as presented in (1). The detailed derivation process can be found in Rajamani (2011). m is the mass of the vehicle, $C_{\alpha f}$ and $C_{\alpha r}$ are the cornering stiffness of the front and rear tires respectively, I_z is yaw moment of inertia. v_x is vehicle longitudinal velocity. Similar to Cho and Huh (2019), a fixed forward velocity assumption is adopted in this paper.

$$\begin{cases} \dot{\beta} = -r + \frac{C_{\alpha f}}{m v_x}(\delta - \beta - \frac{l_f r}{v_x}) + \frac{C_{\alpha r}}{m v_x}(-\beta + \frac{l_f r}{v_x}) \\ \dot{r} = \frac{l_f C_{\alpha f}}{I_z}(\delta - \beta - \frac{l_f r}{v_x}) - \frac{l_r C_{\alpha r}}{I_z}(-\beta + \frac{l_f r}{v_x}) \end{cases} \tag{5}$$

Let $\boldsymbol{x} = [\,\beta \;\; r\,]^{\mathrm{T}}$, then the state space form of lateral dynamics model can be denoted as follows:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}_{cm}\boldsymbol{x} + \boldsymbol{B}_{cm}\boldsymbol{\delta} \tag{6}$$

$$\boldsymbol{A}_{cm} = \begin{pmatrix} -\frac{C_{\alpha f}+C_{\alpha r}}{m v_x} & -1 + \frac{-l_f C_{\alpha f}+l_r C_{\alpha r}}{m v_x^2} \\ \frac{-l_f C_{\alpha f}+l_r C_{\alpha r}}{I_z} & -\frac{l_f^2 C_{\alpha f}+l_r^2 C_{\alpha r}}{I_z v_x} \end{pmatrix} \tag{7}$$

$$\boldsymbol{B}_{cm} = \begin{pmatrix} \frac{C_{\alpha f}}{m v_x} & \frac{l_f C_{\alpha f}}{I_z} \end{pmatrix}^{\mathrm{T}} \tag{8}$$

We discretize the continuous model (6) and consider norm-bounded external disturbances $\boldsymbol{p}$. $\boldsymbol{p}$ satisfies $\|\boldsymbol{p}\| \leq \rho$. Then the system model can be written as follows:

$$\boldsymbol{x}(k+1) = \boldsymbol{A}_m\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{\delta}(k) + \boldsymbol{E}\boldsymbol{p}(k) \tag{9}$$

where $\boldsymbol{E}$ is a constant coefficient matrix, $\boldsymbol{A}_m$ and $\boldsymbol{B}$ can be calculated by the Euler method as:

$$\boldsymbol{A}_m = e^{\boldsymbol{A}_{cm}T}, \boldsymbol{B} = \int_{kT}^{(k+1)T} e^{\boldsymbol{A}_{cm}[(k+1)T-\eta]}\boldsymbol{B}_{cm}d\eta \tag{10}$$

where T is the sampling interval for the discrete state space model.

Then we consider time-varying delay, the system model can be written as follows:

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{A}_d\boldsymbol{x}_d(k) + \boldsymbol{B}\boldsymbol{\delta}(k) + \boldsymbol{E}\boldsymbol{p}(k) \tag{11}$$

where $\boldsymbol{A} = \alpha\boldsymbol{A}_m$, $\boldsymbol{A}_d = (1-\alpha)\boldsymbol{A}_m$, $\boldsymbol{x}_d(k) = \boldsymbol{x}(k-d_k)$, d_k is a time-varying delay bounded in $[d_m, d_M]$, d_m, d_M are known positive integers, the constant α is the retarded coefficient ranging $\alpha \in [0, 1]$, the limits 1 and 0 correspond to no delay term and to a completed delay term, respectively (Jeong and Park, 2005).

It is almost impossible for us to establish a precise model of a vehicle especially when it is moving. So model uncertainties must be considered during the vehicle controller design process.

$$\boldsymbol{x}(k+1) = \widetilde{\boldsymbol{A}}\boldsymbol{x}(k) + \widetilde{\boldsymbol{A}}_d\boldsymbol{x}_d(k) + \widetilde{\boldsymbol{B}}\boldsymbol{\delta}(k) + \boldsymbol{E}\boldsymbol{p}(k) \tag{12}$$

where $\widetilde{\boldsymbol{A}} = \boldsymbol{A}+\boldsymbol{M}\boldsymbol{H}\boldsymbol{N}_A$, $\widetilde{\boldsymbol{A}}_d = \boldsymbol{A}_d+\boldsymbol{M}\boldsymbol{H}\boldsymbol{N}_{Ad}$, $\widetilde{\boldsymbol{B}} = \boldsymbol{B}+\boldsymbol{M}\boldsymbol{H}\boldsymbol{N}_B$, $\boldsymbol{M}$, $\boldsymbol{N}_A$, $\boldsymbol{N}_{Ad}$ and $\boldsymbol{N}_B$ are constant matrices, $\boldsymbol{H}$ is Lebesgue measurable and satisfies $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} \leq \boldsymbol{I}$.

3. ROBUST MPC CONTROLLER DESIGN USING MATRIX INEQUALITIES

In this section, we derive the matrix inequalities involved in the robust MPC controller. MPC is actually an online optimization algorithm. At each iteration time, MPC controller optimizes a cost function and satisfies the corresponding constraints. We first introduce the design of online MPC controller, then we discuss the RPI conditions and develop the computation of a RPI set. Finally, we construct the whole control algorithm.

3.1. Online Robust MPC Design

In this paper, the quadratic cost function with infinite length is defined as

$$J_\infty(k) = \sum_{i=0}^{\infty} \ell(k+i \mid k) \tag{13}$$

where $\ell(k+i \mid k) = \boldsymbol{x}^{\mathrm{T}}(k+i \mid k)\boldsymbol{Q}\boldsymbol{x}(k+i \mid k) + \boldsymbol{\delta}^{\mathrm{T}}(k+i \mid k)\boldsymbol{R}\boldsymbol{\delta}(k+i \mid k) - \tau\boldsymbol{p}^{\mathrm{T}}(k+i \mid k)\boldsymbol{p}(k+i \mid k)$, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are positive definite matrices, τ is a positive constant. $\boldsymbol{x}(k+i \mid k)$

and $\boldsymbol{\delta}(k+i \mid k)$ represent the i-th predictive state and control input at time k respectively. The choice of $l(\cdot)$ is inspired by H_∞ MPC (Mayne et al., 2000) and has already been applied to existing robust MPC research (Yang et al., 2013).

The goal is to find a stabilizing state feedback control $\boldsymbol{\delta}(k) = \boldsymbol{K}\boldsymbol{x}(k)$ for system (12) by using the robust MPC strategy, the online robust MPC optimization can be summarized as:

$$\begin{gathered} \min_{\boldsymbol{\delta}(k+i|k)} \max_{\boldsymbol{p}(k+i|k)} J_\infty(k) \\ \text{s.t.} \quad (12) \\ -\boldsymbol{\delta}_{\max} \leq \boldsymbol{\delta}(k+i \mid k) \leq \boldsymbol{\delta}_{\max} \\ \|\boldsymbol{p}\| \leq \gamma \end{gathered} \tag{14}$$

where $\boldsymbol{\delta}_{\max}$ is the input constraint.

Considering $J_\infty(k)$ can not be minimized directly in reality, an upper bound of the infinite horizon cost function is minimized instead. To find an upper bound of $J_\infty(k)$, a LKF candidate related to time-varying delay is designed as follows:

$$V(\boldsymbol{x}(k)) = V_1(\boldsymbol{x}(k)) + V_2(\boldsymbol{x}(k)) + V_3(\boldsymbol{x}(k)) \tag{15}$$

$$V_1(\boldsymbol{x}(k)) = \boldsymbol{x}^{\mathrm{T}}(k)\boldsymbol{P}\boldsymbol{x}(k) \tag{16}$$

$$V_2(\boldsymbol{x}(k)) = \sum_{i=k-d_k}^{k-1} \boldsymbol{x}^{\mathrm{T}}(i)\boldsymbol{P}_d\boldsymbol{x}(i) \tag{17}$$

$$V_3(\boldsymbol{x}(k)) = \sum_{j=-d_M+1}^{-d_m} \sum_{i=k+j}^{k-1} \boldsymbol{x}^{\mathrm{T}}(i)\boldsymbol{P}_d\boldsymbol{x}(i) \tag{18}$$

where $\boldsymbol{P} = \boldsymbol{P}^{\mathrm{T}} > 0$, $\boldsymbol{P}_d = \boldsymbol{P}_d^{\mathrm{T}} > 0$. And suppose the following inequality is satisfied.

$$V(\boldsymbol{x}(k+i+1 \mid k)) - V(\boldsymbol{x}(k+i \mid k)) \leq -l(k+i \mid k) \tag{19}$$

As shown in Kothare et al. (1996), for the robust performance objective function to be finite, we must have $\boldsymbol{x}(\infty \mid k) = 0$ and hence $V(\boldsymbol{x}(\infty \mid k)) = 0$. Summing (19) from $i = 0$ to $i = \infty$, we get

$$J_\infty(k) \leq V(\boldsymbol{x}(k \mid k)) \leq \xi(k) \tag{20}$$

where $\xi(k) > 0$ is the upper bound of $J_\infty(k)$.

Theorem 1: If there exit matrices with appropriate dimension $\boldsymbol{X} = \boldsymbol{X}^{\mathrm{T}} > 0$, $\boldsymbol{X}_d = \boldsymbol{X}_d^{\mathrm{T}} > 0$, a general matrix $\boldsymbol{Y}$ and a scalar $\eta > 0$, then the following LMIs can be derived to guarantee (19) i.e., to guarantee $V(\boldsymbol{x}(k \mid k))$ is the upper bound of $J_\infty(k)$.

$$\begin{bmatrix} -\boldsymbol{X} & * & * & * & * & * & * & * & * \\ 0 & -\boldsymbol{X}_d & * & * & * & * & * & * & * \\ 0 & 0 & -\tau\xi\boldsymbol{I} & * & * & * & * & * & * \\ \boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Y} & \boldsymbol{A}_d\boldsymbol{X}_d & \xi\boldsymbol{E} & -\boldsymbol{X} & * & * & * & * & * \\ \boldsymbol{X} & 0 & 0 & 0 & -\left(d_s+1\right)^{-1}\boldsymbol{X}_d & * & * & * & * \\ \boldsymbol{Q}\boldsymbol{X} & 0 & 0 & 0 & 0 & -\xi\boldsymbol{Q} & * & * & * \\ \boldsymbol{R}\boldsymbol{Y} & 0 & 0 & 0 & 0 & 0 & -\xi\boldsymbol{R} & * & * \\ \boldsymbol{N}_A\boldsymbol{X}+\boldsymbol{N}_B\boldsymbol{Y} & \boldsymbol{N}_{Ad}\boldsymbol{X}_d & 0 & 0 & 0 & 0 & 0 & -\eta\boldsymbol{I} & * \\ 0 & 0 & 0 & \eta\boldsymbol{M}^{\mathrm{T}} & 0 & 0 & 0 & 0 & -\eta\boldsymbol{I} \end{bmatrix} < 0 \tag{21}$$

where $d_s = d_M - d_m$.

Proof: See **Appendix A**.

Equation (21) can guarantee $\xi(k)$ is the upper bound of $J_\infty(k)$, then we minimize it.

Theorem 2: $\xi(k)$ can be minimized using the following LMIs.

$$\begin{bmatrix} 1 & * \\ \boldsymbol{\zeta}_3(k) & \boldsymbol{\Lambda}_3 \end{bmatrix} \geq 0 \tag{22}$$

where $\boldsymbol{\zeta}_3(k) = \left[\boldsymbol{x}^{\mathrm{T}}(k), \boldsymbol{x}^{\mathrm{T}}(k-1), \cdots, \boldsymbol{x}^{\mathrm{T}}\left(k-d_m\right), \boldsymbol{x}^{\mathrm{T}}\left(k-d_m-1\right), \cdots, \boldsymbol{x}^{\mathrm{T}}\left(k-d_M\right)\right]^{\mathrm{T}}$, $\boldsymbol{\Lambda}_3 = \mathrm{diag}\left(\boldsymbol{P}, \frac{\boldsymbol{P}_d}{d_s+1}, \frac{\boldsymbol{P}_d}{d_s}, \cdots, \boldsymbol{P}_d\right)$.

Proof:

$$\begin{aligned} V_2(\boldsymbol{x}(k)) &= \sum_{i=k-d_k}^{k-1} \boldsymbol{x}^{\mathrm{T}}(i)\boldsymbol{P}_d\boldsymbol{x}(i) \leq \sum_{i=k-d_M}^{k-1} \boldsymbol{x}^{\mathrm{T}}(i)\boldsymbol{P}_d\boldsymbol{x}(i) \\ &= \boldsymbol{\zeta}_1^{\mathrm{T}}(k-1)\boldsymbol{\Lambda}_1\boldsymbol{\zeta}_1(k-1) \end{aligned} \tag{23}$$

where $\boldsymbol{\zeta}_1(k-1) = \left[\boldsymbol{x}^{\mathrm{T}}(k-1), \cdots, \boldsymbol{x}^{\mathrm{T}}\left(k-d_M\right)\right]^{\mathrm{T}}$, $\boldsymbol{\Lambda}_1 = \mathrm{diag}\left(\boldsymbol{P}_d, \cdots, \boldsymbol{P}_d\right)$.

$$V_3(\boldsymbol{x}(k)) = \sum_{j=-d_M+1}^{-d_m} \sum_{i=k+j}^{k-1} \boldsymbol{x}^{\mathrm{T}}(i)\boldsymbol{P}_d\boldsymbol{x}(i) = \boldsymbol{\zeta}_2^T(k-1)\boldsymbol{\Lambda}_2\boldsymbol{\zeta}_2(k-1) \tag{24}$$

where $\boldsymbol{\zeta}_2(k-1) = \left[\boldsymbol{x}^{\mathrm{T}}(k-1), \cdots, \boldsymbol{x}^{\mathrm{T}}\left(k-d_m\right), \boldsymbol{x}^{\mathrm{T}}\left(k-d_m-1\right), \cdots, \boldsymbol{x}^{\mathrm{T}}\left(k-d_M+1\right)\right]^{\mathrm{T}}$, $\boldsymbol{\Lambda}_2 = \mathrm{diag}\left(d_s\boldsymbol{P}_d, \left(d_s-1\right)\boldsymbol{P}_d, \cdots, \boldsymbol{P}_d\right)$.

Summing (16), (23), (24), we can conclude that

$$V(\boldsymbol{x}(k)) = \boldsymbol{\zeta}_3^{\mathrm{T}}(k)\boldsymbol{\Lambda}_3\boldsymbol{\zeta}_3(k) \tag{25}$$

Substituting $\boldsymbol{X} = \xi(k)\boldsymbol{P}^{-1}$, $\boldsymbol{X}_d = \xi(k)\boldsymbol{P}_d^{-1}$, then by using Schur complement, (22) can be obtained.

3.2. Robust Positively Invariant Set Computation

In this section, we design a LRF to compute RPI set. The RPI set is defined as follows:

$$\boldsymbol{\Omega} = \left\{\{\boldsymbol{x}, \boldsymbol{x}_d\} \mid \max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}, \boldsymbol{x}_d^{\mathrm{T}}\boldsymbol{P}_d\boldsymbol{x}_d\right\} \leq \xi\right\} \tag{26}$$

Lemma 3: Consider system (12), the set $\boldsymbol{\Omega}$ is a RPI if there exists a positive scalar $\gamma \in (0,1)$ such that:

$$\frac{1}{\xi}\boldsymbol{x}^{\mathrm{T}}(k+1)\boldsymbol{P}\boldsymbol{x}(k+1) - \frac{1-\lambda}{\xi}\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}, \boldsymbol{x}_d^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}_d\right\} - \frac{\lambda}{\rho^2}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{p} \leq 0 \tag{27}$$

Proof:

According to (27), $\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}, \boldsymbol{x}_d^{\mathrm{T}}\boldsymbol{P}_d\boldsymbol{x}_d\right\} \le \xi$ and $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P} \le \rho^2$, we have:

$$\begin{aligned}\frac{1}{\xi}\boldsymbol{x}^{\mathrm{T}}(k+1)\boldsymbol{P}\boldsymbol{x}(k+1) &\le \frac{1-\lambda}{\xi}\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}, \boldsymbol{x}_d^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}_d\right\} + \frac{\lambda}{\rho^2}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{p} \\ &\le (1-\lambda)+\lambda = 1\end{aligned} \tag{28}$$

So we can get $\boldsymbol{x}^{\mathrm{T}}(k+1)\boldsymbol{P}\boldsymbol{x}(k+1) \le \xi$, that means $\boldsymbol{x}(k+1)$ is still in the RPI set.

Theorem 3: If there exists a positive definite matrix $\boldsymbol{X}$, a general matrix $\boldsymbol{Y}$ and positive scalars σ and λ, such that the following BMIs and LMIs hold.

$$\begin{bmatrix} \gamma(\lambda-1)\boldsymbol{X} & * & * & * & * & * \\ 0 & \gamma_d(\lambda-1)\boldsymbol{X} & * & * & * & * \\ 0 & 0 & -\frac{\lambda}{\rho^2}\boldsymbol{I} & * & * & * \\ \boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Y} & \boldsymbol{A}_d\boldsymbol{X} & \boldsymbol{E} & -\boldsymbol{X} & * & * \\ \boldsymbol{N}_A\boldsymbol{X}+\boldsymbol{N}_B\boldsymbol{Y} & \boldsymbol{N}_{Ad}\boldsymbol{X} & 0 & 0 & -\sigma\boldsymbol{I} & * \\ \sigma\boldsymbol{M}^{\mathrm{T}} & 0 & 0 & 0 & 0 & -\sigma\boldsymbol{I} \end{bmatrix} \le 0 \tag{29}$$

$$\begin{bmatrix} \boldsymbol{Z} & * \\ \boldsymbol{Y}^{\mathrm{T}} & \boldsymbol{X} \end{bmatrix} \ge 0 \tag{30}$$

where $\gamma+\gamma_d = 1$, $Z_{tt} \le \delta^2_{t,max}$, Z_{tt} is the t-th diagonal element of matrix $\boldsymbol{Z}$.

Proof: See **Appendix B**.

3.3. Online Robust MPC Algorithm

We summarize the robust MPC algorithm based on the above-mentioned results. The implementation of the robust controller is summarized as follows.

$$\begin{gathered}\min_{\xi,\boldsymbol{X},\boldsymbol{X}_d,\boldsymbol{Y},\lambda,\eta,\sigma} \xi \\ \text{s.t.}(21),(22),(29),(30)\end{gathered} \tag{31}$$

Then the control input at time k can be computed $\boldsymbol{\delta}(k) = \boldsymbol{K}\boldsymbol{x}(k)$. Note that (29) is BMIs, in our paper, to reduce the computation burden, we only calculate the optimal λ^* of λ at the initial time, and replace λ with λ^* in subsequent iterations.

Theorem 4: If the optimization problem (31) is solvable for vehicle system (12) at the initial time, then the vehicle system (12) is ISS with regard to the external disturbances.

Proof: See **Appendix C**.

Remark: In this paper, two Lyapunov functions (LKF and LRF) are used at the same time. But LRF is only used to ensure that the system enters RPI, which provides a tighter constraint. As shown in Gielen et al. (2012a), the existence of a LRF is a sufficient condition for the existence of a LKF.

4. SIMULATION AND ANALYSIS

In this section, we conduct three simulation scenarios to validate the effectiveness of the designed robust MPC controller. The vehicle model parameters m, I_z, l_f, l_r, $C_{\alpha f}$, $C_{\alpha r}$, I_z and v_x of the simulation vehicle are $1,000$kg, $1,650$kg/m^2, 1.0m, 1.6m, $3,000$N/rad, $3,000$N/rad and 10m/s, respectively.

In the first simulation, we assume there is no time-varying delay, 2% model uncertainties are added, i.e., $\boldsymbol{N}_A = 0.02\boldsymbol{A}$, $\boldsymbol{N}_B = 0.02\boldsymbol{B}$. $\boldsymbol{M}$ is an identity matrix, $\boldsymbol{E} = \begin{bmatrix} 0.01 & 0.1 \end{bmatrix}^{\mathrm{T}}$, $\boldsymbol{Q}$ is diag$\{5,5\}$, $\boldsymbol{R}$ is 1, τ is 1. The external disturbances $\boldsymbol{p}$ and are given by $\boldsymbol{p}(k) = 1 \times 10^{-7}\sin(k), \boldsymbol{H}(k) = \sin(k)$ respectively (At each time step, the magnitude of the increment of β and r are 1×10^{-7} and 1×10^{-6} respectively). The input constraint is $\|\boldsymbol{\delta}\| \le 0.5$ rad. The sampling time is 0.01 s. Under the same conditions, we compared the performance of the proposed robust MPC with the traditional MPC algorithm (Elbanhawi et al., 2018) on the reference signal tracking problem. **Figure 2A** is the sideslip angle tracking performance comparison between the proposed robust MPC and traditional MPC. **Figure 2B** is the yaw rate tracking performance comparison between the proposed robust MPC and traditional MPC. It can be easily seen that the proposed robust MPC has a good tracking performance even there exist model uncertainties and external disturbances. However, traditional MPC controller becomes unstable and can not execute the control. Please note that the performance of MPC control is seriously degraded in the presence of external disturbances and model uncertainties. To observe the performance comparison of the MPC controller and the proposed robust MPC controller in the same figure more clearly, we only show the MPC controller tracking trajectory for 1s in the figure. Otherwise, the tracking trajectory generated by the robust MPC controller and the desired tracking trajectory will look like a straight line because of the scale.

In the second simulation, we assume the vehicle system has time-varying delay, the retarded coefficient α is 0.8, d_m is 1, d_M is 3, γ is 0.8, γ_d is 0.2. 5% model uncertainties are added, i.e., $\boldsymbol{N}_A = 0.05\boldsymbol{A}$, $\boldsymbol{N}_B = 0.05\boldsymbol{B}$, $\boldsymbol{N}_{Ad} = 0.05\boldsymbol{A}_d$. $\boldsymbol{M}$ is an identity matrix, $\boldsymbol{E} = \begin{bmatrix} 0.01 & 0.1 \end{bmatrix}^{\mathrm{T}}$, $\boldsymbol{Q}$ is diag$\{5,5\}$, $\boldsymbol{R}$ is 1, τ is 1. The external disturbances $\boldsymbol{p}$ and $\boldsymbol{H}$ are given by $\boldsymbol{p}(k) = 1 \times 10^{-3}\sin(k), \boldsymbol{H}(k) = \sin(k)$ respectively. The input constraint is $\|\boldsymbol{\delta}\| \le 0.5$ rad. The sampling time is 0.01 s. **Figure 3** shows the proposed controller can track the reference well when there exist model uncertainties, external disturbances, and time-varying delay. The root mean square error (RMSE) of the sideslip angle tracking error is 7.8719×10^{-6}, the RMSE of the yaw rate tracking error is 7.996×10^{-5}.

In order to more convincingly verify the effectiveness of the proposed controller, we chose square wave reference signals as the tracking signal in the third simulation. The upper bound of delay is set as $d_M = 10$, the external disturbances $\boldsymbol{p}$ is set as $\boldsymbol{p}(k) = 5 \times 10^{-3}\sin(k)$. Other parameters are the same as in simulation 2. As shown in **Figure 4**, despite the time-varying delay, model uncertainties and external disturbances, the proposed controller can still track the challenging reference signal well. The system states come into a very small region near the neighborhood of the reference tracking trajectory in the presence of a bigger external disturbances. The RMSE of the sideslip angle tracking error is 6.3077×10^{-5}, the RMSE of the yaw rate tracking error is 4.8308×10^{-4}.

To further verify the robustness of the proposed controller, we consider 10% model uncertainties and the external disturbances $\boldsymbol{p}$ are set as $\boldsymbol{p} = 5 \times 10^{-3}\sin(0.1\pi k)$, where the frequency of the disturbance refers to Wang N. et al. (2016). Other parameters are

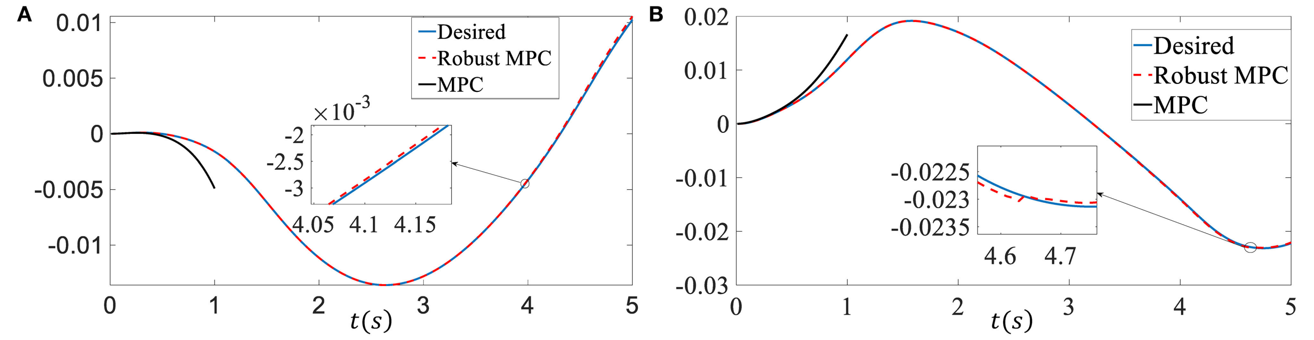


FIGURE 2 | Control performance comparison of the first simulation. **(A)** Sideslip angle. **(B)** Yaw rate.

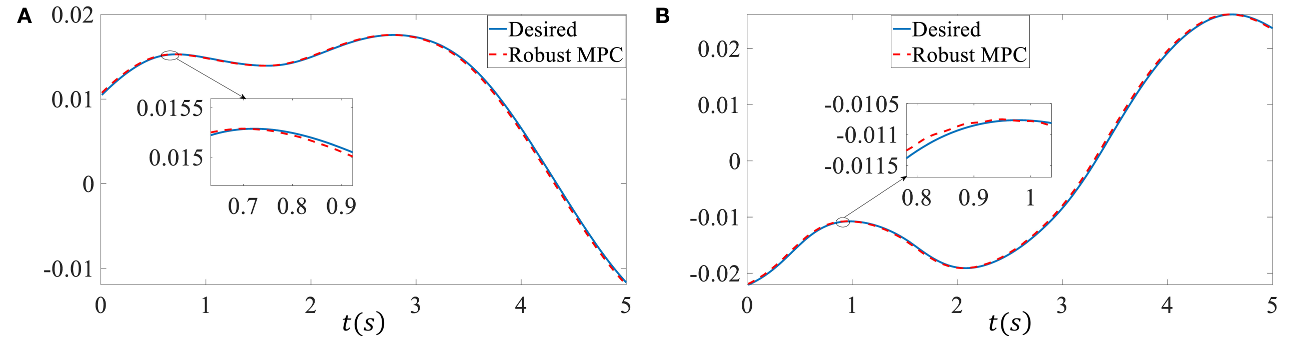


FIGURE 3 | Control performance of the second simulation. **(A)** Sideslip angle. **(B)** Yaw rate.

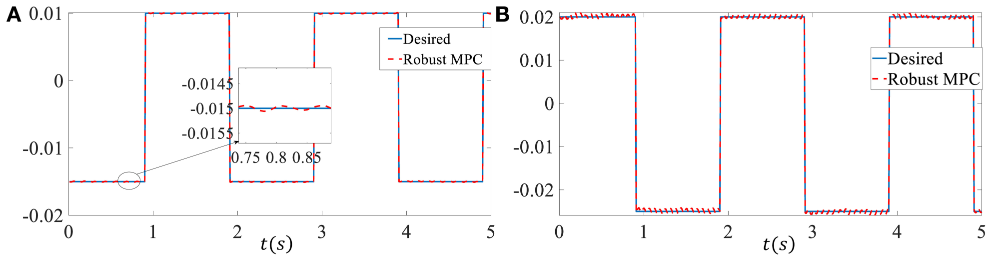


FIGURE 4 | Control performance of the third simulation. **(A)** Sideslip angle. **(B)** Yaw rate.

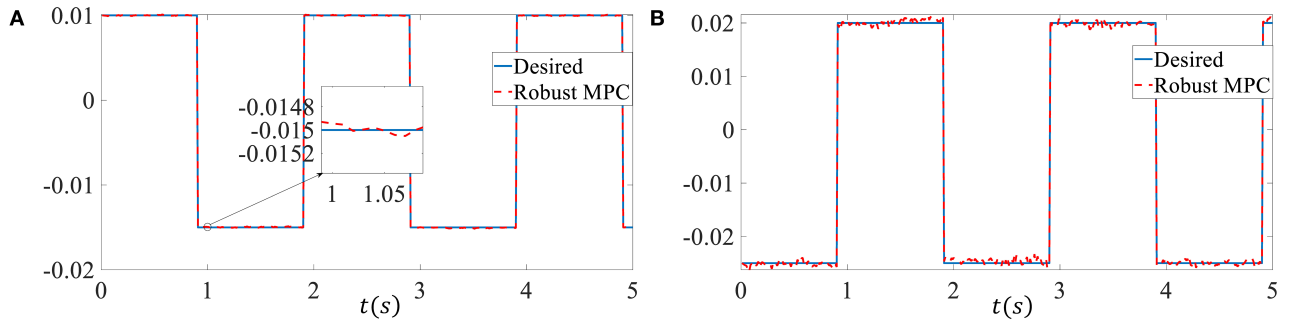


FIGURE 5 | Control performance of the fourth simulation. **(A)** Sideslip angle. **(B)** Yaw rate.

the same as in simulation 3. The square wave reference signals are also selected as the tracking signal in the this simulation. The performance of the proposed robust MPC controller is shown in **Figure 5**. The yaw rate tracking performance is a bit worse but it is still stable in the neighborhood of the reference tracking trajectory in the presence of bigger model uncertainties and bigger disturbance frequency.

5. CONCLUSION

In this paper, we design a robust MPC controller for vehicle system with model uncertainties, external disturbances, and time-varying delay. A LKF candidate related to time-varying delay is constructed to derive an upper bound of the cost function. A LRF is designed to compute the RPI. The computation of RPI and upper bound of the cost function, input constraint and the minimization problem are expressed as LMIs or BMIs by using matrix technology. Then the MPC state feedback law is calculated in terms of LMIs/BMIs. Simulation results show that the proposed controller is effective.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

WL: methodology, software, data curation, and writing original draft. GC: methodology and data curation. AK: conceptualization, supervision, and writing—review. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program. This research has also received funding from the European Union's Horizon 2020 Framework Program for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3), from the National Natural Science Foundation of China (No. 61906138), and from the Shanghai AI Innovative Development Project 2018.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbot.2020.617293/full#supplementary-material

REFERENCES

Bououden, S., Chadli, M., Zhang, L., and Yang, T. (2016). Constrained model predictive control for time-varying delay systems: application to an active car suspension. *Int. J. Control Automat. Syst.* 14, 51–58. doi: 10.1007/s12555-015-2009-4

Chen, G., Cao, H., Conradt, J., Tang, H., Röhrbein, F., and Knoll, A. (2020). Event-based neuromorphic vision for autonomous driving: a paradigm shift for bio-inspired visual sensing and perception. *IEEE Signal Process. Magaz.* 37, 34–49. doi: 10.1109/MSP.2020.2985815

Cheng, S., Li, L., Liu, C.-Z., Wu, X., Fang, S.-N., and Yong, J.-W. (2020). Robust lmi-based h-infinite controller integrating afs and dyc of autonomous vehicles with parametric uncertainties. *IEEE Trans. Syst. Man Cybernet. Syst.* doi: 10.1109/TSMC.2020.2964282

Cho, J., and Huh, K. (2019). Active front steering for driver steering comfort and vehicle driving stability. *Int. J. Automot. Technol.* 20, 589–596. doi: 10.1007/s12239-019-0056-1

Duan, G.-R., and Yu, H.-H. (2013). *LMIs in Control Systems: Analysis, Design and Applications*. Boca Raton, FL: CRC Press. doi: 10.1201/b15060

Elbanhawi, M., Simic, M., and Jazar, R. (2018). Receding horizon lateral vehicle control for pure pursuit path tracking. *J. Vibrat. Control* 24, 619–642. doi: 10.1177/1077546316646906

Gielen, R. H., Lazar, M., and Kolmanovsky, I. V. (2012a). Lyapunov methods for time-invariant delay difference inclusions. *SIAM J. Control Optimizat.* 50, 110–132. doi: 10.1137/100807065

Gielen, R. H., Lazar, M., and Teel, A. R. (2012b). Input-to-state stability analysis for interconnected difference equations with delay. *Math. Control Signals Syst.* 24, 33–54. doi: 10.1007/s00498-012-0080-4

Hu, C., Jing, H., Wang, R., Yan, F., and Chadli, M. (2016). Robust h infinity output-feedback control for path following of autonomous ground vehicles. *Mech. Syst. Signal Process.* 70, 414–427. doi: 10.1016/j.ymssp.2015.09.017

Hu, Y., Chen, G., Ning, X., Dong, J., Liu, S., and Knoll, A. (2019). "Mobile robot learning from human demonstrations with nonlinear model predictive control," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Macau: IEEE), 5057–5062. doi: 10.1109/IROS40897.2019.8968277

Jeong, S. C., and Park, P. (2005). Constrained mpc algorithm for uncertain time-varying systems with state-delay. *IEEE Trans. Automat. Control* 50, 257–263. doi: 10.1109/TAC.2004.841920

Jiang, L., and Wu, Z. (2018). Sliding mode control for intelligent vehicle trajectory tracking based on reaching law. *Trans. Chin. Soc. Agric. Mach.* 49, 381–386. doi: 10.6041/j.issn.1000-1298.2018.03.048

Jiang, Z.-P., and Wang, Y. (2001). Input-to-state stability for discrete-time nonlinear systems. *Automatica* 37, 857–869. doi: 10.1016/S0005-1098(01)00028-0

Kothare, M. V., Balakrishnan, V., and Morari, M. (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica* 32, 1361–1379. doi: 10.1016/0005-1098(96)00063-5

Liu, C., Lee, C., Hansen, A., Hedrick, J. K., and Ding, J. (2014). "A computationally efficient predictive controller for lane keeping of semi-autonomous vehicles," in *Dynamic Systems and Control Conference* (San Antonio, TX: American Society of Mechanical Engineers). doi: 10.1115/DSCC2014-6098

Liu, W., and Li, Z. (2019). Comprehensive predictive control method for automated vehicles with delays. *IEEE Access* 7, 81923–81933. doi: 10.1109/ACCESS.2019.2923762

Liu, Y., Wang, X., Li, L., Cheng, S., and Chen, Z. (2019). A novel lane change decision-making model of autonomous vehicle based on support vector machine. *IEEE Access* 7, 26543–26550. doi: 10.1109/ACCESS.2019.2900416

Mayne, D. Q., Rawlings, J. B., Rao, C. V., and Scokaert, P. O. (2000). Constrained model predictive control: stability and optimality. *Automatica* 36, 789–814. doi: 10.1016/S0005-1098(99)00214-9

Nahidi, A., Khajepour, A., Kasaeizadeh, A., Chen, S.-K., and Litkouhi, B. (2019). A study on actuator delay compensation using predictive control technique with experimental verification. *Mechatronics* 57, 140–149. doi: 10.1016/j.mechatronics.2018.11.007

Rajamani, R. (2011). *Vehicle Dynamics and Control*. Luxemburg: Springer Science and Business Media. doi: 10.1007/978-1-4614-1433-9_2

Ren, H., Karimi, H. R., Lu, R., and Wu, Y. (2019). Synchronization of network systems via aperiodic sampled-data control with constant delay and application to unmanned ground vehicles. *IEEE Trans. Indus. Electron*. 67, 4980–4990. doi: 10.1109/TIE.2019.2928241

Sakhdari, B., and Azad, N. L. (2018). Adaptive tube-based nonlinear mpc for economic autonomous cruise control of plug-in hybrid electric vehicles. *IEEE Trans. Vehicular Technol*. 67, 11390–11401. doi: 10.1109/TVT.2018.2872654

Shamaghdari, S., Nikravesh, S., and Haeri, M. (2015). Integrated guidance and control of elastic flight vehicle based on robust mpc. *Int. J. Robust Nonlinear Control* 25, 2608–2630. doi: 10.1002/rnc.3215

Sun, C., Zhang, X., Zhou, Q., and Tian, Y. (2019). A model predictive controller with switched tracking error for autonomous vehicle path tracking. *IEEE Access* 7, 53103–53114. doi: 10.1109/ACCESS.2019.2912094

Teel, A. R. (1998). Connections between razumikhin-type theorems and the iss nonlinear small gain theorem. *IEEE Trans. Automatic Control* 43, 960–964. doi: 10.1109/9.701099

Teng, L. (2018). *Robustmodel predictive control of nonlinear and time-delay systems* (Ph.D. thesis). Department of Mechanical and Biomedical Engineering, City University of Hong Kong, Hong Kong, China

Teng, L., Wang, Y., Cai, W., and Li, H. (2017). Fuzzy model predictive control of discrete-time systems with time-varying delay and disturbances. *IEEE Trans. Fuzzy Syst*. 26, 1192–1206. doi: 10.1109/TFUZZ.2017.2717798

Wang, N., Lv, S., Er, M. J., and Chen, W.-H. (2016). Fast and accurate trajectory tracking control of an autonomous surface vehicle with unmodeled dynamics and disturbances. *IEEE Trans. Intell. Vehicles* 1, 230–243. doi: 10.1109/TIV.2017.2657379

Wang, P., Liu, Z., Liu, Q., and Chen, H. (2019). An MPC-based manoeuvre stability controller for full drive-by-wire vehicles. *Control Theory Technol*. 17, 357–366. doi: 10.1007/s11768-019-9119-0

Wang, X., Taghia, J., and Katupitiya, J. (2016). Robust model predictive control for path tracking of a tracked vehicle with a steerable trailer in the presence of slip. *IFAC Pap. Online* 49, 469–474. doi: 10.1016/j.ifacol.2016.10.085

Xu, S., Peng, H., and Tang, Y. (2020). Preview path tracking control with delay compensation for autonomous vehicles. *IEEE Trans. Intell. Transport. Syst*. doi: 10.1109/TITS.2020.2978417

Yang, W., and Feng, G. (2013). "Robust model predictive control of discrete-time uncertain linear systems with persistent disturbances," in *2013 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems* (Nanjing: IEEE), 58–63. doi: 10.1109/CYBER.2013.6705420

Yang, W., Feng, G., and Zhang, T. (2013). Robust model predictive control for discrete-time takagi-sugeno fuzzy systems with structured uncertainties and persistent disturbances. *IEEE Trans. Fuzzy Syst*. 22, 1213–1228. doi: 10.1109/TFUZZ.2013.2286635

Yu, J., Guo, X., Pei, X., Chen, Z., Zhu, M., and Gong, B. (2019). *Robust Model Predictive Control for Path Tracking of Autonomous Vehicle*. Detroit, MI: Technical report, SAE Technical Paper. doi: 10.4271/2019-01-0693

Yu, S.-M., Wu, S.-N., Zhao, Y.-B., and He, D.-F. (2018). Delayed feedback mpc algorithms of vehicle platoons subject to constraints on measurement range and driving behaviors. *Asian J. Control* 20, 2260–2270. doi: 10.1002/asjc.1761

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

ORIGINAL RESEARCH
published: 26 January 2021
doi: 10.3389/fnbot.2021.627157

Adaptive Locomotion Control of a Hexapod Robot via Bio-Inspired Learning

Wenjuan Ouyang [1], *Haozhen Chi* [1], *Jiangnan Pang* [1], *Wenyu Liang* [2] *and Qinyuan Ren* [1]*

[1] *College of Control Science and Engineering, Zhejiang University, Hangzhou, China,* [2] *Department of Electrical and Computing Engineering, National University of Singapore, Singapore, Singapore*

In this paper, an adaptive locomotion control approach for a hexapod robot is proposed. Inspired from biological neuro control systems, a 3D two-layer artificial center pattern generator (CPG) network is adopted to generate the locomotion of the robot. The first layer of the CPG is responsible for generating several basic locomotion patterns and the functional configuration of this layer is determined through kinematics analysis. The second layer of the CPG controls the limb behavior of the robot to adapt to environment change in a specific locomotion pattern. To enable the adaptability of the limb behavior controller, a reinforcement learning (RL)-based approach is employed to tune the CPG parameters. Owing to symmetrical structure of the robot, only two parameters need to be learned iteratively. Thus, the proposed approach can be used in practice. Finally, both simulations and experiments are conducted to verify the effectiveness of the proposed control approach.

Keywords: hexapod robot, two-layer CPG, reinforcement learning, adaptive control, bio-inspired

OPEN ACCESS

Edited by:
Zhan Li,
University of Electronic Science and Technology of China, China

Reviewed by:
Ning Tan,
Sun Yat-Sen University, China
Mingchuan Zhou,
Technical University of Munich, Germany

Correspondence:
Qinyuan Ren
latepat@gmail.com

Received: *08 November 2020*
Accepted: *04 January 2020*
Published: *26 January 2021*

Citation:
Ouyang W, Chi H, Pang J, Liang W and Ren Q (2021) Adaptive Locomotion Control of a Hexapod Robot via Bio-Inspired Learning. Front. Neurorobot. 15:627157. doi: 10.3389/fnbot.2021.627157

1. INTRODUCTION

In the past decades, a big step has been taken toward the study of legged robots, such as the study of biped robot (Kim et al., 2020), quadruped robot (Hyun et al., 2014), hexapod robot (Yu et al., 2016), octopod robot (Grzelczyk et al., 2018), and etc. Most of these legged robots have exhibited astonish maneuverabilities in a typically structured environment. Among these legged robots, the hexapod robots have been increasingly attracting attention from scientists and a lot of hexapod robotic prototypes have been successfully developed (Stelzer et al., 2012; Li et al., 2019; Sartoretti et al., 2019; Lele et al., 2020; Zhao and Revzen, 2020). Even though these hexapod robots in shape look very much like the arthropod that the scientists are animating, such as ants or spiders, the robots developed hitherto are still pretty away from real arthropods. One of the main challenges lies in the difficulty of controlling the multi-legs of the robots with coordination to a complex dynamic environment.

To control the locomotion of hexapod robots, from a perspective of cybernetics, two methods are generally adopted, namely kinematics-based and bio-inspired. The former models the locomotion patterns via kinematics analysis. As pointed from the study of Ramdya et al. (2017), three basic locomotion patterns of Drosophila melanogaster have been extracted through biological study, namely tripod locomotion, quadruped locomotion, and five legs support locomotion. Based on the analysis of these three basic locomotion patterns, a foot-force distribution model is established for a hexapod robot walking on an unstructured terrain (Zhang et al., 2014). The study of Zarrouk and Fearing (2015) investigates the kinematics of a hexapod robot using only one actuator and explores the turning issue of the robot. In the work of Sun et al. (2018), the inverse kinematics of an 18-degree-of-freedom (DoF) hexapod robot is calculated to control the dynamicly alternating tripod locomotion of the robot. Since it is hard to accurately model the kinematics of all the six-leg crawling modes, most obtained locomotion patterns from kinematics analysis are rough and trail-and-error strategy is usually necessary for tuning the rough patterns applied on the robots. The study from Delcomyn (1980) indicates that center pattern generators (CPGs), which are mainly located in the central nervous system of vertebrates or in relevant ganglia of invertebrates, are primarily responsible for generating coordinated, rhythmic locomotion patterns of animals in real time, such as crawling, flying, swimming, and running. Inspired by the characteristics of the stability and self-adaption of biological CPGs, artificial CPGs have been extensively studied, namely the bio-inspired approach, for locomotion generation of hexapod robots. The notable examples include the studies in Chung (2015), Zhong et al. (2018), Yu et al. (2020), and Bal (2021). Through these previous studies, it can be found out that the bio-inspired method can greatly simplify the locomotion control problem underlying coordination of multiple legs.

Although the bio-inspired method has been widely and fruitfully applied in locomotion control of many biomimetic robots, it still remains a challenge for modulating the CPG parameters to generate adaptive locomotion for hexapod robots. The CPG parameters in many studies are determined by experiences and some researchers adopt data-driven optimization methods, such as particle swarm optimization (PSO) method and reinforcement learning (RL), to tune the parameters. In the work of Juang et al. (2011), a symbiotic species-based PSO algorithm is proposed to automate the parameter design for evolving dynamic locomotion of a hexapod robot, but reducing the computing complexity of the PSO algorithm is still under research. In addition, the study of Kecskés et al. (2013) points out that PSO method easily suffers from the partial optimism and causes the loss of accuracy in a coordinate system. In locomotion control, there has been recent success in using RL to learn to walk for hexapod robots. In the work of Barfoot (2006), a cooperative Q-learning RL approach is utilized to experimentally learn locomotion for a 6-DoF hexapod robot, but this RL approach may be unable to deal with the hexapod robots that have higher DoF. The researchers in Sartoretti et al. (2019) employ A3C RL algorithm to learn hexapodal locomotion stochastically. Nevertheless, the performance of the learned controller proposed in the study is dependent on a large number of iterations. For the different terrains, the locomotion of a hexapod robot is controlled through training several artificial neural networks via RL method separately (Azayev and Zimmerman, 2020), but the training scenario is limited to the expert policies and thus the adaptivity of the controller may be inflexible for a dynamic environment.

In this paper, a bio-inspired learning approach is proposed for locomotion control of a hexapod robot with environment change. The proposed bio-inspired learning approach can be characterized by the structure of the learning mechanism. Biologists have proved the motor patterns of animals are controlled by neuro systems hierarchically (Fortuna et al., 2004) and functional configuration of CPGs can be regulated according to sensory feedback to produce different motor outputs (Hooper, 2000). Therefore, inspired from biological control systems, a two-layer CPG motion control system is firstly proposed in this paper to generate locomotion for the robot and then the parameters of the CPG tuning issue is explored to enhance the adaptability of the robot. In the proposed bioinspired control system, the outputs of the first layer of the CPG are

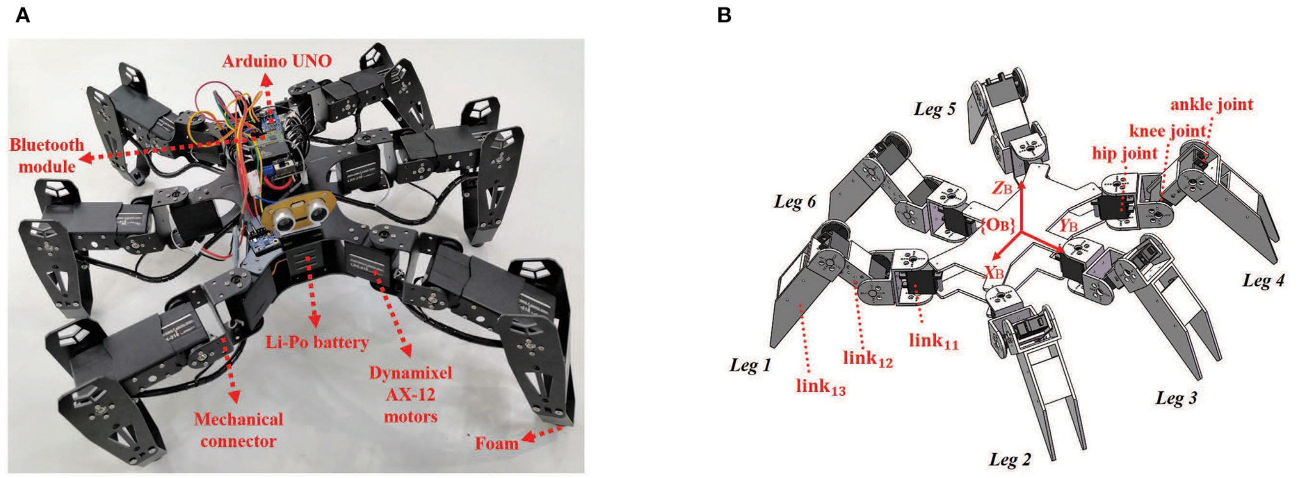


FIGURE 1 | (A) The hexapod robot prototype. **(B)** The mechanical schematic of the hexapod robot. $\{O_B\}$ is the body frame whose origin is fixed on the body's mass center. The robot possesses a rectangular body trunk and six mechanical legs numbered from leg 1 to leg 6. There are three joints and three links in each leg. The joints are named as hip joint, knee joint and ankle joint from the direction of body to foot tip. The legs can be labeled as $link_{i1}, link_{i2}, link_{i3}$ with the leg number $i = 1, 2, \cdots, 6$. For example, the $link_{11}$ (between hip joint and knee joint), $link_{12}$ (between knee joint and ankle joint) and $link_{13}$ (between ankle joint and foot tip) are the three links of leg 1.

TABLE 1 | Technical specifications of the prototype.

Parameter		Prototype	
		Value	Unit
Number of servo motor		18	\
Power supply		7.4	DC(V)
Total weight		1.995	kg
Body dimension	Length	24	cm
	Width	18.5	cm
	Height	4.5	cm
Limb $link_{i1}$	Weight	18.6	g
	Length	4.5	cm
Limb $link_{i2}$	Weight	128	g
	Length	7.5	cm
Limb $link_{i3}$	Weight	56.3	g
	Length	13.5	cm

Where $i = 1, 2, \cdots, 6$ *is the number of six legs.*

responsible for generating the basic locomotion patterns, such as tripod locomotion, quadruped locomotion, and five legs support locomotion. The second layer of the CPG acting as a Behavior Layer controls the limb motion of the hexapod robot. In order to adapt to environment change, through sensory feedback, basic locomotion patterns can be switched accordingly, and the limb behavior of the robot is regulated via a RL-based learning approach. Compared to the pure data-driven locomotion control approach, only few of the CPG parameters involved with the limb behavior control need to be learned iteratively. Hence, the proposed locomotion control approach can be adopted to the robot practically.

The rest of this paper is organized as follows. Section 2 introduces the model of the hexapod robot. Section 3 details the two-layer CPG controller and explores its dynamics with numerical studies. Following that, the RL-based learning approach for refining the CPG parameters is presented in section 4. In section 5, both simulations and experiments are conducted to verify the proposed locomotion control approach. Finally, the conclusions and future work are given.

2. MODELING OF A HEXAPOD ROBOT

2.1. The Prototype of the Hexapod Robot

The prototype of the hexapod robot is investigated in this paper shown in **Figure 1A**, and the specifications are given in **Table 1**.

Figure 1B illustrates the mechanical schematic of the hexapod robot, which consists of 18 servo motors, a microprocessor, a Bluetooth communication module, a set of mechanical connectors and several other peripherals. Three motors (Dynamixel AX-12) equipped in a leg are concatenated together to act as three joints. A microprocessor (Arduino UNO) is used for processing sensor data, transferring diagnostic information via the Bluetooth module, making decisions and controlling servo motors. Besides that, an external camera (Logitech C930) is employed to track the position of the robot as feedback signals.

2.2. Modeling

To establish the kinematic/dynamic model of the hexapod robot, the joint coordinates of each leg i are defined as depicted in **Figure 2**.

The kinematic model is represented by Denavit-Hartenberg (DH) parameters for resolving inverse kinematic of the leg. According to these fixed frames, the transformation parameters and DH parameters are demonstrated in **Tables 2**, **3**, respectively.

The relative translation and rotation between the $(j-1)th$ and the jth joint coordinates are computed by the transformation matrix (1):

$$ {}^{i}\mathbf{T}_{j}^{j-1} = \begin{bmatrix} \cos\theta_{ij} & -\cos\alpha_{ij}\sin\theta_{ij} & \sin\alpha_{ij}\sin\theta_{ij} & a_{ij}\cos\theta_{ij} \\ \sin\theta_{ij} & \cos\alpha_{ij}\cos\theta_{ij} & -\sin\alpha_{ij}\cos\theta_{ij} & a_{ij}\sin\theta_{ij} \\ 0 & \sin\alpha_{ij} & \cos\alpha_{ij} & d_{ij} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1} $$

where especially, the transition matrix between the body coordinate $\{O_B\}$ and the hip joint coordinate $\{O_{i0}\}$ is represented by (2):

$$ {}^{i}\mathbf{T}_{0}^{B} = \begin{bmatrix} \cos\varphi_i & -\sin\varphi_i & 0 & d_{xi} \\ \sin\varphi_i & \cos\varphi_i & 0 & d_{yi} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2} $$

Consequently, the foot tip coordinate $\{O_{i3}\}$ can be transformed into the body coordinate $\{O_B\}$ by multiplying the previous matrixs sequentially shown in (3):

$$ \begin{aligned} {}^{i}\mathbf{T}_{3}^{B} &= {}^{i}\mathbf{T}_{0}^{B} \cdot {}^{i}\mathbf{T}_{1}^{0} \cdot {}^{i}\mathbf{T}_{2}^{1} \cdot {}^{i}\mathbf{T}_{3}^{2} \\ &= \begin{bmatrix} \cos(\varphi_i+\theta_{i1})\cos(\theta_{i2}+\theta_{i3}) & -\cos(\varphi_i+\theta_{i1})\sin(\theta_{i2}+\theta_{i3}) & \sin(\varphi_i+\theta_{i1}) & d_{xi}+\cos(\varphi_i+\theta_{i1})(l_{i1}+l_{i2}\cos(\theta_{i2})+l_{i3}\cos(\theta_{i2}+\theta_{i3})) \\ \sin(\varphi_i+\theta_{i1})\cos(\theta_{i2}+\theta_{i3}) & -\sin(\varphi_i+\theta_{i1})\sin(\theta_{i2}+\theta_{i3}) & -\cos(\varphi_i+\theta_{i1}) & d_{yi}+\sin(\varphi_i+\theta_{i1})(l_{i1}+l_{i2}\cos(\theta_{i2})+l_{i3}\cos(\theta_{i2}+\theta_{i3}) \\ \sin(\theta_{i2}+\theta_{i3}) & \cos(\theta_{i2}+\theta_{i3}) & 0 & l_{i2}\sin(\theta_{i2})+l_{i3}\sin(\theta_{i2}+\theta_{i3}) \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{3} $$

Thus, the position of the foot tip with respect to the body coordinate $\{O_B\}$ can be derived as given below:

$$ \begin{aligned} \begin{bmatrix} p_{xi} \\ p_{yi} \\ p_{zi} \end{bmatrix} &= \begin{bmatrix} d_{xi}+\cos(\varphi_i+\theta_{i1})(l_{i1}+l_{i2}\cos(\theta_{i2})+l_{i3}\cos(\theta_{i2}+\theta_{i3}) \\ d_{yi}+\sin(\varphi_i+\theta_{i1})(l_{i1}+l_{i2}\cos(\theta_{i2})+l_{i3}\cos(\theta_{i2}+\theta_{i3}) \\ l_{i2}\sin(\theta_{i2})+l_{i3}\sin(\theta_{i2}+\theta_{i3}) \end{bmatrix}, \end{aligned} \tag{4} $$

where $[p_{xi}\ p_{yi}\ p_{zi}]^T$ is the position coordinate of the ith foot hip and θ_{ij} is the joint angle.

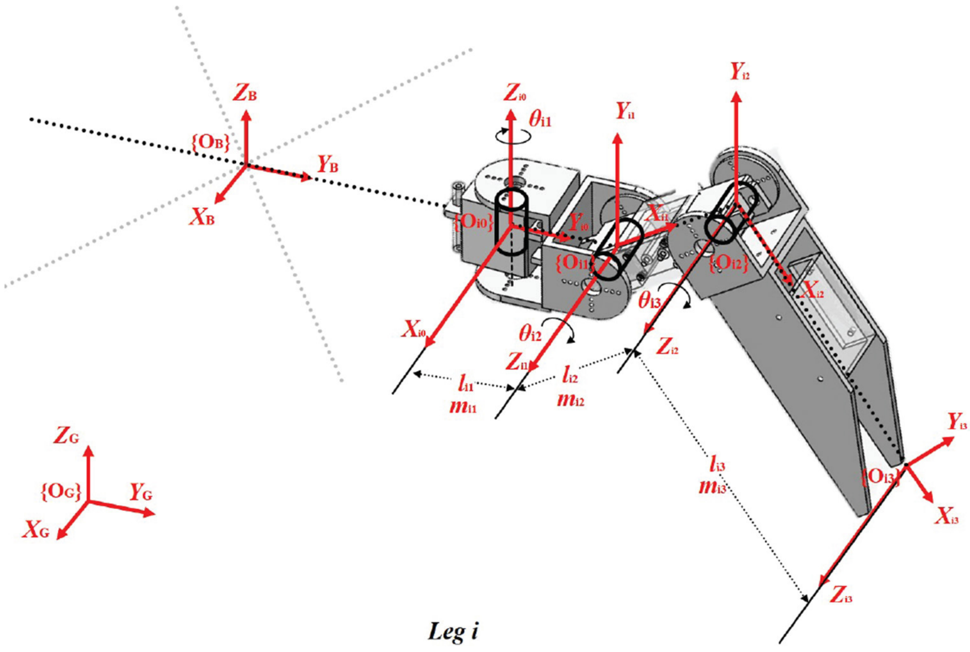


FIGURE 2 | Coordinates at different joints of each leg *i*. $\{O_G\}$ denotes the global coordinate, and $\{O_{ik}\}$ $(k = 0, 1, 2, 3)$ represents the floating frame whose origin fixed on the joints or the foot tip. l_{i1}, l_{i2}, l_{i3} and m_{i1}, m_{i2}, m_{i3} are the length and mass of the leg links, respectively. And $\theta_{i1}, \theta_{i2}, \theta_{i3}$ are the rotational joint angles around Z_{i0}, Z_{i1}, Z_{i2} axis of the leg.

TABLE 2 | Transformation parameters from the $\{O_{i0}\}$ to the $\{O_B\}$.

Leg *i*	**1**	**2**	**3**	**4**	**5**	**6**
d_{xi}(mm)	33.5	67	33.5	−33.5	−67	−33.5
d_{yi}(mm)	58	0	−58	−58	0	58
φ_i(°)	−60	0	60	120	180	−120

The hip joint position is defined as (d_{xi}, d_{yi}) *and* φ_i *denotes the direction angle in the body frame* $\{O_B\}$*.*

TABLE 3 | Denavit-Hartenberg parameters.

Joint *j*	α_{ij}	a_{ij}	d_{ij}	θ_{ij}
1	$\pi/2$	l_{i1}	0	θ_{i1}
2	0	l_{i2}	0	θ_{i2}
3	0	l_{i3}	0	θ_{i3}

Where $j = 1,2,3$ *is the joint number from hip joint to knee joint of each leg. And* α_{ij} *is the link twist indicating the angle from* $Z_{i(k-1)}$ *to* Z_{ik} *around* X_{ik}*,* a_{ij} *is the link length representing the distance from the intersection of* $Z_{i(k-1)}$ *and* X_{ik} *to the origin of* X_{ik}*,* d_{ij} *is the joint distance meaning the distance from the intersection of* $Z_{i(k-1)}$ *and* X_{ik} *to the origin of* $Z_{i(k-1)}$*,* θ_{ij} *is the joint angle showing the angle from* $X_{i(k-1)}$ *to* X_{ik} *around* $Z_{i(k-1)}$*.*

The leg of the hexapod robot is a complex joint-link system connecting the body trunk with the ground. Hence, closed kinematics chains can be found in the robot system. Since forces and moments propagate via the kinematics chains among different legs (Roy and Pratihar, 2013), the kinematics and dynamics are coupled. The dynamic model of such a coupled hexapod robot with 18 actuators is derived via Lagrangian-Euler method as follows:

$$\boldsymbol{\tau}_i = \mathbf{M}_i(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}}_i + \mathbf{H}_i(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})\dot{\boldsymbol{\theta}}_i + \mathbf{G}_i(\boldsymbol{\theta}) - \bar{\mathbf{J}}_i^T \mathbf{F}_i \quad , \tag{5}$$

where $\boldsymbol{\tau}_i = [\tau_{i1}\ \tau_{i2}\ \tau_{i3}]^T \in \mathbb{R}^3$ is the joint torque vector of the *ith* leg consisting of hip joint torque τ_{i1}, knee joint torque τ_{i2} and ankle joint torque τ_{i3}. $\boldsymbol{\theta}_i = [\theta_{i1}\ \theta_{i2}\ \theta_{i3}]^T \in \mathbb{R}^3, \dot{\boldsymbol{\theta}}_i = [\dot{\theta}_{i1}\ \dot{\theta}_{i2}\ \dot{\theta}_{i3}]^T \in \mathbb{R}^3, \ddot{\boldsymbol{\theta}}_i = [\ddot{\theta}_{i1}\ \ddot{\theta}_{i2}\ \ddot{\theta}_{i3}]^T \in \mathbb{R}^3$ are joint angle, joint angle acceleration, and joint angle jerk vector of the *ith* leg, respectively. $\mathbf{M}_i(\boldsymbol{\theta}) \in \mathbb{R}^{3\times3}$ is a inertia matrix of the *ith* leg. $\mathbf{H}_i(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) \in \mathbb{R}^{3\times3}$ is Coriolis forces matrix of the *ith* leg. $\mathbf{G}_i(\boldsymbol{\theta}) \in \mathbb{R}^3$ is a link gravitational forces vector of the *ith* leg. $\mathbf{F}_i = [f_{ix}\ f_{iy}\ f_{iz}]^T \in \mathbb{R}^3$ represents ground reaction forces of the *ith* support foot tip with the coordinate $\{O_{i3}\}$. $\bar{\mathbf{J}}_i \in \mathbb{R}^{3\times3}$ is the Jacobian matrix of the *ith* leg, computed by (6). Moreover, the position and velocity of the hexapod robot in this work are transformed to the global coordinate $\{O_G\}$.

$$\bar{\mathbf{J}}_i = \begin{bmatrix} -(l_{i1} + l_{i2}\cos(\theta_{i2}) + l_{i3}\cos(\theta_{i2}+\theta_{i3}))\sin(\theta_{i1}) & -(l_{i2}\sin(\theta_{i2}) + l_{i3}\sin(\theta_{i2}+\theta_{i3}))\cos(\theta_{i1}) & -l_{i3}\sin(\theta_{i2}+\theta_{i3})\cos(\theta_{i1}) \\ (l_{i1} + l_{i2}\cos(\theta_{i2}) + l_{i3}\cos(\theta_{i2}+\theta_{i3}))\cos(\theta_{i1}) & -(l_{i2}\sin(\theta_{i2}) + l_{i3}\sin(\theta_{i2}+\theta_{i3}))\sin(\theta_{i1}) & -l_{i3}\sin(\theta_{i2}+\theta_{i3})\sin(\theta_{i1}) \\ 0 & l_{i2}\cos(\theta_{i2}) + l_{i3}\cos(\theta_{i2}+\theta_{i3}) & l_{i3}\cos(\theta_{i2}+\theta_{i3}) \end{bmatrix}. \tag{6}$$

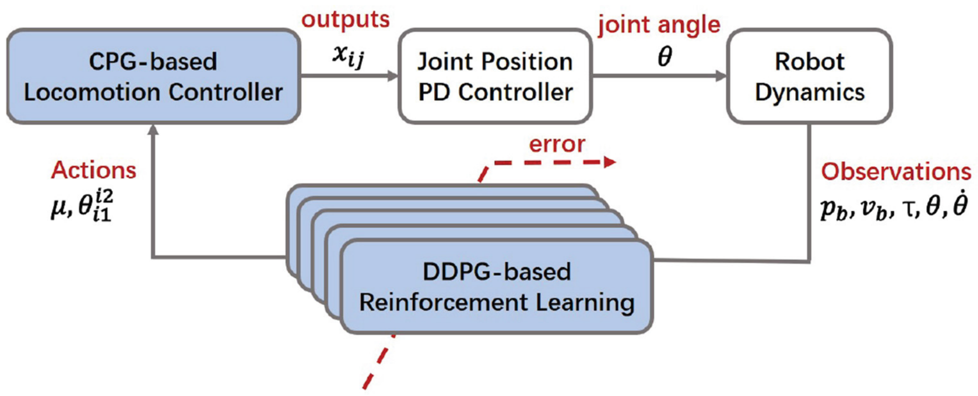


FIGURE 3 | Diagram of the proposed bio-inspired control scheme. The proposed control scheme has a cascaded structure with a feedback loop. It consists of three parts: (1) A dynamic model (with a embedded PD controller) that computes torque commands to handle robot dynamics subject to mechanical constraints. The dynamics parameters $p_b = [p_x, p_y, p_z]^T$, $v_b = [v_x, v_y, v_z]^T$ are the robot body position and velocity vector,respectively; τ, θ, $\dot{\theta}$ indicate the joint torque, angle and angle velocity, respectively. (2) A two-layer CPG locomotion controller that outputs coordinated signals to generate the basic locomotion. The CPG parameters μ and θ_{i1}^{i2} are the inputs representing the amplitude and the phase difference between the hip joint *i*1 and the knee joint *i*2 of the leg i, respectively; x_{ij} is the output signal. (3) A DDPG-based RL motion controller that trains the optimal locomotion via the cost function.

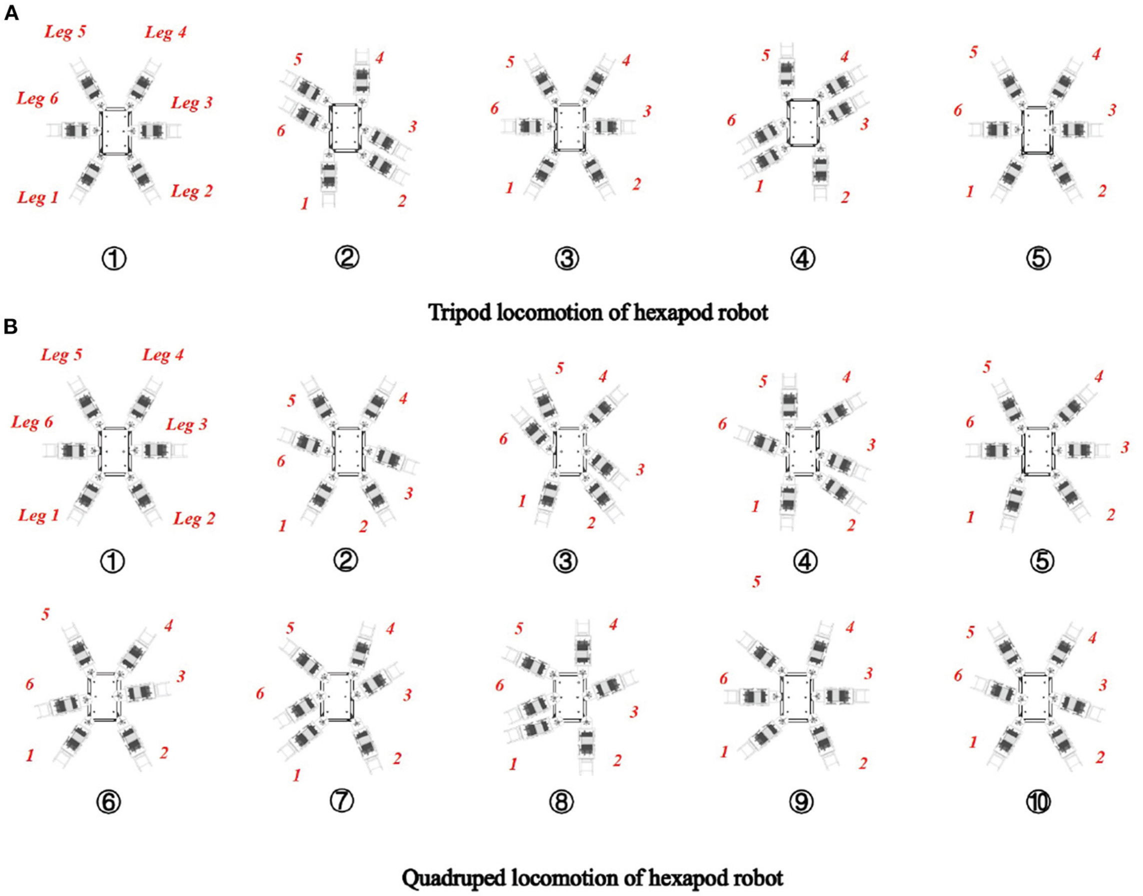


FIGURE 4 | Locomotion patterns via the body layer. During each cycle, six legs in the tripod locomotion are separated into two sets of {*leg* 1, *leg* 3, *leg* 5} and {*leg* 2, *leg* 4, *leg* 6} moving alternately, while in the quadruped locomotion are separated into four sets of {*leg* 3}, {*leg* 1, *leg* 4}, {*leg* 6}, and {*leg* 2, *leg* 5} to move successively.

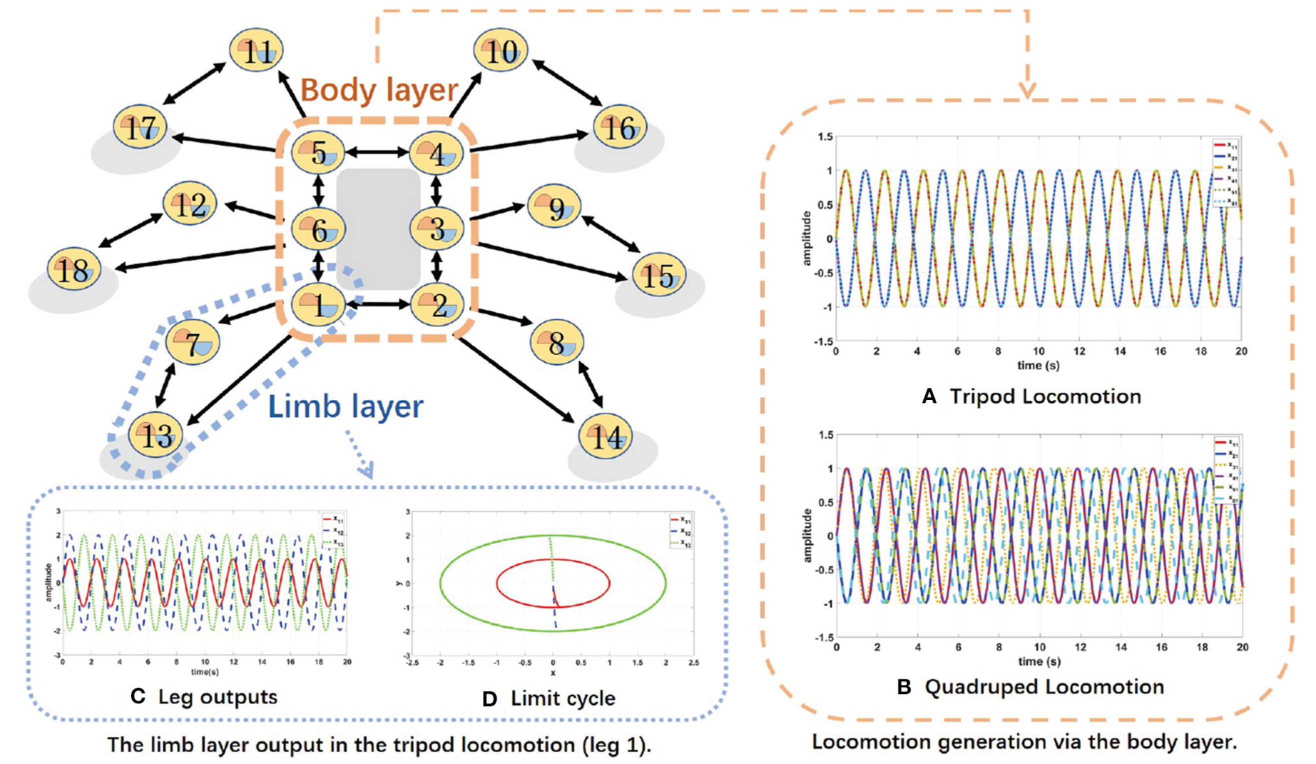


FIGURE 5 | The topology network of the proposed two-layer CPG.

3. LOCOMOTION CONTROLLER VIA CPG

Based on the analysis of the aforementioned mathematical model, the whole control scheme is proposed as shown in **Figure 3**. Inspired by biological arthropods, a hexapod robot is supposed to exhibit various locomotion in different terrains, such as tripod locomotion, quadruped locomotion, and five legs support locomotion (Zhong et al., 2018). Among these locomotion patterns, the tripod locomotion can achieve the fastest movement, while the quadruped and five legs support locomotion are more flexible. In this work, the locomotion patterns can be judged by velocity criterion according to the change of terrains.

In nature, CPGs are mainly used for generating coordinated and rhythmic movements for the locomotion of animals. Based on the similarity between biological legged animals and hexapod robots as well as the attractive capability of the CPG-based model on coupling the dynamics of robots, artificial CPG-based locomotion controllers are widely adopted to generate the locomotion behaviors of the biological counterparts. The basic locomotion patterns of the hexapod robot and the phase relations of the locomotion patterns are illustrated in **Figures 4A,B**.

3.1. Two-Layer CPGs Model

Due to complicated couplings and high degrees of freedom on the hexapod robot, the proposed CPG-based locomotion control is decomposed into two layers: (1) The body layer consists of six hip oscillators with bidirectional couplings. (2) The limb layer includes three oscillators in association with the hip joint, the knee joint and the ankle joint in every leg, where the knee joint oscillator and ankle joint oscillator are interconnected with bidirectional coupling, but the oscillator pair is unidirectionally controlled by the corresponding hip oscillator in the body layer.

Therefore, the body layer acting as a Conscious Layer shown in **Figures 5A,B** provides knowledge to determine the locomotion mode of the hexapod robot, while the limb layer acting as a Behavior Layer shown in **Figures 5C,D** has a major impact on final motion states and performance.

Considering the stable limit cycle and the explicit interpretable parameters, Hopf oscillator is a suitable element to construct CPGs for robotic locomotion (Seo et al., 2010). Hence, in this work, our CPG model can be described as a set of coupled Hopf oscillators and each Hopf oscillator is formulated by (7):

$$\begin{cases} \dot{x} = \alpha(\mu^2 - x^2 - y^2)x - \omega y \\ \dot{y} = \beta(\mu^2 - x^2 - y^2)y - \omega x \end{cases}, \tag{7}$$

where x and y are two state variables, ω is the frequency, α and β are the positive constants which determine the convergence rate of the limit cycle. In this paper, x is defined as the output signal of the oscillator.

Since the hexapod robot in this work has six legs and each leg has 3 DoF, a network consisted of 18 Hopf oscillators is proposed. According to the proposed CPG model shown in **Figure 5**, to achieve desired motion of the hexapod robot, multiple oscillators are needed to be coupled together to guarantee robotic system synchronization and coordination. Motivated by the work presented by Campos et al. (2010), the proposed CPG model connected by the diffusive coupling is described by:

$$\begin{aligned}\begin{bmatrix}\dot{x}_{ij}\\ \dot{y}_{ij}\end{bmatrix} &= \begin{bmatrix}\alpha(\mu^2 - x_{ij}^2 - y_{ij}^2) & -\omega_{ij}\\ \omega_{ij} & \beta(\mu^2 - x_{ij}^2 - y_{ij}^2)y_{ij}\end{bmatrix}\begin{bmatrix}x_{ij}\\ y_{ij}\end{bmatrix}\\ &\quad + k \cdot \sum_{mn \neq ij} \overline{R}(\theta_{mn}^{ij})\begin{bmatrix}0\\ \frac{x_{mn}+y_{mn}}{\sqrt{x_{mn}^2+y_{mn}^2}}\end{bmatrix}\end{aligned}, \quad (8)$$

where x_{ij}, y_{ij} with $i = 1, 2, \cdots, 6$ and $j = 1, 2, 3$ denote two state variables. The constant coupling strength $k = 0.1$ and the convergence coefficients $\alpha = \beta = 100$ are set for all oscillators, which are determined through a trial-and-error simulation on the stability of limit cycle in this work. The oscillator frequencies are unified as $\omega_{ij} = \omega$ for simplifying the high-level optimization. Besides, θ_{mn}^{ij} with $m = 1, 2, \cdots, 6$ and $n = 1, 2, 3$ represents the phase difference between the joint ij and the joint mn, then an associated 2D rotation matrix $\overline{R}(\theta_{mn}^{ij})$ is defined as:

$$\overline{R}(\theta_{mn}^{ij}) = \begin{bmatrix}\cos(\theta_{mn}^{ij}) & -\sin(\theta_{mn}^{ij})\\ \sin(\theta_{mn}^{ij}) & \cos(\theta_{mn}^{ij})\end{bmatrix}. \quad (9)$$

Compared with (7), the coupling relations among different Hopf oscillators are embedded into the artificial CPG model. This proposed 3D two-layer CPG model not only can regulate the basic locomotion patterns of the hexapod robot, but also fine-tune the motion performance for adapting to environment change. More information about the superiority of the 3D topology are demonstrated in our previous work (Niu et al., 2014). Through this CPG-based locomotion controller, the coordination can be adjusted with fewer parameters, which effectively reduce the control dimension and complexity of the system.

3.2. Simulation of Locomotion Generation

To verify the performance of the proposed CPG-based locomotion controller, several simulations are conducted. In the first layer of the network, the phase differences of the body layer among different hip joints are set as shown in **Table 4** to generate the tripod locomotion or quadruped locomotion. The six body oscillator parameters in the tripod and quadruped locomotion are set as *amplitude* = 1 and *frequency* = 3.14.

As can be seen from **Figures 5A,B**, the outputs of the body layer network are stable and periodic, while the phase differences between the neighboring oscillators maintain strictly 180 *deg* for tripod locomotion and 90 *deg* for quadruped locomotion, respectively.

Take the tripod locomotion patterns in leg 1 as an example, the limb layer network firstly receives the corresponding hip joint signal from the body layer. Secondly, the limb network outputs two signals to control the knee joint and the ankle joint interacting with environment. The phase difference between the knee joint and the ankle joint is limited to 180 *deg* in each leg with *amplitude* = 2 and *frequency* = 3.14.

As shown in **Figure 5C**, the phase difference between the knee joint and the ankle joint is locked in the limb layer. Moreover, **Figure 5D** presents the stable limit cycle of the coupled Hopf oscillators, which alleviates the influence of disturbances and ensures the smooth tuning of the robot locomotion. These simulation results show that the proposed CPG-based locomotion controller carry the potential of excellent controllability and robustness in unknown and unstructured terrains via online adjustment.

It should be noted that several parameters play important roles in the two-layer CPG controller, namely, amplitudes, frequencies, and phase differences. The CPG allows direct modulation of these parameters to enhance locomotion adaptability of the hexapod robot, but the manual tuning process still remain a challenge. Motivated by the movement of the six-legged arthropods modulated further via higher controller from brain-stem level (Yu et al., 2020), a RL-based controller is proposed to optimize the specialized locomotion patterns automatically in the next section.

4. LOCOMOTION OPTIMIZATION VIA REINFORCEMENT LEARNING

4.1. Problem Statement

Locomotion of a hexapod robot can be considered as a Markov Decision Process (MDP), which is described as an agent interacting with the environment in discrete time steps. At each time step t, the state of the agent and the environment can be jointly described by a state vector $s \in S$, where S is the state space. The agent takes an action $a_t \in A$, after which the environment advances a single time step and reaches a new state s_{t+1} through an environment state-transition distribution $\mathscr{P} : S \times A \times S \rightarrow [0, 1]$. Each state-action transition process is evaluated by a scalar reward function $\mathscr{R} : S \times A \rightarrow \mathbb{R}$. At the beginning of each episode, the agent finds itself in a random initial state $s_0 \sim \rho(s)$. Thus, the MDP is defined as a tuple $(S, A, \mathscr{R}, \mathscr{P}, \rho)$ (Tan et al., 2018).

In the MDP, the agent selects the action a under the state s through a stationary policy $\pi : S \rightarrow P(A)$, which is defined as a function mapping states into probability distributions over actions. The set of all stationary policies is denoted as Π. Give a performance measurement function as:

$$J(\pi) = \underset{\varsigma \sim \pi}{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t, s_{t+1})\right], \quad (10)$$

where $\gamma \in [0, 1)$ is the discount factor and ς denotes a trajectory under the policy π. The objective of the RL is to select a optimal policy π^* that maximizes $J(\pi)$, i.e.,

$$\pi^* = \underset{\pi \in \Pi}{argmax}\ J(\pi). \quad (11)$$

However, lack of complete freedoms when training the hexapod robot could suffer from some failed actions, such as collisions, falls, and inaccessible locomotion for a real robot. To tackle these issues, the actions of the hexapod robot should be constrained by several conditions such as acceleration, velocity, and torque constraints, which ensures the robot safe exploration.

TABLE 4 | Phase differences in corresponding locomotion.

Phase differences	**Locomotion patterns**			
	Tripod *(deg)*		**Quadruped *(deg)***	
θ_{11}	0	$\theta_{mn}^{ij} = 360 \cdot \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 \end{bmatrix}$	0	$\theta_{mn}^{ij} = 360 \cdot \begin{bmatrix} 0 & \frac{2}{4} & \frac{1}{4} & 0 & \frac{2}{4} & \frac{3}{4} \\ -\frac{2}{4} & 0 & -\frac{1}{4} & -\frac{2}{4} & 0 & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{2}{4} \\ 0 & \frac{2}{4} & \frac{1}{4} & 0 & \frac{2}{4} & \frac{3}{4} \\ -\frac{2}{4} & 0 & -\frac{1}{4} & -\frac{2}{4} & 0 & \frac{1}{4} \\ -\frac{3}{4} & -\frac{1}{4} & -\frac{2}{4} & -\frac{3}{4} & -\frac{1}{4} & 0 \end{bmatrix}$
θ_{21}	180		180	
θ_{31}	0		90	
θ_{41}	180		0	
θ_{51}	0		180	
θ_{61}	180		270	

Similar to the MDP, a constrained Markov Decision Process (CMDP) is defined as a tuple $(S, A, \mathscr{R}, C, \mathscr{P}, d, \rho)$. The difference between the MDP and the CMDP is that the policies are trained under additional cost constrains C in the latter. Each cost function $C_l : S \times A \times S \rightarrow \mathbb{R}$ maps transition tuples into costs with the limit c_l. Thus, the discounted cost of policy π with cost function C_l (Achiam et al., 2017) is represented by:

$$J_{C_l}(\pi) = \underset{\varsigma \sim \pi}{E}[\sum_{t=0}^{\infty} \gamma^t C_l(s_t, a_t, s_{t+1})]. \tag{12}$$

where l is the number of the constraints.

The set of feasible stationary policies in a CMDP is

$$\Pi_C = \{\pi \in \Pi : \forall l,\ J_{C_l}(\pi) <= c_l\}, \tag{13}$$

and the reinforcement learning problem in a CMDP is formulated as:

$$\pi^* = \underset{\pi \in \Pi_C}{argmax}\ J(\pi). \tag{14}$$

4.2. Deep Deterministic Policy Gradient Algorithm

Hexapod robots are multiple-input-multiple-output (MIMO) systems, so generally both the state space and the action space are high-dimensional and continuous. While many of stochastic policy gradient-based RL methods require massive and time-consuming search in such a vast space, deterministic policy greatly improve learning rates without sampling in the action space. Deep deterministic policy gradient (DDPG) (Lillicrap et al., 2016) as a model-free, off-policy RL algorithm, which could deal with unprocessed, high-dimensional sensory inputs and learn policies in a high-dimensional continuous action space via deep function approximators, has been widely accepted for robot control. Adaptive locomotion control of a hexapod robot is a challenging task due to the high-dimensional observations and continuous actions. In this work, the DDPG-based reinforcement learning optimization approach is proposed and illustrated in **Figure 6**.

Significantly, the DDPG combines an actor-critic method with deep neural networks (DNNs), and it shows stable performance in tough physical control problems including complex multi-joint movements and unstable contact dynamics. Besides, compared with the on-policy and stochastic counterparts such as proximal policy optimization (PPO), the off-policy and deterministic feature of DDPG ensures a more sample-efficient learning owing to the ability of generating a deterministic action.

The proposed DDPG algorithm is applied on learning the adaptive control policy π for the hexapod robot. The control policy π is assumed to be parameterized by θ^π. Specifically, the RL problem of learning the optimal control policy is converted into learning the optimal value θ^π. Considering that Policy Gradient is utilized for continuous action space, DDPG algorithm actually combines Policy Gradient with an actor-critic network. The parameter vector θ^π is updated in the direction that maximizes the performance measurement function $J(\pi)$. The direction, defined as the action gradient, is the gradient of $J(\pi)$ with respect to θ^π which can be calculated as follows:

$$\nabla_{\theta^\pi} J(\pi) = E[\nabla_a Q(s, a) \nabla_{\theta^\pi} \pi(s)], \tag{15}$$

where the action gradient of the performance measurement function $J(\pi)$ depends on the action-value function $Q(s, a)$, which is unknown and need to be estimated. To achieve the estimation, a critic network Q parameterized by θ^Q is used to approximate the action-value function and an actor network based on the current state offers control policy π that outputs the deterministic action in continuous space. In DDPG, the actor network and critic network are approximated by DNNs which can be learned via policy gradient method and error back propagation, respectively.

The use of neural networks for learning action-value function and control policy tends to be unstable. Thus, DDPG employs two important ideas to solve this problem.

REMARK 1. *A copy of the critic network and actor network: a target critic network and a target actor network are constructed and parameterized by vector $\theta^{Q'}$ and $\theta^{\pi'}$, respectively. These two target networks are adopted to calculate the target values, and the parameters $Q'$and π' in the two target networks slowly track the parameters Q and π in the original critic and actor network as follows:*

$$\theta^{Q'} = \kappa \theta^Q + (1 - \kappa)\theta^{Q'}, \tag{16}$$

$$\theta^{\pi'} = \kappa \theta^\pi + (1 - \kappa)\theta^{\pi'}, \tag{17}$$

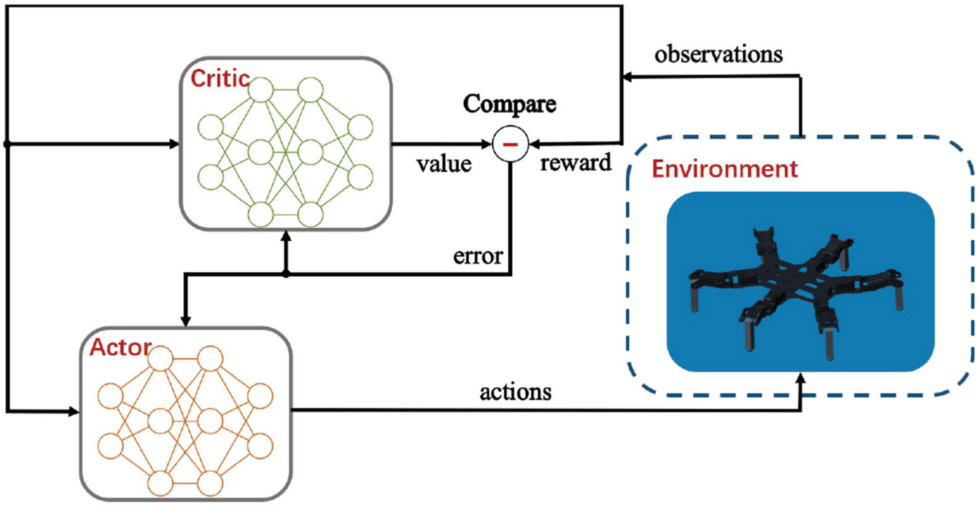


FIGURE 6 | The DDPG-based reinforcement learning structure.

where κ is positive and $\kappa << 1$. The updating mechanism is called soft update, which avoids non-stationary target values and enhances the stability.

REMARK 2. *Another challenge using neural networks for RL is the assumption that the samples are independently and identically distributed. Obviously, when the samples are generated from sequential exploration in an environment for the robot locomotion, this assumption is violated. To solve this, the replay buffer is used in DDPG. The replay buffer is a finite-size cache filled with transition samples. At each time step, both the actor network and the critic network are updated by sampling a mini batch uniformly from the buffer. Since DDPG is an off-policy learning algorithm, the replay buffer can be large where the algorithm benefits from a set of uncorrelated transitions. At each time step, the critic network θ^Q is updated by minimizing the loss:*

$$L = \frac{1}{H}\sum_{h=1}(Y_h - Q(s_h, a_h|\theta^Q))^2, \tag{18}$$

where

$$Y_h = r_h + \gamma Q'(s_{h+1}, \pi'(s_{h+1}|\theta^{\pi'})|\theta^{Q'}), \tag{19}$$

and h is the time step. H is the size of the mini batch sample.

The actor network θ^π is updated using the sampled policy gradient:

$$\nabla_{\theta^\pi} J = \frac{1}{H}\sum_{h=1} \nabla_a Q(s, a|\theta^Q)|_{s=s_h, a=\pi(s_h)} \nabla_{\theta^\pi} \pi(s|\theta^\pi)|_{s_h}, \tag{20}$$

4.3. Observation Space

The hexapod robot interacts with the environment through observations and actions. In order to apply DDPG on a practical system, the observation space is required to match the real robot and provides enough information for the agent to learn the task. In this work, a MDP observation vector o_t at time t is defined as:

$$o_t =< p_b, v_b, \mathcal{O}, \theta, \dot{\theta}, \tau, A >, \tag{21}$$

where $\mathcal{O}$ is the body orientation. A is the policy output including the amplitude and phase difference in the limb layer of the CPG.

In addition, the observation space consists of only part of the states, so the MDP can not be fully described. For example, the hexapod robot can not identify terrain types without any use of exteroceptive sensors. Hence, the process is referred as a Partially Observable Markov Decision Process (POMDP). Since in our work, the hexapod robot interacts with the environment through a continuous trajectory rather than a discrete action, we find that our observation space is sufficient enough for learning the desirable tasks.

4.4. Action Space

The control policy outputs the coupling parameters of the limb layer of the CPG which determine the intra-limb coordination as well as the adaptation to different terrain types. The action space is defined as follows:

$$a_t =< \mu, \theta_{i1}^{i2} >, \tag{22}$$

The action vector is transmitted as the input of the CPG network which generates the joint positions for each joint actuators.

The two-layer CPG network is chosen as the locomotion generator instead of learning joint position commands directly like most of the other studies (Hwangbo et al., 2019; Tsounis et al., 2020). There are three reasons for this: (1) the CPG network constrains the basic locomotion of the robot, which reduces the search space and accelerates the learning; (2) compared to 18 joint position or joint torque commands, learning symmetric CPG coupling parameters lowers the dimension of the action space; (3) the CPG network outputs smooth joint position commands, which are easier to be realized in the real robot.

TABLE 5 | Reward terms.

Term	Expression
Forward velocity	v_x
Energy consumption	$\lvert\tau \cdot \dot{\theta}\rvert + \tau^2$

4.5. Network Architecture

In DDPG, the critic network and actor network are parameterized as deep neural networks. **Figure 7** provides a graphical depiction of the NN model. The model is similar to the network architecture implemented in Fujimoto et al. (2018) and is proved to perform well. The critic network is composed of five hidden layers including three fully-connected (FC) layers and two ReLU layers. The actor network consists of six hidden layers including three fully-connected (FC) layers, two ReLU layers and a Tanh layer. The output modulates the proposed two-layer CPG parameters.

4.6. Reward Function

In this work, the environmental adaptability of the hexapod robot is measured by two criteria: one is the heading velocity of the robot and the other is the energy consumption of the robot. In general, precise reward function in robotics is one of the main challenges in solving the RL problem. Due to the constraints in RL, the reward function is simplified to encourage faster heading velocity and penalize higher energy consumption.

Table 5 shows the detailed reward terms. The velocity reward term motivates the robot to move forward as fast as possible, and it is tuned so that the robot receives a reward for a positive velocity up to a certain point. The penalty term is used to optimize the energy consumption of the robot. Hence, the reward term and the penalty term is intergraded into the reward function r_t as follows:

$$r_t = K_v \cdot v_x - K_e \cdot (\lvert\tau \cdot \dot{\theta}\rvert + \tau^2), \tag{23}$$

where K_v and K_e are the positive weights.

4.7. Guided Constrained Costs

In this work, two types of constrains are introduced into the proposed RL method (Gangapurwala et al., 2020). The first is the performance constraint, which restricts the hexapod robot into the region with potential good performance. The second is the safety constraint to avoid the robot exploring the region where damages may occur.

(1) Performance Constraint Costs: These costs are directly added to the reward function, as shown in **Table 6**. The constraint costs are guided by the kinematic model of the hexapod robot and help to improve the locomotion performance. For example, the Joint Speed term and Torque term are the limits of the actuator performance of the robot. In our control scheme, each supporting leg of the hexapod robot moves symmetrically, so the Orientation term and Height term are given to limit the robot from swinging too much.

TABLE 6 | Performance constraint costs.

Term	Expression
Joint speed	$\lVert max(\lvert\dot{\theta}\rvert - \dot{\theta}^{max}, 0)\rVert^2$
Torque	$\lVert max(\lvert\tau\rvert - \tau^{max}, 0)\rVert^2$
Orientation	$\lVert \mathscr{O} \rVert^2$
Height	$\lVert z_c - z_{c0} \rVert^2$

$\dot{\theta}^{max}$ and τ^{max} refer to the maximum outputs of the robot actuator. z_c and z_{c0} are the current and original body height of the hexapod robot. Orientation and Height limit the swing range of robot body.

TABLE 7 | Safety constraint costs.

Term	Expression
Joint speed	$\text{bool}(\dot{\theta} > \dot{\theta}^{max})$
Torque	$\text{bool}(\tau > \tau^{max})$
Fall	$\text{bool}(z_c < 0)$
Roll	$\text{bool}(\mathscr{O} > \mathscr{O}^{limit})$

bool(·) is a general Boolean judgement function.

(2) Safety Constraint Costs: For the cases when control policy outputs actions that cause the robot to land on unstable and unrecoverable states and damage the robot, the safety constraints are introduced in **Table 7**. The Fall term and Roll term are given to judge whether the robot falls or rolls over. If the control policy outputs the commands that robot can not carry on (see Joint Speed and Torque) or the robot falls and rolls over (see Fall and Roll), the training episode is terminated directly. The training steps explored in this episode are abandoned and a negative terminal reward is added to the last training step in the reformatted episode samples. This control policy avoids inefficient explorations of some constrained regions because the training episode is terminated if any safety constraint costs is true.

5. SIMULATIONS AND EXPERIMENTS

The proposed bio-inspired learning scheme is used to shape the hexapod robot locomotion. We evaluate the feasibility and performance of the motion policy via four different terrains in both simulations and experiments.

5.1. Simulations

The aim of these simulations is to guarantee the convergence of RL algorithm and obtain the theoretical maximum velocity of the hexapod robot in forward motion under different terrains.

The hexapod robot is modeled corresponding to the dimensions and mass of the actual hexapod robot prototype where the contact friction dynamics are determined by material surfaces. There are five main parameters for simulations: (1) learning rates for actor-critic network are set as 0.005 and 0.001, respectively; (2) the maximum number of episodes and steps in an episode are set as 1,400 and 200, respectively; (3) the sampling time is similar to the CPG cycle time which is 1 s; (4) the

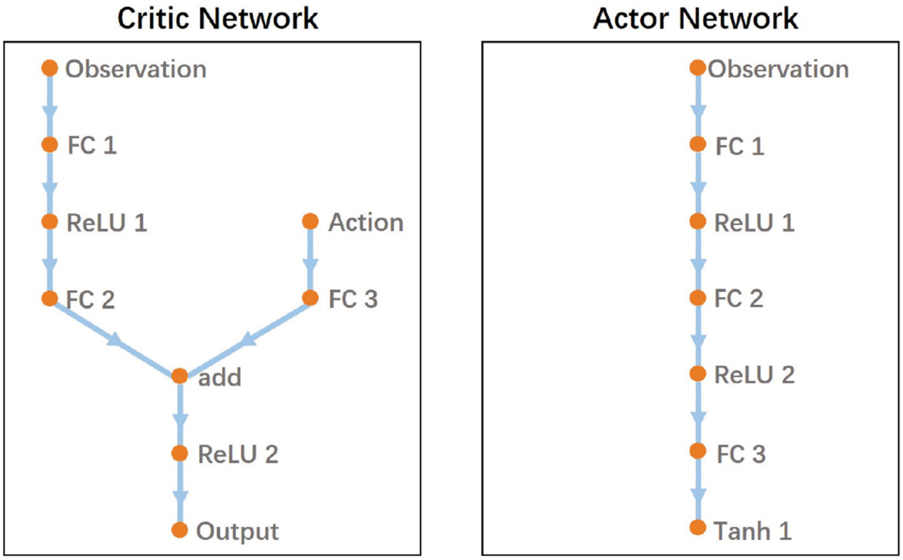


FIGURE 7 | The network architecture of the proposed DDPG-based RL.

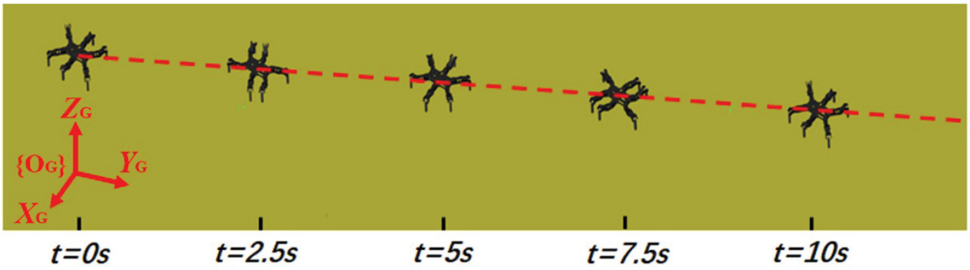


FIGURE 8 | The RL training simulation in an episode.

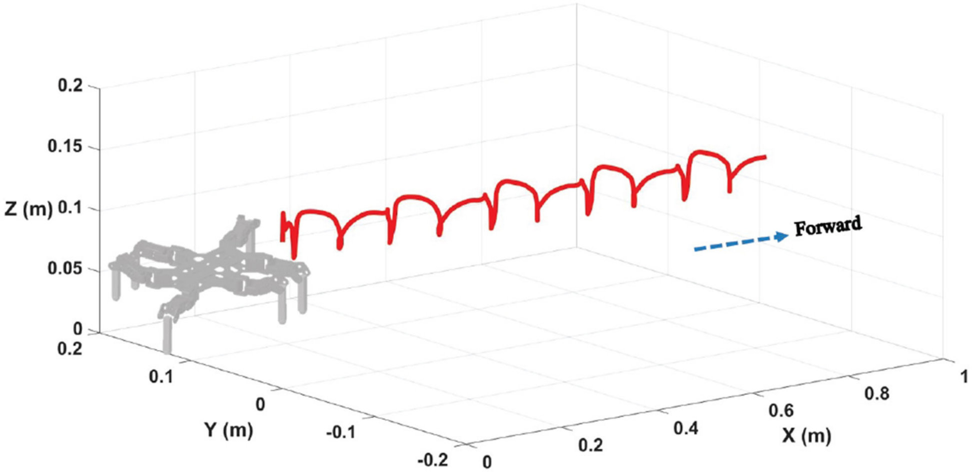


FIGURE 9 | The motion trajectory of the hexapod robot.

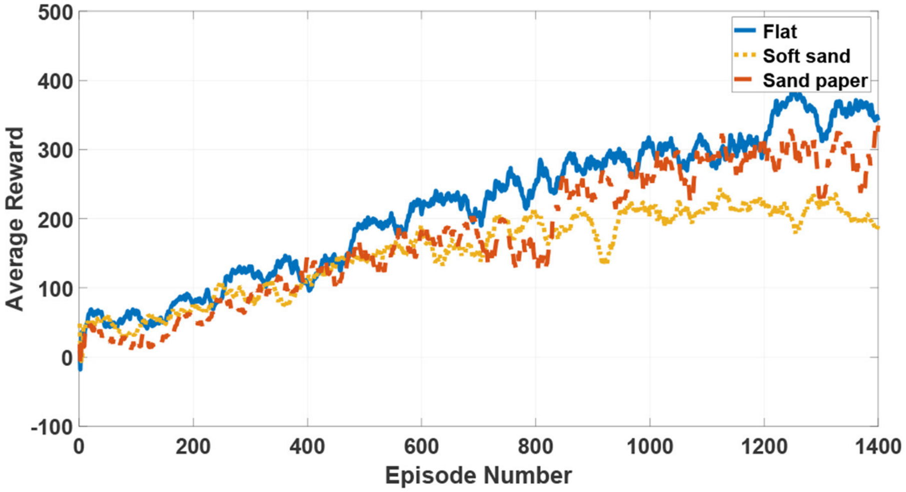


FIGURE 10 | The average reward of reinforcement learning in the tripod locomotion.

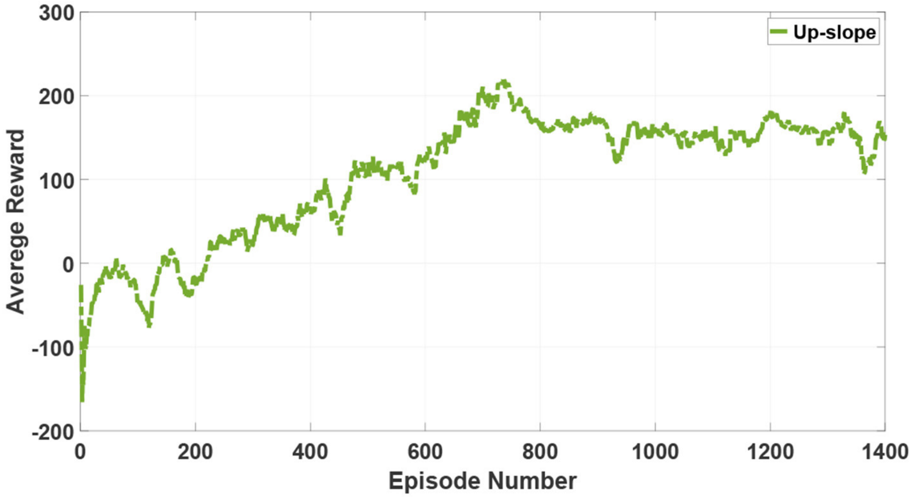


FIGURE 11 | The average reward of reinforcement learning in the quadruped locomotion.

frequency of the proposed two-layer CPG network is fixed on 0.5. The contact friction coefficients are modified according to different ground materials. The training process of an episode is randomly chosen in **Figure 8** and the motion trajectory is displayed in **Figure 9**.

From the two figures, it is noted that the hexapod robot walks well in the tripod locomotion without obvious offset in the direction of Y and Z axis. For a forward motion, we would like to emphasize that the sideslip in the vertical direction will cause an extremely uncertain deviation for the whole movement direction and posture. Therefore, the nearly tiny offset in Y axis illustrates the effectiveness of the proposed motion control scheme. As can be seen on Z axis, the slight fluctuation with the body height also reflects the control stability of the robot under the benefit of physical constraints.

In the training process, the observations are acquired from the hexapod dynamic model and the actions directly modulate the amplitude μ and phase difference θ_{mn}^{ij} in the limb layer of the CPG network. The reward function is given in the aforementioned section. At the end of 1,400 episodes, the average reward converges to a stable value in three terrain types as shown in **Figure 10**. The average training time in these terrains is approximately 6 h.

During learning processes, zero initial values drive the hexapod robot to swing around the origin and the value of reward function is equal to zero. After postures adapting and actions updating, the robot locomotion continuously becomes smoother. As the motion stability performance is improved, the reward value increases over time. Finally, the accumulative data samples help the robot reach the best motion state under the

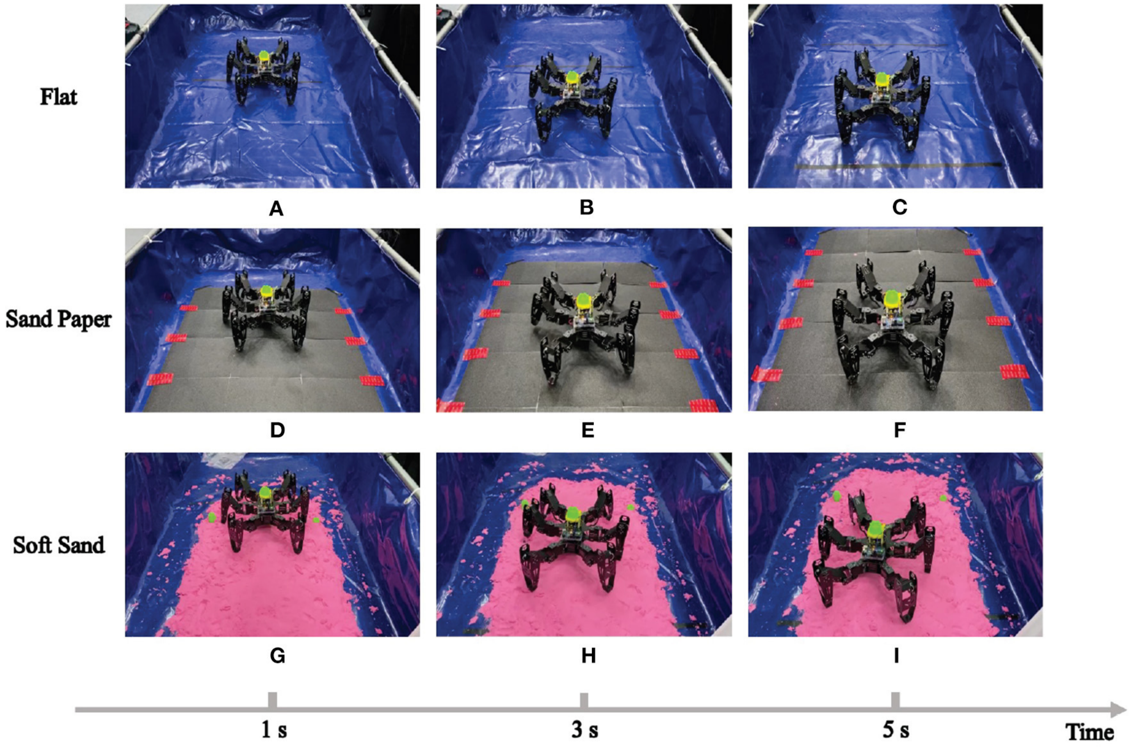


FIGURE 12 | The RL training experiment in different terrains.

specific locomotion mode and the reward function also converges to the biggest value. Compared with the movements in a sand paper and a soft sand, the steady velocity in a flat environment is a bit faster, but the convergence rate is conversely slower. As can be seen from the whole learning episodes, there are no obvious differences of the learning trend among three terrains. Besides, although the learning processes may suffer from several asymmetric and non-natural looking, even defective locomotion, the hexapod robot will finally converge to a stable and optimal locomotion under the limitation of several presetting constraints in the proposed DDPG-based learning approach.

As mentioned in section 3, tripod locomotion can be the fastest but inflexible. Therefore, when encountering complex and harsh terrains such as a slope or stones, the robot will switch to flexible locomotion modes such as quadruped locomotion or five legs support locomotion. In order to exhibit the locomotion flexibility derived from the proposed 3D two-layer CPG controller, an up-slope (10 deg) terrain is simulated and the quadruped locomotion is generated for repeating the aforementioned training process. The result of the average reward is represented in **Figure 11**.

Although the hexapod robot also accomplishes a fast convergence in an up-slope after 1,400 episodes, the average reward in such a tougher terrain is obviously lower than the stable value in a flat. Moreover, based on excellent adjustment characteristics of the proposed 3D two-layer CPG controller, the hexapod robot is endowed with the capability of locomotion transitions for adapting to complicated and unstructured terrains.

5.2. Experiments

Similarly, four experiments on different terrain surfaces, namely, flat, sand paper, soft sand, and up-slope, are conducted to validate the adaptivity and robustness of the proposed bio-inspired learning approach in practical scenarios. The training time is set as 5 s in each episode and other parameters set in these experiments are the same as simulation parameters. In addition, the simulation results can reduce the experimental training time through offering the hexapod robot an initial policy that performs the best in the simulations.

Firstly, environmental adaptability under individual locomotion mode has been tested. Here, the most common locomotion, tripod locomotion, is adopted as the training mode in three different terrain surfaces. The neural network in RL-based algorithm evolves three corresponding policies to make the robot perform well on the specific surfaces. The experiment scenes are shown in **Figures 12A–I**, where the same robot crawls on different contact surfaces.

During the repetitive episodes in the three mentioned terrains, the actual velocity of the hexapod robot with regard to the CPG tuning parameters, namely, the amplitude μ and phase difference θ_{mn}^{ij} in the limb layer are recorded. All these raw data are fitted and the learning results are shown in **Figures 14A–C**.

As can be seen from these figures, the results in different terrains have similar characteristic (the convex surface), but there are obvious differences in specific nodes. For example, when the amplitude is 1 with phase difference is 0.1, the velocity of the robot crawling on the flat will drop significantly from its maximum speed, but it decreased slowly on the soft sand.

Next, the quadruped locomotion in an up-slope (10 deg) is trained to evaluate the adaptability of different locomotion mode generated by the 3D two-layer CPG network (especially the body layer). The experimental platform and the learning result are illustrated in **Figures 13A–C**, **14D**, respectively. As observed in this experiment, the quadruped locomotion performs well and stably in up-slope environment showing a different trend compared with tripod locomotion, which explains the impact of basic locomotion patterns to the robot behavior.

Finally, the maximum velocities in the four terrains are calculated as listed in **Table 8**. It is noticed from **Table 8** that the hexapod robot runs fastest on the flat under tripod locomotion while it runs slowest under quadruped locomotion on the up-slope. Compared with the body length (BL) of the hexapod prototype (24 cm), the maximum velocity is 1/3 ∼ 1/2 BL from 7.35 to 13.10 (cm/s) in the preset terrains. As can be seen, since the surface friction coefficients of the chosen Sand Paper and Soft Sand terrain belong to the category of sand which may be similar to some extent, the difference of maximum velocity is not obvious.

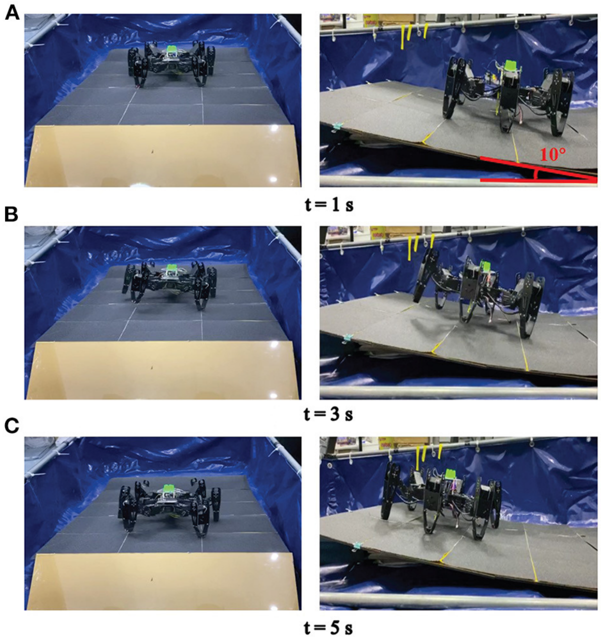
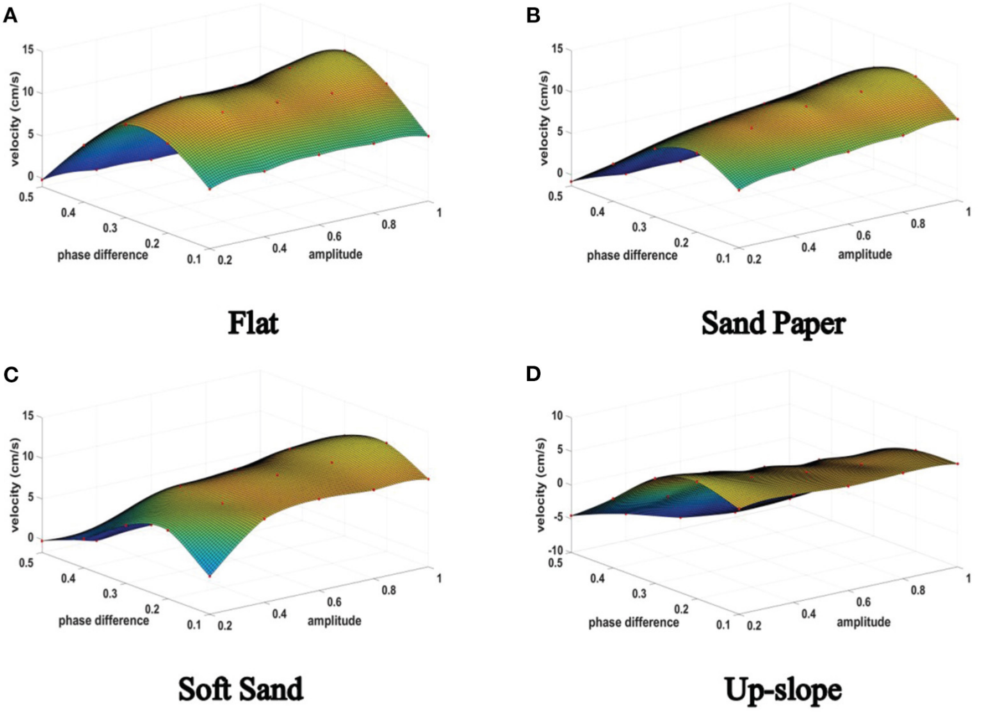


FIGURE 13 | The RL training experiment in an up-slope.

TABLE 8 | The maximum velocity in different terrains.

Terrain types	**Velocity(cm/s)**	
	Simulations	**Experiments**
Flat	22.27	13.10
Sand paper	19.87	11.86
Soft sand	15.40	11.72
Up-slope	13.05	7.35



Flat

Sand Paper

Soft Sand

Up-slope

FIGURE 14 | The hexapod robot velocity in different terrains.

It should be also emphasized that, suffering from inaccuracies in modeling process as well as environment construction uncertainties, several deviations between simulation model and actual experiment exist inevitably. For instance, the parametric variables in simulation are fixed, while all the variables are inherently floating in the experiments. Nevertheless, the simulation results still have a certain association with the experimental results, which effectively offer prior information at the beginning of experiment settings and greatly accelerate the convergence rate in the actual system.

In summary, it can be concluded from the experimental results that the proposed 3D two-layer CPG network and the DDPG-based RL algorithm can provided the hexapod robot with excellent maneuverability and environmental adaptability performance while the stability and robustness of the overall control scheme can be also achieved.

6. CONCLUSION

This paper aims to investigate an adaptive locomotion control approach for a hexapod robot. Inspired by biological neuron control systems, the proposed locomotion controller is composed of a set of coupled oscillators, namely an artificial CPG network. The novelty of the CPG controller lies in its 3D two-layer. The first layer of the CPG is able to control the basic locomotion patterns according to the environment information, while a RL-based learning algorithm is adopted for fine-tuning the second layer of the CPG to regulate the behavior of robot limbs. Several numerical studies and experiments have been conducted to demonstrate the valid and effectiveness of the proposed locomotion controller. The navigation of the robot in a complex and dynamic environment will be explored in the next research phase.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors contributed to the theory and implementation of the study. WO designed the whole locomotion control scheme, proposed the two-layer CPG, and wrote the first draft of the manuscript. HC modeled the hexapod robot and carried on the experiments. JP offered the simulation of the reinforcement learning part. WL corrected the paper formation and text required for the journal. QR determines the final Abstract, Introduction, and Conclusion. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 61773271) and Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT20066).

REFERENCES

Achiam, J., Held, D., Tamar, A., and Abbeel, P. (2017). "Constrained policy optimization," in *2017 34th International Conference on Machine Learning (ICML)* (Ningbo), 30–47.

Azayev, T., and Zimmerman, K. (2020). Blind hexapod locomotion in complex terrain with gait adaptation using deep reinforcement learning and classification. *J. Intell. Robot. Syst*. 99, 659–671. doi: 10.1007/s10846-020-01162-8

Bal, C. (2021). Neural coupled central pattern generator based smooth gait transition of a biomimetic hexapod robot. *Neurocomputing* 420, 210–226. doi: 10.1016/j.neucom.2020.07.114

Barfoot, T. D., (2006). Experiments in learning distributed control for a hexapod robot. *Robot. Auton. Syst*. 54, 864–872. doi: 10.1016/j.robot.2006.04.009

Campos, R., Matos, V., and Santos, C. (2010). "Hexapod locomotion: a nonlinear dynamical systems approach," in *2010 36th Annual Conference on IEEE Industrial Electronics Society (IECON)* (Glendale, CA), 1546–1551. doi: 10.1109/IECON.2010.5675454

Chung, H.-Y., Hou, C. C., and Hsu, S. Y. (2015). Hexapod moving in complex terrains via a new adaptive cpg gait design. *Indus. Robot* 42, 129–141. doi: 10.1108/IR-10-2014-0403

Delcomyn, F. (1980). Neural basis of rhythmic behavior in animals. *Science* 210, 492–498. doi: 10.1126/science.7423199

Fortuna, L., Frasca, M., and Arena, P. (2004). *Bio-Inspired Emergent Control of Locomotion Systems*. Singapore: World Scientific. doi: 10.1142/5586

Fujimoto, S., Van Hoof, H., and Meger, D. (2018). "Addressing function approximation error in actor-critic methods," in *35th International Conference on Machine Learning* (Stockholm), 2587–2601.

Gangapurwala, S., Mitchell, A., and Hacoutis, I. (2020). Guided constrained policy optimization for dynamic quadrupedal robot locomotion. *IEEE Robot. Autom. Lett*. 5, 3642–3649. doi: 10.1109/LRA.2020.2979656

Grzelczyk, D., Szymanowska, O., and Awrejcewicz, J. (2018). Kinematic and dynamic simulation of an octopod robot controlled by different central pattern generators. *Proc. Instit. Mech. Eng. Part I J. Syst. Control Eng*. 233, 400–417. doi: 10.1177/0959651818800187

Hooper, S. L. (2000). Central pattern generators. *Curr. Biol*. 10, 176–177. doi: 10.1016/S0960-9822(00)00367-5

Hwangbo, J., Lee, J., Dosovitskiy, A., Bellicoso, D., Tsounis, V., Koltun, V., et al. (2019). Learning agile and dynamic motor skills for legged robots. *Sci. Robot*. 4:eaau5872. doi: 10.1126/scirobotics.aau5872

Hyun, D. J., Seok, S. O., Sang, O., Lee, J., and Kim, S. (2014). High speed trot-running: implementation of a hierarchical controller using proprioceptive impedance control on the MIT cheetah. *Int. J. Robot. Res*. 33, 1417–1445. doi: 10.1177/0278364914532150

Juang, C.-F., Chang, Y. C., and Hsiao, C. M. (2011). Evolving gaits of a hexapod robot by recurrent neural networks with symbiotic species-based particle swarm optimization. *IEEE Trans. Indus. Electron*. 58, 3110–3119. doi: 10.1109/TIE.2010.2072892

Kecskés, I., Székics, L., Fodor, J. C., and Odry, P. (2013). "PSO and GA optimization methods comparison on simulation model of a real hexapod robot," in *2013 IEEE 9th International Conference on Computational Cybernetics, Proceedings (ICCC)* (Tihany), 125–130. doi: 10.1109/ICCCyb.2013.6617574

Kim, D., Jorgensen, S., Lee, J., Ahn, J., Luo, L., and Sentis, L. (2020). Dynamic locomotion for passive-ankle biped robots and humanoids using whole-body locomotion control. *Int. J. Robot. Res*. 39, 936–956. doi: 10.1177/0278364920918014

Lele, A. S., Fang, Y., Ting, J., and Raychowdhury, A. (2020). "Learning to walk: spike based reinforcement learning for hexapod robot central pattern generation," in *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)* (Genova), 208–212. doi: 10.1109/AICAS48895.2020.9073987

Li, T., Lambert N., Calandra, R., Meier, F., and Rai, A. (2019). Learning generalizable locomotion skills with hierarchical reinforcement learning. *arXiv:1909.12324v1*. doi: 10.1109/ICRA40945.2020.9196642

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., et al. (2016). "Continuous control with deep reinforcement learning," in *International Conference on Learning Representations* (San Juan).

Niu, X., Xu, X., Ren, Q., and Wang, Q. (2014). Locomotion learning for an anguilliform robotic fish using central pattern generator approach. *IEEE Trans. Indus. Electron*. 61, 4780–4787. doi: 10.1109/TIE.2013.2288193

Ramdya, P., Thandiackal, R., Cherney, R., Asselborn, T., Benton, R., Ijspeert, A. J., et al. (2017). Climbing favours the tripod gait over alternative faster insect gaits. *Nat. Commun*. 8:14494. doi: 10.1038/ncomms14494

Roy, S. S., and Pratihar, D. K. (2013). Kinematics, dynamics and power consumption analyses for turning motion of a six-legged robot. *J. Intell. Robot. Syst*. 74, 663–688. doi: 10.1007/s10846-013-9850-6

Sartoretti, G., Paivine, W., Shi, Y., Wu, Y., and Choset, H. (2019). Distributed learning of decentralized control policies for articulated mobile robots. *IEEE Trans. Robot*. 35, 1109–1122. doi: 10.1109/TRO.2019.2922493

Seo, K., Chung, S. J., and Slotine, J. J. E. (2010). CPG-based control of a turtle-like underwater vehicle. *Auton. Robots* 28, 247–269. doi: 10.1007/s10514-009-9169-0

Stelzer, A., Hirschmüller, H., and Görner, M. (2012). Stereo-vision-based navigation of a six-legged walking robot in unknown rough terrain. *Int. J. Robot. Res*. 31, 381–402. doi: 10.1177/0278364911435161

Sun, Q., Gao, F., and Chen, X. (2018). Towards dynamic alternating tripod trotting of a pony-sized hexapod robot for disaster rescuing based on multi-modal impedance control. *Robotica* 36, 1048-1076. doi: 10.1017/S026357471800022X

Tan, J., Zhang, T., Coumans, E., Iscen, A., Bai, Y., Hafner, D., et al. (2018). Sim-to-real: Learning agile locomotion for quadruped robots. *arXiv [Preprint] arXiv:1804.10332*.

Tsounis, V., Alge, M., and Lee, J. (2020). Deepgait: planning and control of quadrupedal gaits using deep reinforcement learning. *IEEE Robot. Autom. Lett*. 5, 3699–3706. doi: 10.1109/LRA.2020.2979660

Yu, H., Gao, H., and Deng, Z. (2020). Enhancing adaptability with local reactive behaviors for hexapod walking robot via sensory feedback integrated central pattern generator. *Robot. Auton. Syst*. 124:103401. doi: 10.1016/j.robot.2019.103401

Yu, H., Gao, H., Ding, L., Li, M., Deng, Z., Liu, G., et al. (2016). Gait generation with smooth transition using cpg-based locomotion control for hexapod walking robot. *IEEE Trans. Indus. Electron*. 63, 5488–5500. doi: 10.1109/TIE.2016.2569489

Zarrouk, D., and Fearing, R. S. (2015). Controlled in-plane locomotion of a hexapod using a single actuator. *IEEE Trans. Robot*. 31, 157–167. doi: 10.1109/TRO.2014.2382981

Zhang, H., Liu, Y., Zhao, J., Chen, J., and Yan, J. (2014). Development of a bionic hexapod robot for walking on unstructured terrain. *J. Bion. Eng*. 11, 176–187. doi: 10.1016/S1672-6529(14)60041-X

Zhao, D., and Revzen, S. (2020). Multi-legged steering and slipping with low dof hexapod robots. *Bioinspir. Biomimet*. 15:045001. doi: 10.1088/1748-3190/ab84c0

Zhong, G., Chen, L., Jiao, Z., and Li, J. (2018). Locomotion control and gait planning of a novel hexapod robot using biomimetic neurons. *IEEE Trans. Control Syst. Technol*. 26, 624–636. doi: 10.1109/TCST.2017.2692727

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

ORIGINAL RESEARCH
published: 01 March 2021
doi: 10.3389/fnbot.2021.642733

Mobile Robot Path Planning Based on Time Taboo Ant Colony Optimization in Dynamic Environment

Ni Xiong [1], Xinzhi Zhou [1]*, Xiuqing Yang [2,3]*, Yong Xiang [2,3]* and Junyong Ma [2,3]

[1] College of Electronics and Information Engineering, Sichuan University, Chengdu, China, [2] The Second Research Institute of Civil Aviation Administration of China, Chengdu, China, [3] Civil Aviation Logistics Technology Company Limited, Chengdu, China

This article aims to improve the problem of slow convergence speed, poor global search ability, and unknown time-varying dynamic obstacles in the path planning of ant colony optimization in dynamic environment. An improved ant colony optimization algorithm using time taboo strategy is proposed, namely, time taboo ant colony optimization (TTACO), which uses adaptive initial pheromone distribution, rollback strategy, and pheromone preferential limited update to improve the algorithm's convergence speed and global search ability. For the poor global search ability of the algorithm and the unknown time-varying problem of dynamic obstacles in a dynamic environment, a time taboo strategy is first proposed, based on which a three-step arbitration method is put forward to improve its weakness in global search. For the unknown time-varying dynamic obstacles, an occupancy grid prediction model is proposed based on the time taboo strategy to solve the problem of dynamic obstacle avoidance. In order to improve the algorithm's calculation speed when avoiding obstacles, an ant colony information inheritance mechanism is established. Finally, the algorithm is used to conduct dynamic simulation experiments in a simulated factory environment and is compared with other similar algorithms. The experimental results show that the TTACO can obtain a better path and accelerate the convergence speed of the algorithm in a static environment and can successfully avoid dynamic obstacles in a dynamic environment.

Keywords: path planning, mobile robot, ant colony algorithm, dynamic environment, time taboo strategy

Edited by:
Zhan Li,
University of Electronic Science and Technology of China, China

Reviewed by:
Dechao Chen,
Hangzhou Dianzi University, China
Ameer Hamza Khan,
Hong Kong Polytechnic University, Hong Kong
Xin Luo,
Chongqing Institute of Green and Intelligent Technology (CAS), China

***Correspondence:**
Xinzhi Zhou
xz.zhou@scu.edu.cn
Xiuqing Yang
yangxiuqing@caacetc.com
Yong Xiang
xiangyong@caacetc.com

Received: *16 December 2020*
Accepted: *25 January 2021*
Published: *01 March 2021*

Citation:
Xiong N, Zhou X, Yang X, Xiang Y and Ma J (2021) Mobile Robot Path Planning Based on Time Taboo Ant Colony Optimization in Dynamic Environment. Front. Neurorobot. 15:642733. doi: 10.3389/fnbot.2021.642733

INTRODUCTION

In mobile robot navigation, global path planning has always been one of the research hotspots. At present, the research on path planning of mobile robot in static environment has been relatively mature, bringing many excellent kinds of algorithms. There are some traditional algorithms, such as A* algorithm, artificial potential field method (Rimon and Koditschek, 1992), and Dijkstra. Besides, heuristic optimization algorithms are also among the lists, including genetic algorithm, neural network algorithm (Khan et al., 2020c), particle swarm optimization algorithm, ant colony algorithm (Fan et al., 2003), cuckoo algorithm (Mohanty and Parhi, 2016), bug algorithm (Khan et al., 2020a), and so on. However, all kinds of algorithms are more or less limited to algorithm deficiencies. Compared with the traditional gradient descent algorithm, the metaheuristic algorithm performs better in convergence speed, and global optimization ability and hence is widely used in trajectory planning (Khan et al., 2020b), prediction, resource scheduling, and other fields. Ant colony algorithm with its good robustness, positive feedback, and parallel computing ability has been widely used in robot path planning and achieved good results.

In Zhou et al. (2013), an improved ant colony optimization (ACO) algorithm is proposed. By modifying the initial environment pheromone and state transition probability, the search deadlock can be eliminated. The combination of deterministic search and random search can reduce redundant paths. Jiang et al. (2019) adopt a method to make the initial pheromone uneven, which can reduce the blind search path of ants, cut down the running time of the algorithm, and improve the convergence rate of the algorithm. In Luo et al. (2019), the improved pheromone updating strategy was used to update the excellent path and punish the poor path; meanwhile, the upper and lower limits of pheromone were set. In Liu and Zhang (2016), the simulated annealing algorithm is added to the process of simulated annealing ant colony algorithm formed by ant colony algorithm, and it is applied to path planning to solve the local optimal problem caused by premature ant colony algorithm. However, the working environment of mobile robot is always dynamic, such as intelligent factory, hospital, supermarket, and so on. Therefore, it is a difficult problem to avoid obstacles successfully and make quadratic optimal path planning after encountering dynamic obstacles. In Li et al. (2019), an inertial positioning strategy is proposed to enable the robot to predict the position of the target in advance. From the predicted position, the robot path is generated by cubic spline interpolation, and then the improved particle swarm optimization algorithm with random positive feedback factor in speed update is used to optimize the path. As it is very important to track and predict dynamic obstacles in dynamic environment, Ferguson et al. (2008) proposed that if a dynamic obstacle runs on a straight road, it is likely that it will continue to travel along the same straight line in the future. Li et al. (2017) constructed a prediction model to avoid obstacles. In Kim et al. (2017), a trajectory prediction method based on occupancy grid and neural network is proposed. In Qu and Huang (2015), different obstacle avoidance strategies were set up to avoid dynamic obstacles when the mobile robot met with dynamic obstacles. The size of dynamic obstacles is considered in Xu et al. (2019) on the basis of Qu and Huang (2015). The above literatures predict that the dynamic obstacles are moving on a straight line without considering the change of the speed of the dynamic obstacles. There are unknown dynamic obstacles in the dynamic environment. Ant colony algorithm also has the problems of slow convergence speed and poor global search ability.

This article proposes a novel, improved ACO algorithm using time taboo strategy, namely, time taboo ant colony algorithm, which aims to solve the path planning problem of mobile robot in dynamic environment. It utilizes the adaptive initial pheromone uneven distribution to reduce the blindness of ants in early path finding. And, the problem of deadlock is solved by the rollback strategy; pheromone preferential limited update is adopted to reduce pheromone redundancy. The above improvements are designed to improve the convergence speed and global search ability of the algorithm. The corresponding strategies, mechanisms, and prediction models are proposed to avoid dynamic obstacles. Based on the improvement, strategy, and prediction model mentioned above, the dynamic obstacle avoidance of mobile robot is effectively realized and verified by simulation.

ANT COLONY OPTIMIZATION

ACO is a heuristic global optimization algorithm in evolutionary algorithm. In the process of searching the path, ants release pheromones on the path through which they have passed. The shorter the path is, the higher the concentration of pheromone released will be. Therefore, the following ants will favorably choose the path with highest concentration of pheromone, as a result of which an optimal path can be obtained.

In order to reduce the detour of ants in the process of searching, the algorithm introduces the concept of taboo table. In the process of ant searching, the taboo table is added to the path nodes that have been explored, and the next time ants select the path nodes, optional nodes will exclude those in the taboo table. At the same time, the heuristic function η is introduced, and the following transfer function is used to improve the search efficiency of ant colony. The transfer formula (1) and the heuristic function (2) are as follows:

$$P = \begin{cases} \frac{\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}}{\sum_{s\in A}\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}} & if\ j \in A \\ 0 & else \end{cases} \tag{1}$$

$$\eta_{ij}(t) = \frac{1}{d_j} \tag{2}$$

In the formula,τ_{ij} represents the pheromone concentration from the ith node to the jth node on the feasible path node. η_{ij} represents the heuristic value from the ith path node to the jth path on the feasible path, and its value is the reciprocal of the distance from the jth node to the end point, as shown in formula (2), d_j represents the distance from the jth node to the end point. S represents the current path node. A represents the nodes that ants can still choose after removing obstacles and path nodes in the taboo table. α and β represent the importance of pheromone and heuristic pheromone, respectively. According to this formula, we can know that the higher the pheromone concentration is, the higher the heuristic pheromone will be, and the greater the probability of ants selecting the node will be.

At the same time, the pheromone concentration on the map will gradually evaporate with the running time. The updating of pheromone is shown in the following formulas (3) and (4).

$$\tau_{ij}(t+\Delta t) = (1-\rho)\,\tau_{ij}(t) + \Delta\tau_{ij}(t) \tag{3}$$

$$\Delta\tau_{ij}(t) = \sum\nolimits_{k=1}^{m} \Delta\tau_{ij}^{k}(t) \tag{4}$$

ρ stands for the evaporation coefficient of pheromone. The larger the value is, the more the pheromone evaporates, with its value between (0 and 1). Pheromone evaporation can avoid excessive accumulation of pheromone. $\Delta\tau_{ij}(t)$ represents the pheromone increment from the ith node to the jth node at time t. Its definition is shown by formula (5).

$$\Delta\tau_{ij}^{k}(t)=\begin{cases}\frac{Q}{L_k} & P(i,j)\in P\\ 0 & P(i,j)\notin P\end{cases} \tag{5}$$

In the formula, K represents the kth ant, L_k represents the total length of the kth ant's path, P represents the nodes in the path, $P(i, j)$ represents the path nodes from i to j, and Q represents the pheromone strength. The formula shows that only the ants that have reached to the destination can update pheromones, and the pheromone concentration is inversely proportional to the length of the path. Although the ants in the ant colony do not communicate directly, they communicate through pheromone, an indirect medium, to achieve the goal of optimal path planning.

ANT COLONY OPTIMIZATION IMPROVEMENT

Some factors will directly affect the performance of ant colony algorithm, such as pheromone, pheromone updating rule, heuristic pheromone, state transition rule, and deadlock and self-locking of ant colony algorithm. However, the improvement of the initial pheromone is still rigid, which is not conducive to the global search of the algorithm, and the processing of pheromone change caused by the improved deadlock problem is not perfect yet. In this article, pheromone, deadlock and self-locking, and pheromone update mechanism are improved. Based on these improvements, the time taboo ACO (TTACO) is further proposed.

Adaptive Initial Pheromone Distribution

In the basic ant colony algorithm, the pheromone distribution is quite even; even pheromone distribution may cause ants wander around and backtrack, rendering too-long search time of ants in the early stage of the algorithm, and the length of the search path is increased. In this article, the pheromone near the connection line is added according to the distance from the starting point to the end point, and the pheromone of the node closer to the connection line is higher in concentration. At the same time, in order to prevent the algorithm from falling into local optimum due to the initial pheromone, the algorithm will determine the number of pheromones added according to the proportion of obstacles on and near the starting and ending lines. The formula is as follows:

$$\begin{cases}\tau=\tau_0+\vartheta C\\ \vartheta=\mu\varepsilon\end{cases} \tag{6}$$

In the formula, τ is the pheromone of ant colony algorithm, τ_0 is the basic pheromone, ϑ is the adaptive parameter, and μ is the distance between the node and the line; the closer the distance is, the greater the value will be, with the range of the value being [0, 1]. When the distance between the obstacle and the connection exceeds a set value, the value of u is taken as zero. The closer the grid is to the line, the higher the pheromone concentration will be. Because of the existence of obstacles, ε will decrease, resulting in the decrease of the overall pheromone. According to the connection between the starting point and the end point, it is highly probable that the optimal path is the path near the line, and the condition of the obstacle near the line will hinder the degree of the optimal path near the line. Therefore, the uneven distribution of the initial pheromone is conducive to the ants' searching for the best path in the early stage.

Solving Deadlock and Self-Locking Problems

When the ants explore the path in the early stage, because of the restriction of the map environment and the taboo table, the individual ant may fall into the dilemma of finding nowhere to go before reaching the end point. It can be divided into deadlock and self-locking. Inspired by Dai et al. (2019), this article uses a rollback strategy to solve the deadlock and self-locking problems and divides the taboo table into a global taboo table and a local taboo table. The global taboo table records information of deadlock location, and the local taboo table, inherited from the global taboo table, records the information of the ant's walking path node and the self-locking position node. The ant chooses the path according to the local taboo table. When the ant has a deadlock problem caused by the environment, the ant is made to retreat two steps, and the place where the ant falls into the deadlock will be added to the global taboo table to reduce the possibility that the ant falls into the same deadlock again. When the ant is self-locked because of its walking path, the ant is made to go back two steps. Different from the deadlock, the position where the ant self-locking occurs will be added to the local taboo table to prevent the ant from repeatedly walking into the same self-locking node.

Pheromone Update Rules

Because of the introduction of the rollback strategy, all ants can successfully find the end point. Although the search ability of the algorithm is improved, the increase in the number of ants that successfully reach the end point leads to pheromone redundancy. So pheromones of ants will be updated only with the shortest path and shorter path to reduce the redundancy of pheromones on the map. The number of ants updated with the pheromone is determined by the scale of the map. The larger the map is, the more ants can get pheromone updates.

TTACO

The basic ant colony algorithm introduces a taboo table, but its use of the taboo table is limited to preventing ants from walking nodes that have been already walked. In order to further improve the algorithm's global search capability, algorithm convergence speed, and the ability to avoid dynamic obstacles, this article proposes a time taboo strategy: In the case of path redundancy or detection of dynamic obstacles, ants will be prohibited from accessing certain path nodes for a certain period of time until path redundancy is avoided, or dynamic obstacles are successfully avoided, and then the prohibition is canceled in a specific way. And based on this strategy, a three-step arbitration method and an occupancy grid prediction model are proposed to improve

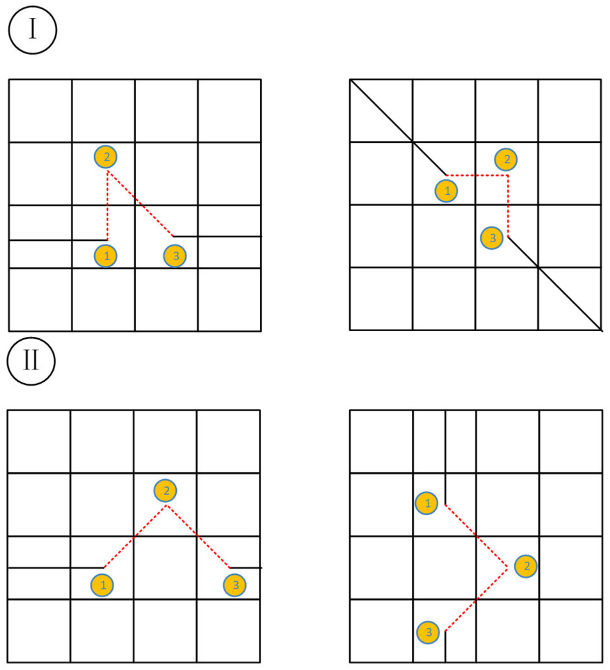


FIGURE 1 | Schematic diagram of path redundancy.

TABLE 1 | Determination rules.

Path condition	Determination formula
Situation I	$S_{13} < 2$
Situation II	$S_{13} == 2 \&\& S_{12} != 1$

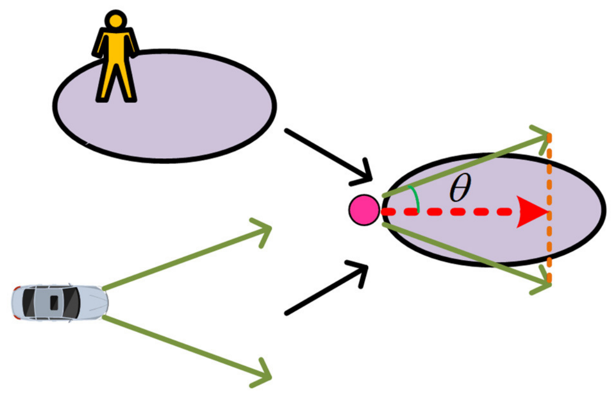


FIGURE 2 | Schematic diagram of dynamic obstacle movement area.

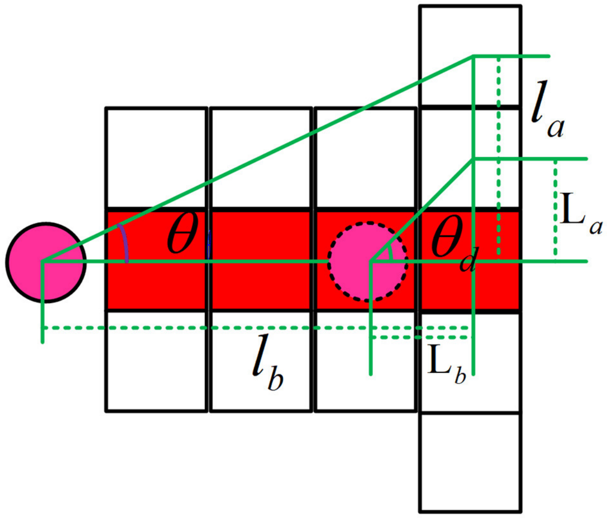


FIGURE 3 | Schematic diagram of dynamic obstacle prediction angle.

the algorithm's global search ability and the ability to avoid dynamic obstacles.

Three-Step Arbitration Method

In the crawling process of ants, there will inevitably be path redundancy, as for which this article divides it into two situations: one is redundancy situation I, and the other is redundancy situation II. The two cases are shown in **Figure 1**, where the top two grids show situation I, and the bottom two are situation II.

When the redundancy situation occurs, the path is divided into three steps. The first step is before the situation occurs, the second step is when the situation occurs, and the third step is after the situation occurs. The specific judgment method is as follows. S_{12} and S_{13} are used to represent the distance between the first step and the second or third step, respectively. The specific determination rules are shown in **Table 1** below.

When it is determined that the second step is redundant, the algorithm will record the node where the problem occurs. When the ant crawls to the first step, the algorithm will temporarily add the path node of the second step to the taboo table, so that the ant cannot choose the second step. In this way, the algorithm will accelerate the convergence speed and improve global search capability. The crawling paths of ants in the early stage are more chaotic, and more redundant ones will also occur, so the method is disabled in this period. The algebra that specifically enables the method is shown in formula (7):

$$K_s = mod(MM, 5) \tag{7}$$

K_s is the algebra of activating the method, and MM is the scale of the map. The larger the scale of the map is, the later the method will be activated. In order to enable a suitable number of ants to invoke the method and avoid multiple redundant path nodes caused by an individual ant to trigger the method at the same time, which will lead to redundant path, and having referred to the cuckoo algorithm (Luan et al., 2015), the following trigger mechanism is proposed.

1. This method is only enabled from the generation of ants crawling a relatively stable path.
2. The use of this method or not will be determined by the probability before each ant searches.

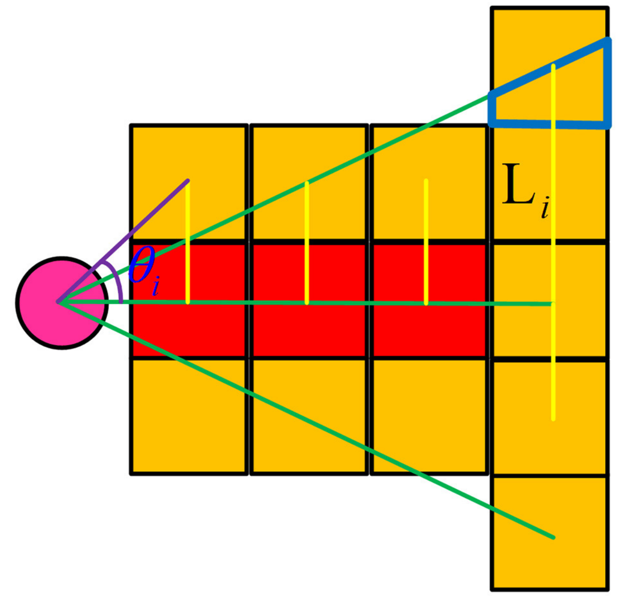


FIGURE 4 | Schematic diagram of prediction of dynamic obstacle occupancy grids.

3. When the ant using this method encounters a redundant node, the use of this method or not will be once again determined at the node according to the probability.

Occupancy Grid Prediction Model

Highly dynamic scenes have brought great challenges to robot path planning. In Zhao et al. (2020), the classical method of detecting and planning at the same time cannot meet the requirement of safely avoiding dynamic obstacles. Kim et al. (2017) use a prediction method based on occupancy grids, where the scene is divided into grids, and the occupancy grids are used to represent the possible locations of dynamic obstacles in the future. Ferguson et al. (2008) propose that if a dynamic obstacle is driving on a straight road, it is very likely that it will continue to drive along the same straight line in the future. Rummelhard et al. (2014) proposes a new grid-based collision risk prediction method. Li et al. (2017) take the mobile robot's kinematics model as the algorithm's trajectory prediction model. Riosmartinez et al. (2013) propose a pedestrian area based on Gaussian distribution. Two individual areas of different sizes are constructed with pedestrians as the center, with the area in front of the pedestrian being lager, assuming that our dynamic obstacles are most likely to be moving mechanical vehicles and pedestrians, as shown in **Figure 2**.

The kinematics of mobile robot is described in Boukens et al. (2019) and Keyvan et al. (2020). The motion model of a dynamic obstacle can be expressed by the following differential equation, where (x, y) is the position, θ is the angle of the dynamic obstacle velocity, v is the velocity, k is the curvature, and ω is the angular velocity. It can be seen from the formula that when the velocity is constant, the greater the curvature is, the greater the value of θ will be, and the greater the angular velocity ω of the dynamic obstacle will be.

$$\begin{cases} \dot{x} = vcos(\theta) \\ \dot{y} = vsin(\theta) \\ \dot{\theta} = kv \end{cases} \tag{8}$$

$$\omega = vk \tag{9}$$

It is assumed that the dynamic obstacle has two predicted steering angles, a safe steering predicted angle and an emergency steering predicted angle. Let θ in the above formula be the safe steering prediction angle, that is, the angle at which a slight path deviation of dynamic obstacles may occur is considered, and the safe steering angle is considered only when starting to predict. θ_d is the emergency steering predicted angle, which represents the angle when the emergency steering of the dynamic obstacle occurs, and each step of the prediction has an emergency steering angle. The schematic diagram is shown in **Figure 3**.

In the figure, θ is the predicted angle of safe steering, and θ_d is the predicted angle of emergency steering, and its formula is as follows:

$$\theta = \arctan(\frac{l_a}{l_b}) \tag{10}$$

$$\theta_d = \arctan(\frac{L_a}{L_b}) \tag{11}$$

In the above formulas, L_a is the offset distance in the vertical direction when the dynamic obstacle makes an emergency turn, and L_b is the distance in the horizontal direction when the dynamic obstacle is making an emergency turn. l_a is the vertical distance when the dynamic obstacle slightly deviates from the path, and l_b is the horizontal distance when the dynamic obstacle slightly deviates from the path. And the angle of θ is set as $(0, 26.12°)$ and the angle of θ_d $(0, 45°)$.

Figure 4 shows a schematic diagram of path nodes that have been marked as occupied. According to the previously set angle, further specific analysis of whether a path node is affected by dynamic obstacles is determined by the following formula, where m is the ratio of the affected area of a grid to the whole grid area. As shown in **Figure 4**, the ratio of the area of the blue trapezoid to the area of the whole grid and the value is affected by θ. L_i is the vertical distance from the center of the ith grid to the velocity direction of the dynamic obstacle, as shown by the golden yellow straight line in **Figure 4**. γ is an adaptive parameter whose value is shown in formula (13). θ_i is the angle between the center of the ith grid and the speed direction of the dynamic obstacle, which includes θ_d. When the value of f is > 0.5, it means that the grid is affected and will be tinted yellow.

$$f = \frac{\gamma}{L_i} + m \tag{12}$$

$$\gamma = sin(\theta_i) \tag{13}$$

Combining the possible movement of the dynamic obstacle and the dynamic obstacle movement model, the following occupancy grid prediction model is proposed. When the mobile robot

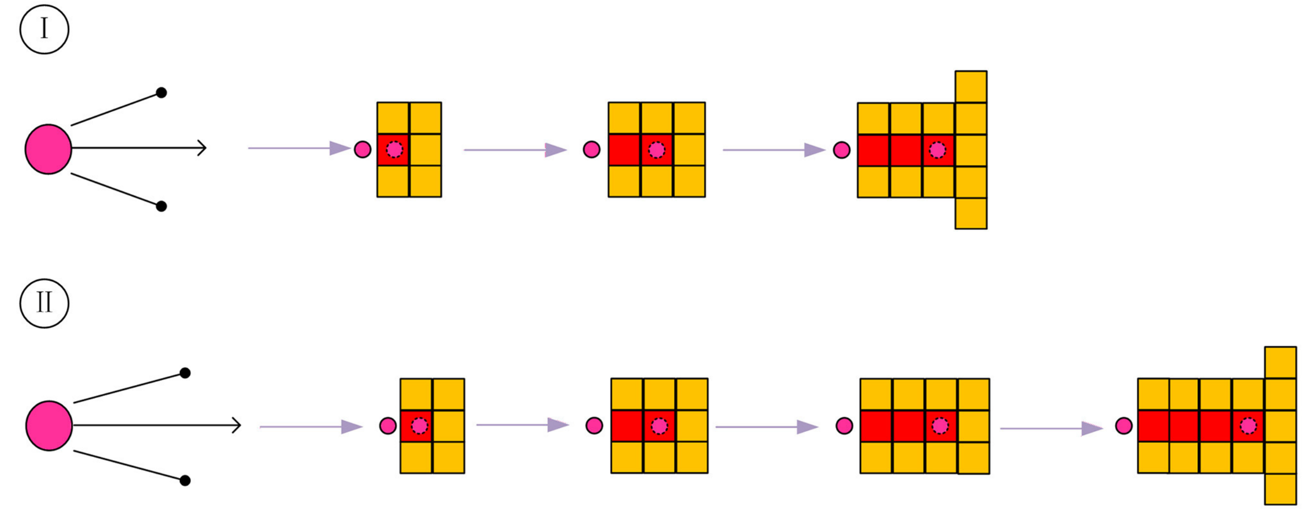


FIGURE 5 | Schematic illustration of step-by-step occupancy grids prediction for dynamic obstacles with different angular velocities.

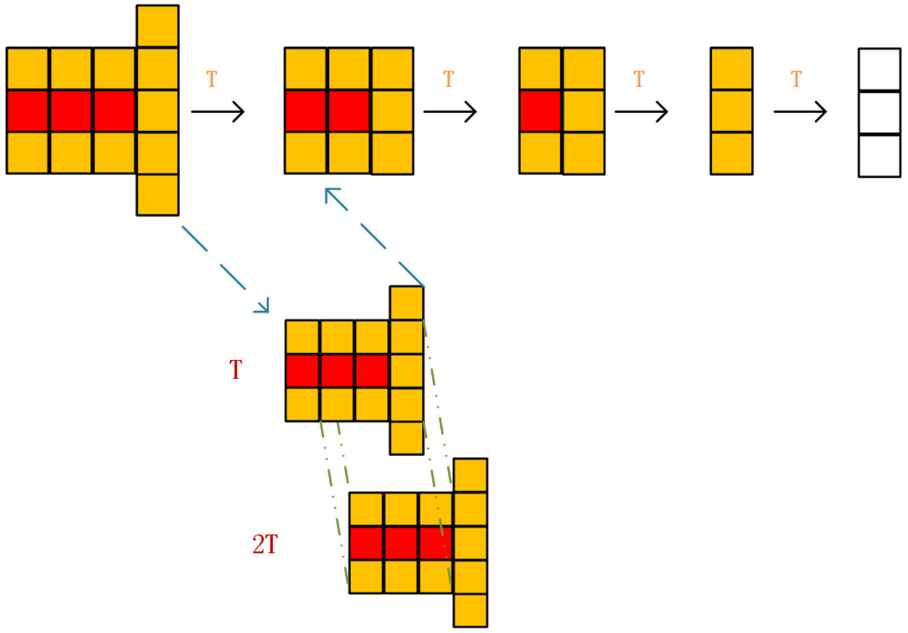


FIGURE 6 | Schematic diagram of misplaced overlap elimination of occupancy grids.

detects a dynamic obstacle, it is predicted that the future path of the dynamic obstacle will continue to follow the current speed direction of the dynamic obstacle, as shown in the red grids in **Figure 5** below, and a certain probability of deviation to the original speed direction, as indicated in yellow grids. The longer the path is, the higher the probability of deviating from the red path may be. We set that the predicted footprint of the dynamic obstacle is related to the angular velocity of the movement of the dynamic obstacle. The forecast is shown in **Figure 5** below. v is the current speed of the dynamic obstacle and remains unchanged. The upper part of **Figure 5** shows the step-by-step occupancy grid prediction map when the angular velocity ω of dynamic obstacle movement is large, and the lower part of **Figure 5** shows the step-by-step occupancy grid prediction map when the angular velocity ω of dynamic obstacle movement is small. The solid circle in the figure represents the location of the dynamic obstacle at the current time, and the dotted circle represents the location of the dynamic obstacle at the future time.

According to the original path and the predicted path of the dynamic obstacle, it should be judged whether it is necessary

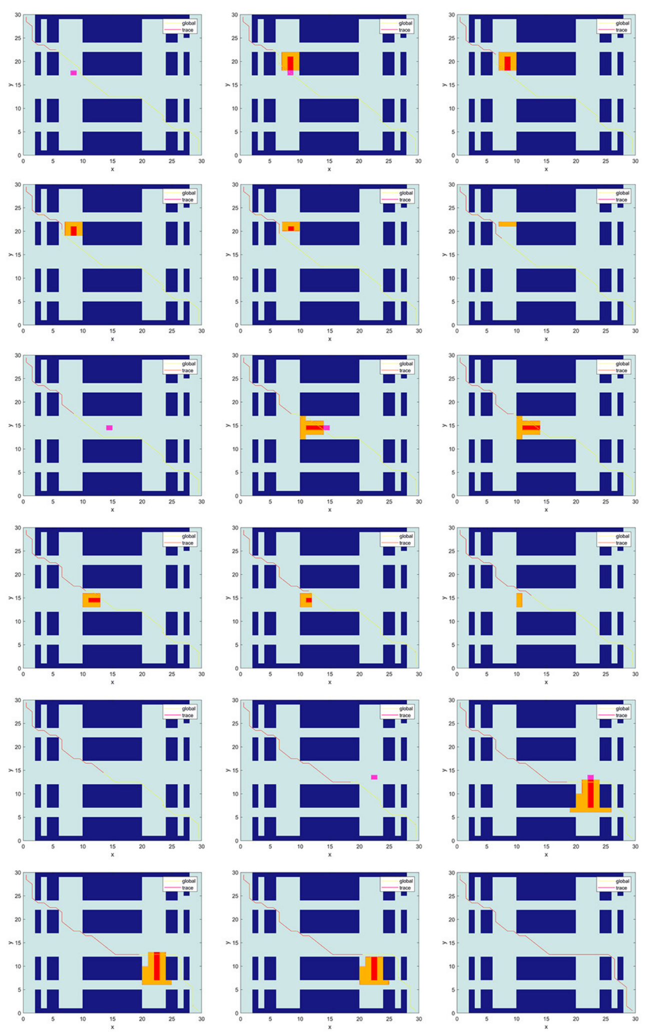


FIGURE 7 | Dynamic obstacle avoidance simulation in Environment 1.

TABLE 2 | Description of TTACO for solving path planning.

```
Algorithm TTACO
Begin
Create grid environment
Adaptive initial pheromone distribution according to formula (6)
Repeat
for each ant k do
    Trigger the three-step arbitration taboo method based on probability
    Add the grids involved in the Occupancy grid prediction model and three-step arbitration taboo method to $TABU_s$
    if grid i $\epsilon$ $allow_k$ then
        if grid i $\epsilon$ $TABU_{lock}$ then
        Rollback
end if
    According to formula (1) and (5) select next grid j
    Update taboo
    end if
Pheromone preferential limited update for each iteration
Until Meet the iteration end condition
Return best grid serial number
END
```

to call the model, and the time point to call the occupancy grid prediction model should be calculated. Suppose the mobile robot is walking along the original path, when the distance between the mobile robot and the dynamic obstacle is less than the safe distance, the occupancy grid prediction model is called in the first two-unit time when it is less than the safety distance. After the time and location of the occupancy grid prediction model are obtained, the nodes affected by the occupancy grid prediction model are recorded. The formula for the distance between the mobile robot and the dynamic obstacle is as follows, where (x_a, y_a) represents the position of the mobile robot, (x_b, y_b) represents the position of the dynamic obstacle, S represents the safety distance, and let S_0 be the judgment limit for safety distance.

$$S = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \tag{14}$$

Occupancy grid prediction model predicts the distance of dynamic obstacle based on the track length of dynamic obstacle when the distance between mobile robot and dynamic obstacle is less than safety distance. As shown in the following formula (15), $F(s)$ represents the cumulative length of the dynamic obstacle track length S_b when the distance between the mobile robot and dynamic obstacle is less than the safety distance, and the rounded-up value is the predicted distance of the algorithm for dynamic obstacles.

$$F(s) = \sum_{S<S_0} S_b \tag{15}$$

$$S_T = \lceil F(s) \rceil \tag{16}$$

The rules for calling the occupancy grid prediction model are as follows:

1. When the mobile robot detects an unknown dynamic obstacle, it is predicted based on the occupancy grid prediction model to obtain the initial predicted position, time, and occupied grids of the dynamic obstacle.
2. According to the position, time, and occupied grids obtained in the first step, the path nodes that may be affected will be added to the local taboo table. The TTACO will be run, and every time a moment has passed, the path nodes that the dynamic obstacle has passed and are no longer possible to pass are removed from the local taboo table until all the occupancy grids do not exist in the local taboo table.
3. We only consider the possible path nodes during the period of time when the dynamic obstacle has an impact on the mobile robot. Therefore, if the dynamic obstacle deviates from the red predicted line during this process, we repredict and repeat the above processes 1, 2, and 3.

When removing path nodes from the local taboo table, we adopt dislocation overlap elimination. To be specific, we overlap the

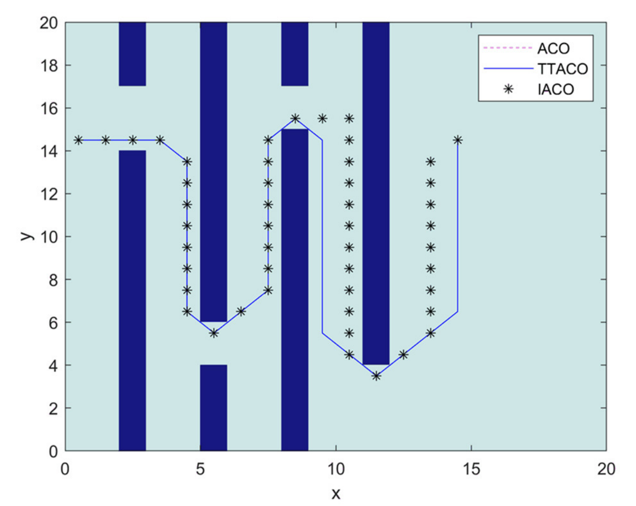


FIGURE 8 | The path planning results of the three algorithms in Environment 2.

TABLE 3 | Comparison of simulation results in static environment.

Environment	Algorithm	Optimal solution of the algorithm	Average shortest distances	Average iteration times
E2②	ACO	–	–	–
	IACO①	50.7280	51.0605	8.4
	TTACO	49.5564	50.6422	6.7
E3	ACO	31.556	–	27
	IACO1	31.556	–	8
	TTACO	30.97	31.4974	3.5
E4	ACO	47	55.1	80
	IACO2	45.69	–	40
	TTACO	44.12	45.219	9
E5	ACO	47	54.5727	82
	JPACSPF	44.5269	–	9
	TTACO	43.9411	44.234	9

①: IACO, IACO1, IACO2, JPACSPF represent the algorithms in Xiaoxu et al. (2018), Dai et al. (2019), Zhang et al. (2019), and Ma and Mei (2020).

②: E2, E3, E4, E5 represent the environment maps in Xiaoxu et al. (2018), Dai et al. (2019), Zhang et al. (2019), and Ma and Mei (2020).

occupancy grid prediction maps of two moments to obtain the occupancy grid map of the next moment. The schematic diagram is shown in **Figure 6** below.

The predictive model is used to predict the possible path nodes of dynamic obstacles in the future moments, study the influence of the possible path nodes of dynamic obstacles on the path of the mobile robot, and give the mobile robot enough time to meet the real-time requirements of the algorithm in the high dynamic environment.

DYNAMIC OBSTACLE AVOIDANCE

Two Cases of Dynamic Obstacle Avoidance

Dynamic obstacle avoidance is divided into two situations: "encounter" and "chase." Such division is based on the angle between the speed direction of the mobile robot and the dynamic obstacle speed. When the angle is $> 90°$, the "encounter" happens, and when it is $\leq 90°$, the "chase" occurs. In both cases, the occupancy grid prediction method is used to avoid obstacles. However, when the mobile robot is in a chasing situation with a dynamic obstacle and as the distance is getting closer and closer, it cannot safely pass through. That is, in the TTACO, when the movement of the dynamic obstacle reaches the last predicted ending position, the ant still has not avoided the dynamic obstacle, and the occupancy grid prediction model is called again. If a better path cannot be obtained after calling the model again or several times, the next-generation ants of the TTACO will choose a safe position according to the occupancy grid prediction model and wait for a period of time when avoiding obstacles.

Ant Colony Information Inheritance Mechanism

When using the occupancy grid prediction model for dynamic obstacle avoidance, in order to improve the calculation speed of the algorithm, we put forward the ant colony information inheritance mechanism. The initial pheromone and taboo table of ACO are determined as follows:

$$Tua_d = Tua + Tua_s \tag{17}$$

$$TABU_d = TABU_s \tag{18}$$

Among them, Tua_d is the initial pheromone when the algorithm is called to avoid dynamic obstacles. Tua is the original ant colony pheromone matrix, and the values in the matrix are all constant C. Tua_s is the pheromone matrix left by the global ACO algorithm after the upper and lower limits are optimized. $TABU_d$ is the global taboo table that calls the ant colony algorithm when the mobile robot avoids dynamic obstacles, and $TABU_s$ is the global taboo table obtained after the global ACO algorithm runs, and it contains all deadlock position information. Through the inheritance mechanism of the above initial pheromone and the global taboo table, the running speed of the algorithm can be accelerated.

Algorithm Flow

The pseudocode of this algorithm is shown in **Table 2**. $TABU_s$ is a local taboo list; $allow_k$ is the list of optional nodes; and $TABU_{lock}$ is the deadlock taboo list. And the computational complexity of the algorithm is O (n ^ 3). The robot first performs global path planning according to the TTACO and then detects whether there are unknown dynamic obstacles and dynamic obstacle information through its sensors. If there are dynamic obstacles, the occupancy grid prediction model will be used to predict based on the information of the detected dynamic obstacles, and the TTACO is called to replan the path after the ant colony pheromone inheritance mechanism is activated.

SIMULATION EXPERIMENT

In order to verify TTACO's algorithm convergence speed, global search ability, and capability to avoid dynamic obstacles, a simulation experiment of dynamic obstacle avoidance was

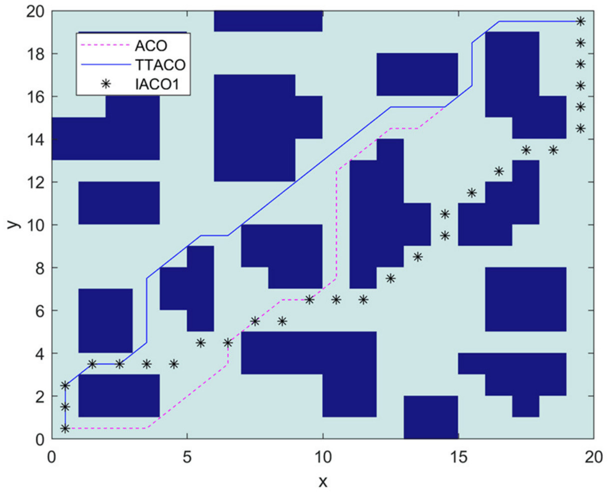


FIGURE 9 | The path planning results of the three algorithms in Environment 3.

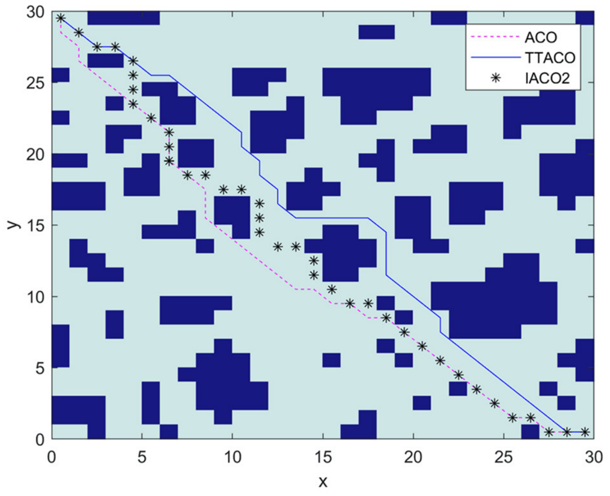


FIGURE 10 | The path planning results of the three algorithms in Environment 4.

carried out in a 30 × 30 simulated factory environment. In order to further verify the performance of the algorithm, it was compared with the algorithms produced by similar articles. The computer performance parameters for the simulation are Intel Core i5-6300HQ processor with a main frequency of 2.30 GHz, a memory size of 8G, the running system of Windows 10, and the simulation software of MATLAB.

Simulation Experiment in Dynamic Environment

The following simulation is to verify the path planning ability and dynamic obstacle avoidance ability of TTACO in dynamic environment. As shown in **Figure 7**, Environment 1 is a simulated 30 × 30 factory environment, with three dynamic obstacles set up along the way. Obstacle 1 is a dynamic obstacle moving upward at a speed slightly faster than the mobile robot; Obstacle 2 is a dynamic obstacle moving to the left with a slightly slower speed; and Obstacle 3 is a dynamic obstacle moving downward whose speed is the same as that of the mobile robot. The golden path in the picture is the current planned global path, and the red path is the trajectory of the mobile robot after it has walked along the path. The pink path node is the position of the dynamic obstacle when calling the occupancy grid prediction model. The red path node is the predicted walking path for dynamic obstacles. The yellow path node is the path node where the dynamic obstacle may deviate and make a turn. When the mobile robot encounters a dynamic obstacle, the occupancy grid prediction model is called, and the path is replanned. From the results, we can see that the mobile robot first moves forward along the preset path; then, it calls the obstacle avoidance algorithm and occupation prediction model three times and replans the path three times. The first two calls of occupancy grid prediction model have successfully enabled the robot to avoid the obstacles, and the last call also achieved the purpose of obstacle avoidance through changing the path.

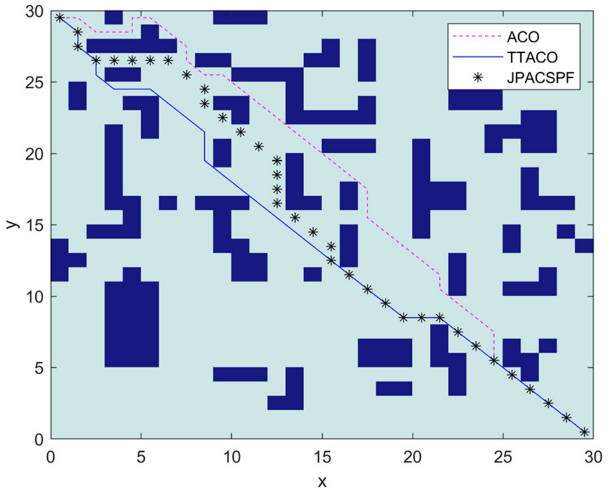


FIGURE 11 | The path planning results of the three algorithms in Environment 5.

Comparison of Similar Algorithms

In order to further verify the performance of the TTACO algorithm, the following four sets of comparative experiments are designed to compare the algorithm in this article with the algorithms from other articles.

Environment 2 is shown in **Figure 8** as the environment map in Dai et al. (2019). The improved ant colony algorithm in the literature uses the characteristics of A * algorithm and maximum–minimum ant system and introduces the retraction mechanism to solve the deadlock problem. IACO is used to represent the algorithm in Dai et al. (2019). It can be seen from **Figure 8** and **Table 3** that both TTACO and IACO successfully find a path; however, the path obtained by TTACO is shorter and converges faster, with the optimal path length being 49.5564 and the average time of iteration 6.7, whereas ACO fails to plan a path.

Environment 3 is shown in **Figure 9** as the environment map in Zhang et al. (2019). In the algorithm, the ant colony is inspired to search the planned path by the improved artificial potential field algorithm. At the same time, the negative feedback channel is constructed by the convergence times, and the size of the environment map is 20 × 20. IACO1 is used to represent the algorithm in Zhang et al. (2019). It can be seen from **Figure 9** and **Table 3** that the path and times of iteration obtained by the algorithm in this article are better than those given by Zhang et al. (2019) and the ACO algorithm. The optimal path length is 30.97, and the average time of iteration is 3.5.

Environment 4 is shown in **Figure 10** as the environment map in Xiaoxu et al. (2018); this article mainly adopts the strategy of death and rollback for the deadlock problem of ant colony, improves the state transition rules of ant colony algorithm, optimizes the composition structure of pheromone, and replaces the algorithm in the literature with IACO2, with the map size of 30 × 30. From **Figure 10** and **Table 3**, we can see that the path obtained by the algorithm in this article is shorter, which is 44.12, and the average time of iteration is nine times, which is better than the algorithm in Xiaoxu et al. (2018) and ACO.

Environment 5 is shown in **Figure 11** as the environment map in Ma and Mei (2020). The reference algorithm is JPACSPF, which combines the search strategy of ant colony algorithm and jump point search algorithm, and introduces the decreasing coefficient of potential field resultant force. The size of environment map is 30 × 30. **Figure 11** shows the paths planned by different algorithms. It can be seen from **Figure 11** and **Table 3** that the path obtained by the TTACO algorithm is shorter than the path obtained by JPACSPF and ACO, and the average time of iteration of TTACO is better than that of the ACO algorithm. The optimal path length is 43.9411, and the average time of iteration is 9.

By comparing with other algorithms, it can be concluded that both the length of the path and the time of iteration in TTACO are better than those in similar literatures, and the algorithm of this article can get a better path in a shorter time than the ACO.

CONCLUSION

This article introduces a novel approach using the TTACO to solve the path planning problem of mobile robots in a dynamic environment. Through the improved adaptive initial pheromone, the death rollback strategy, and the improved pheromone update strategy, the algorithm's convergence speed and global search ability are effectively improved. The three-step arbitration method and occupancy grid prediction model based on the time taboo strategy further improve the algorithm's global search ability and the capability to avoid dynamic obstacles. Through simulation on MATLAB, the experimental results prove that the algorithm can plan a better path in a dynamic environment, so as to realize the navigation of the mobile robot in a dynamic environment. Although the algorithm proposed in this article is novel and has some practical significance, the experimental results need to be further improved, and the details of the article need to be further polished.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

NX proposed time taboo ant colony optimization (TTACO). XZ provided algorithm improvement guidance and general direction control. YX tested and annotated the algorithm. XY carried out MATLAB simulation in different environments based on the algorithm. JM participated in the preparation of the algorithm and the writing of the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. U1933123).

REFERENCES

Boukens, M., Boukabou, A., and Chadli, M. J. A. S. (2019). A real time self-tuning motion controller for mobile robot systems. *IEEE/CAA J. Automat. Sin.* 6, 84–96. doi: 10.1109/JAS.2018.7511216

Dai, X., Long, S., Zhang, Z., and Gong, D. (2019). Mobile robot path planning based on ant colony algorithm with A* heuristic. *Method* 13:15. doi: 10.3389/fnbot.2019.00015

Fan, X., Luo, X., Yi, S., Yang, S., and Zhang, H. (2003). Optimal path planning for mobile robots based on intensified ant colony optimization algorithm. *IEEE Int. Conf. Robot.* 1, 131–136. doi: 10.1109/RISSP.2003.1285562

Ferguson, D., Darms, M., Urmson, C., and Kolski, S. (2008). Detection, prediction, and avoidance of dynamic obstacles in urban environ ments. in 2008 IEEE intelligent vehicles *Symposium* 1, 1149–1154. doi: 10.1109/IVS.2008.4621214

Jiang, M., Wang, F., and Ge, Y. (2019). Research on path planning of mobile robot based on improved ant colony algorithm. *Chin. J. Sci. Instrum.* 40, 113–121. doi: 10.19650/j.cnki.cjsi.J1804429

Keyvan, M., Razeghi-Jahromi, M., Homaifr, A., and Sinica, H. (2020). A stable analytical solution method for car-like robot trajectory tracking and optimization. *IEEE/CAA J. Automat. Sin.* 7, 42–50. doi: 10.1109/JAS.2019.1911816

Khan, A. H., Cao, X., Li, S., and Luo, C. (2020a). Using social behavior of beetles to establish a computational model for operational management. *IEEE Trans. Comput. Soc. Syst.* 7, 492–502. doi: 10.1109/TCSS.2019.2958522

Khan, A. H., Hamza, K., Xinwei, C., and Shuai (2020b). BAS-ADAM: An ADAM based approach to improve the performance of beetle antennae search optimizer. *IEEE/CAA J. Automat. Sin.* 7, 150–160. doi: 10.1109/JAS.2020.1003048

Khan, A. H., Li, S., and Luo, X. (2020c). Obstacle avoidance and tracking control of redundant robotic manipulator: an RNN-based metaheuristic approach. *IEEE Trans. Indus. Inform.* 16, 4670–4680. doi: 10.1109/TII.2019.2941916

Kim, B., Kang, C. M., Kim, J., Lee, S. H., Chung, C. C., and Choi, J. W. (2017). "Probabilistic vehicle trajectory prediction over occupancy grid map via recurrent neural network," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems*. Yokohama: ITSC. doi: 10.1109/ITSC.2017.8317943

Li, W., Tan, M., Wang, L., and Wang, Q. (2019). A cubic spline method combing improved particle swarm optimization for robot path planning in dynamic uncertain environment. *Int. J. Adv. Robot. Syst.* 16:172988141989166. doi: 10.1177/1729881419891661

Li, X., Sun, Z., Cao, D., Liu, D., and He, H. (2017). Development of a new integrated local trajectory planning and tracking control framework for autonomous ground vehicles. *Mech. Syst. Signal Process.* 87, 118–137. doi: 10.1016/j.ymssp.2015.10.021

Liu, K., and Zhang, M. (2016). "Path planning based on simulated annealing ant colony algorithm," in *2016 9th International Symposium on Computational Intelligence and Design*. Hangzhou: ISCID, 461–466. doi: 10.1109/ISCID.2016.2114

Luan, X. L., Gong, F. X., Wei, Z. Q., Yin, B., and Sun, Y. T. (2015). "Using ant colony optimization and cuckoo search in AUV 3D path planning," in *The 2015 International Conference on Software Engineering and Information Technology*. Guilin: SEIT2015, 208–212. doi: 10.1142/9789814740104_0041

Luo, Q., Wang, H., Zheng, Y., and He, J. (2019). Research on path planning of mobile robot based on improved ant colony algorithm. *Neural Comput. Appl.* 32, 1555–1566. doi: 10.1007/s00521-019-04172-2

Ma, X. L., and Mei, H. (2020). Mobile robot global path planning based on improved ant colony system algorithm with potential field. *J. Mech. Eng.* 19–27. doi: 10.3901/JME.2021.01.019

Mohanty, P. K., and Parhi, D. R. (2016). Optimal path planning for a mobile robot using cuckoo search algorithm. *J. Exp. Theor. Artif. Intell.* 28, 35–52. doi: 10.1080/0952813X.2014.971442

Qu, H., and Huang, L. W. (2015). Research of improved ant colony based robot path planning under dynamic environment. *J. Univ. Electr. Sci. Technol. China* 44, 260–265. doi: 10.3969/j.issn.1001-0548.2015.02.017

Rimon, E., and Koditschek, D. E. (1992). Exact robot navigation using artificial potential functions. *IEEE Trans. Robot. Autom.* 8, 501–518. doi: 10.1109/70.163777

Riosmartinez, J., Escobedo, A., Spalanzani, A., and Laugier, C. (2013). "Intention Driven Human Aware Navigation for Assisted Mobility," in *Workshop on Assistance & Service Robotics in A Human Environment at Iros* (Tokyo).

Rummelhard, L., Nègre, A., Perrollaz, M., and Laugier, C. (2014). *Probabilistic Grid-based Collision Risk Prediction for Driving Application.* San Francisco, CA: ISER. doi: 10.1007/978-3-319-23778-7_54

Xiaoxu, W., Guoliang, W., Yan, S., and Xuegang, H. J. (2018). Improved ACO-based path planning with rollback and death strategies. *Syst. Sci. Cont. Eng.* 6, 102–107. doi: 10.1080/21642583.2018.1471426

Xu, K. B., Lu, H. Y., and Huang, Y. (2019). Robot path planning based on double-layer ant colony optimization algorithm and dynamic environment. *Acta Electr. Sin.* 47, 2166–2176. doi: 10.3969/j.issn.0372-2112.2019.10.019

Zhang, Q., Cheng, B. K., Liu, X. Y., Liu, X. Y., and Machinery, H. Y. (2019). "Ant colony optimization with improved potential field heuristic for robot path planning". *in 2018 37th Chinese Control Conference (CCC)* (Wuhan: Transactions of the Chinese Society for Agricultural Machinery), 5317–5321. doi: 10.6041/j.issn.1000-1298.2019.05.003

Zhao, Q., Chen, Y., Luo, B., and Zhang, L. P. (2020). A local path planning algorithm integrating pedestrian forecast information. *Trans. Chin. Soc. Agric. Mach.* 50, 30–39. doi: 10.13203/j.whugis20200105

Zhou, Z., Nie, Y., and Min, G. (2013). "Enhanced ant colony optimization algorithm for global path planning of mobile robots," in *Fifth International Conference on Computational & Information Sciences* (Shiyan), 698–701. doi: 10.1109/ICCIS.2013.189

Conflict of Interest: XY was employed by the company "Civil Aviation Logistics Technology Company Limited."

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurorobotics

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fnbot.2021.672582

Exoskeleton Active Walking Assistance Control Framework Based on Frequency Adaptive Dynamics Movement Primitives

Shiyin Qiu[1], Wei Guo[1], Fusheng Zha[1,2]*, Jing Deng[2] and Xin Wang[2]*

[1] State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin, China, [2] Robotics Institute, Shenzhen Academy of Aerospace Technology, Shenzhen, China

This paper introduces a novel exoskeleton active walking assistance control framework based on frequency adaptive dynamics movement primitives (FADMPs). The FADMPs proposed in this paper is an online learning and prediction algorithm which is able to online estimate the fundamental frequency of human joint trajectory, learn the shape of joint trajectory and predict the future joint trajectory during walking. The proposed active walking assistance control framework based on FADMPs is a model-based controller which relies on the human joint torque estimation. The assistance torque provided by exoskeleton is estimated by human lower limb inverse dynamics model which is sensitive to the noise in the joint motion trajectory. To estimate a smooth joint torque profile, the joint motion trajectory must be filtered first by a lowpass filter. However, lowpass filter will introduce an inevitable phase delay in the filtered trajectory. Both simulations and experiments in this paper show that the phase delay has a significant effect on the performance of exoskeleton active assistance. The active assistant control framework based on FADMPs aims at improving the performance of active assistance control by compensating the phase delay. Both simulations and experiments on active walking assistance control show that the performance of active assistance control can be further improved when the phase delay in the filtered trajectory is compensated by FADMPs.

Keywords: exoskeleton, DMPs, active walking assistance, frequency estimation, motion learning, motion prediction

Edited by:
Yan Wu,
Institute for Infocomm Research (A*STAR), Singapore

Reviewed by:
Yangwei You,
Institute for Infocomm Research (A*STAR), Singapore
Zhenshan Bing,
Technical University of Munich, Germany

***Correspondence:**
Fusheng Zha
zhafusheng@hit.edu.cn
Xin Wang
xin.wang@chinasaat.com

Received: 26 February 2021
Accepted: 21 April 2021
Published: 20 May 2021

Citation:
Qiu S, Guo W, Zha F, Deng J and Wang X (2021) Exoskeleton Active Walking Assistance Control Framework Based on Frequency Adaptive Dynamics Movement Primitives.
Front. Neurorobot. 15:672582.
doi: 10.3389/fnbot.2021.672582

1. INTRODUCTION

Since the 1960s, the United States General Electric Company launched the world's first full-body exoskeleton robot Hardiman (Mosher, 1967), exoskeleton robot has gradually became a hot research direction of robotics. Exoskeleton robot is a typical man-machine coordinated control system. The core of this system is the human who provides intelligent decision for the whole system. Exoskeleton provides power assistance for the human body. By combining human intelligence with the powerful assistance of the exoskeleton, the exoskeleton can accomplish some tasks that cannot be completed by the conventional robots, such as individual combat, disaster relief, and rehabilitation.

After entering the 20th century, the progress of science and technology has promoted the rapid development of exoskeleton, and many research achievements have emerged

(Kalita et al., 2020). At present, according to the different application and users, exoskeletons can be roughly divided into three categories: load carrying exoskeleton (Fontana et al., 2014), motion assistance exoskeleton (Witte et al., 2020), and rehabilitation exoskeleton (Jamwal et al., 2020). The main control target and strategies for the three kinds of exoskeleton are different (Kalita et al., 2020). The control target of load carrying exoskeleton is to offset loads and follow the human motion. In order to achieve these goals, the dynamics of loads and exoskeletons should be totally or partially compensated by exoskeleton and the interaction force/torque between human and exoskeleton should be controlled as small as possible. Sensitivity amplification control (SAC) (Kazerooni et al., 2005; Huang et al., 2018, 2019) is a typical control strategy for load carrying exoskeleton. For rehabilitation exoskeleton, the main control target is to drive the patient's paralyzed limb to follow a predefined trajectory for rehabilitation purpose. Both of the dynamics of exoskeleton and patient's limb should be fully compensated by the exoskeleton. The predefined gait trajectory control is a common control strategy for rehabilitation exoskeleton (Quintero et al., 2011; Esquenazi et al., 2012; Lu et al., 2014). Being different from the load carrying exoskeleton and rehabilitation exoskeleton, the motion assistance exoskeleton aims at reducing the user's muscle activity and metabolic cost during motion. Therefore, active assistance control strategy should be used to completely compensate the dynamics of exoskeleton and partially compensate the dynamics of human limb.

The main difficulty in active assistance control is how to provide positive work to the user while ensuring the initiative of user. Hence, estimating the user's motion intention is the first important step of active assistance control. The most direct way to detect the user's intention is to measure the biosignals, such as electromyogram (EMG) (Zeng et al., 2020), electroencephalogram (EEG) (Ortiz et al., 2020), and muscle stiffness (Chao et al., 2018). And based on these biosignals some active assistance control strategies have been proposed including proportional EMG based control (Young et al., 2017; Lorenzo et al., 2018), EEG based control (Al-Quraishi et al., 2018), and muscle stiffness control (Kim et al., 2013). However, due to the high sensitivity of the electrode position, muscle fatigue, sweat, and the deformation of skin, the biosignal based active assistance control strategy is not widely used in the real application. To avoid the above drawbacks of biosignals, some non-biosignal based active assistance control strategies have been proposed, such as gait phase based control (Asbeck et al., 2015), integral admittance shaping control (Nagarajan et al., 2016), motor primitive based control (Ruiz Garate et al., 2017), hybrid assisted control (Oh et al., 2015), admittance control (Liang and Hsiao, 2020), and adaptive oscillator based control (Seo et al., 2015). Integral admittance shaping control and motor primitive based control are model-based control strategies relying on the human-exoskeleton interaction model and musculoskeletal model, respectively. These two control strategies are highly depending on the model parameters which should be precisely estimated for the different user and exoskeleton. The complicated parameter identification process make these control strategies inconvenient to be used in daily life. Therefore, more and more researches are trying to find a kind of active assistance controller that is capable for self-adaption and self-learning (Young and Ferris, 2017).

Adaptive oscillator (AO) based control (Ronsse et al., 2010) is a promising active assistance control framework which can online learn and adapt to the features (frequency components, amplitudes and phases) of a periodic joint trajectory during walking, and provide effective assistance torque to the user's limb joints without a user-specific calibration (Ronsse et al., 2011). The basic ideal of AO is establishing a pool of adaptive oscillators to learn the fundamental frequency of input signal and then constructing a supervised learning problem to learn the profile of input signal (Ronsse et al., 2011). But there are two drawbacks existed in AO: (1) the initial frequency of each oscillator will determine whether the AO can convergent to the fundamental frequency of input signal (Seo et al., 2018); (2) if the amplitude of input signal changes, AO will take a significant amount of time to converge to the new amplitude (Chinimilli et al., 2019). Therefore, to enhance the convergence of AO, a particularly-shaped adaptive oscillator (PSAO) (Seo et al., 2018) was proposed to make AO less dependent on the initial parameters. And to make AO converge rapidly to the varying amplitude of input signal, an amplitude omega adaptive oscillator (AωAO) was proposed in Chinimilli et al. (2019).

In this paper, to avoid the drawbacks of AO, a novel frequency adaptive dynamics movement primitives (FADMPs) and an active assistance control framework based on FADMPs are proposed. FADMPs is able to online predict a smooth joint trajectory without phase delay, which is benefit for improving the performance of active walking assistance of exoskeleton. There are three advantages in the proposed control framework, which are follows:

1. the frequency of human limb joint trajectory can be precisely online estimated and adapted by FADMPs algorithm, the initial parameters of FADMPs and the sudden change of walking frequency and motion amplitude have no effect on the results of frequency estimation;
2. the profile of human limb joint trajectory can be online learned and predicted. And the phase compensation has little effect on the profile of the predicted trajectory of FADMPs;
3. the active walking assistance control framework based on FADMPs is suitable for both stable and unstable gaits. It is able to automatically choose active assistance mode based on walking frequency. When human walks in a stable frequency the exoskeleton will work in an assistance mode. And if walking frequency is unstable exoskeleton will work in a transparent mode.

The rest of this paper is organized as follows: the related works on human motion trajectory online learning and prediction are discussed in section 2. The derivation FADMPs and the active assistance control framework based on FADMPs are proposed in section 3. The simulation results of the proposed control framework are shown in section 4. The experiment results of active assistance control are discussed in section 5. Section 6 concludes this paper.

2. RELATED WORK

2.1. Adaptive Frequency Oscillator (AFO)

AFO were invented in Righetti et al. (2006) using some oscillators in parallel to learn the frequency component of the input signal. The learning process of AFO is a kind of real-time Fourier decomposition. The dynamics of AFO is given by the following equations:

$$\begin{cases} \dot{\phi}_i(t) = \omega_i(t) - \upsilon e(t) sin\phi_i(t) \\ \dot{\omega}_i(t) = -\upsilon e(t) sin\phi_i(t) \\ \dot{\alpha}_i(t) = \eta cos\phi_i(t) e(t) \\ e(t) = \theta(t) - \hat{\theta}(t) \\ \hat{\theta}(t) = \sum_{i=0}^{K} \alpha_i(t) cos\phi_i(t) \end{cases} \tag{1}$$

where ϕ_i is the phase of oscillator i, ω_i is the frequency of each oscillator, K is the total number of oscillators, α_i is the amplitude of each oscillator, η is the integrator gain, υ determines the speed of phase synchronization to the input signal $\theta(t)$, $\hat{\theta}(t)$ is the weighted sum of oscillators.

The basic idea of AFO is using a feedback structure to make each oscillator converge to the frequency components of the input signal. However, the initial frequency $\omega_i(0)$, amplitude $\alpha_i(0)$, and phase $\phi_i(0)$ of each oscillator have significant effect on the convergence of oscillator. If the initial parameters of each oscillator are set inappropriate, the frequency of oscillator may not converge to the frequency components of the input signal and the frequency of oscillator may even become negative (Gams et al., 2009).

2.2. Adaptive Oscillator (AO)

To improve the convergence of AFO, AO was proposed in Ronsse et al. (2011) based on the assumption that the input signal is periodic. Hence, AO only learns the fundamental frequency of the input signal. The dynamics of AO can be given by the following equations:

$$\begin{cases} \dot{\phi}_i(t) = i\omega(t) + \upsilon e(t) cos\phi_i(t) \\ \dot{\omega}(t) = \upsilon e(t) cos\phi_1(t) \\ \dot{\alpha}_i(t) = \eta sin\phi_i(t) e(t) \\ e(t) = \theta(t) - \hat{\theta}(t) \\ \hat{\theta}(t) = \sum_{i=0}^{K} \alpha_i(t) sin\phi_i(t) \end{cases} \tag{2}$$

Comparing (2) and (1) we can find that the main difference between AO and AFO is that all oscillators share the same fundamental frequency ω. Hence, only the initial value of the fundamental oscillator $[\omega(0)]$ needs to be set. However, the initial value of the phase and amplitude of each oscillator $[\phi_i(0), \alpha_i(0)]$ can still affect the convergence of AO. To make AO less depend on the initial values, a particularly-shaped adaptive oscillator (PSAO) was proposed in Seo et al. (2018). Being different from AO, the basic function of PSAO was established by a nominal pattern function of input signal, therefore with the guidance of nominal pattern function the oscillators can converge more quickly to the fundamental frequency of input signal. But the nominal pattern needs to be recorded by walking experiments in advance, and PSAO needs walking pattern classification algorithm to choose right nominal pattern function. To improve the convergence speed of AO when the walking amplitude suddenly changed, an amplitude omega adaptive oscillator ($A\omega AO$) was proposed in Chinimilli et al. (2019). $A\omega AO$ algorithm firstly calculates the amplitude and frequency of human motion, and then uses support vector machine (SVM) and discrete hidden Markov model (DHMM) to determine whether the movement pattern of the human body has changed. When the change of human motion pattern is detected, the AO algorithm can converge to the amplitude and frequency of the new motion pattern faster by reinitializing the AO parameters. However, the accuracy of human motion pattern recognition based on SVM and DHMM is 95.2%, so it is still possible that the AO parameters cannot be reinitialized correctly because of the motion pattern recognition error. Although PSAO and $A\omega AO$ algorithm can reduce the dependence on the initial setting of the oscillator, they still cannot completely get rid of the influence of the initial value of the algorithm and the sudden change of human motion pattern.

In order to avoid the problems existing in AO algorithm, this paper proposes a frequency adaptive dynamics movement primitive (FADMPs) algorithm. Instead of using multiple oscillators to learn the input signal, FADMPs learns and predicts the input signal based on dynamics movement primitives (DMPs). FADMPs can completely avoid the convergence failure caused by the improper setting of the initial parameters. In addition, FADMPS algorithm can quickly converge to the new amplitude, frequency and phase of the input signal when the human motion changes. And FADMPS can also conveniently change the phase of the output trajectory to realize a real-time prediction of a smooth human motion trajectory.

3. METHODOLOGY

In this section, FADMPs is introduced in section 3.1. Then, the control framework based on FADMPs is established for active walking assistance in section 3.2.

3.1. Frequency Adaptive Dynamics Movement Primitives

FADMPs is an online learning and prediction algorithm which can be used to learn and predict periodic signal in real time. FADMPs algorithm includes three parts: trajectory frequency estimation, trajectory learning and prediction. In this paper, we use zero crossing detection method to estimate the frequency of input trajectory first, and then we use dynamics movement primitives (DMPs) (Ijspeert et al., 2013) to learn the input trajectory and predict the future trajectory.

Frequency estimation is the first step of FADMPs algorithm. The human lower limb joint trajectory during stable walking is a non-sinusoidal periodic signal. Hence, the frequency of human joint trajectory can be estimated by using zero crossing detection method. In order to estimate the trajectory frequency, the time between two zero-crossing point should be recorded. In this paper, the zero-crossing point is defined as a time stamp $[t(n)]$ which should meet the condition: $\theta(t(n)) \geq 0$ and $\theta(t(n-1)) < 0$, where θ denotes the joint angle and n denotes the time step. The time between two zero-crossing point denotes as the period of trajectory (T_m) and the frequency of input trajectory can be estimated by (3), where m denotes the sequence number of zero-crossing points.

$$F_m = 1/T_m \tag{3}$$

When the frequency of input trajectory has been estimated, the shape of input trajectory can be online learned and predicted by using FADMPs algorithm. The fundamental learning mechanism of FADMPs is to use Gaussian-like kernel function as building blocks to establish a non-linear forcing term to make the output of a globally stable second-order linear system converge to the input trajectory. The globally stable second-order linear system is chosen as a damped spring model shown in (4).

$$\begin{cases} \dot{z} = \Omega(\alpha_z(\beta_z(g-y)-z)+f) \\ \dot{y} = \Omega z \\ \ddot{y} = \Omega \dot{z} \\ \Omega = 2\pi F_m \end{cases} \tag{4}$$

where y is the input trajectory. g is an oscillation baseline of the learning trajectory and it is set as $g = 0$ in this paper. α_z and β_z are positive constants, and in order to make the system stable β_z should be set as $\beta_z = \alpha_z/4$. In this paper α_z is set as 25. f is the non-linear forcing term given by (6). ϕ is the phase of Canonical Dynamical System which makes the model (4) depend on phase (ϕ) rather than time (t) (Gams et al., 2009). Ω is the frequency of Canonical Dynamical System. F_m given by (1) is the estimated frequency of input trajectory.

$$\dot{\phi} = \Omega, \quad \phi \in [0, 2\pi] \tag{5}$$

$$f = \frac{\sum_{i=1}^{N} \Psi_i \omega_i}{\sum_{i=1}^{N} \Psi_i} \tag{6}$$

Being different from the traditional DMPs algorithm (Ijspeert et al., 2013), our FADMPs algorithm is not only able to online learn the input trajectory, but also online predict the future trajectory. When controlling exoskeleton, the time-delay or phase-delay caused by the computation time, sensing time, communication delay and filter will make exoskeleton lag behind the human intention. To reduce the time-delay, the user's future motion should be predicted and the control command of exoskeleton should be send before user's motion (Ding et al., 2020). In order to online predict the future trajectory, two sets of Gaussian-like kernel functions are needed. As shown in (7), one set (Ψ_i) is for learning and the other set (Ψ_{pi}) is for prediction. Prediction also means phase lead in the future, as shown in (7) and (8), Δ_ϕ represents a phase lead and $\Psi_{pi}(i = 1 \cdots N)$ denotes a set of Gaussian-like kernel functions with phase lead, here N is the total number of Gaussian-like kernel functions. In this paper N is set as 50.

$$\begin{cases} \Psi_i = exp(h_i(cos(\phi - c_i) - 1)) \\ \Psi_{pi} = exp(h_i(cos(\phi + \delta_\phi - c_i) - 1)) \end{cases} \tag{7}$$

$$\dot{\delta_\phi} = \alpha_\phi(\Delta_\phi - \delta_\phi) \tag{8}$$

where $h_i(i = 1 \cdots N)$ is the width of Gaussian-like kernel function and it is set as $h = 2.5N$ in this paper. $c_i(i = 1 \cdots N)$ is the center of each Gaussian-like kernel function and c_i is evenly distributed over the range $[0, 2\pi]$. δ_ϕ is the state variable of Δ_ϕ. To make the predicted trajectory smoother, the sudden change of the phase of Canonical Dynamical System should be avoided. Hence, a simple first-order differential equation given by (8) is utilized in this paper to filter the discontinuous change of the goal phase (Δ_ϕ). α_ϕ decides the phase changing speed and it is set as $\alpha_\phi = \alpha_z/2$ in this paper.

But before prediction the input trajectory needs to be learned first by using $\Psi_i(i = 1 \cdots N)$, which is a set of Gaussian-like kernel functions without phase lead. To determine the weight $\omega_i(i = 1 \cdots N)$ of Gaussian-like kernel functions, a recursive least squares algorithm with a forgetting factor of λ is adopted in this paper.

In order to explain the recursive least squares algorithm more clearly, we firstly assume the input trajectory is $[y_d(t), \dot{y}_d(t), \ddot{y}_d(t)]$. And then, according to (4), we can get the target forcing term by (9).

$$f = \ddot{y}_d/\Omega^2 - \alpha_z(\beta_z(g - y_d) - \dot{y}_d/\Omega) \tag{9}$$

And then putting (9) into (6), we can establish a supervised learning problem to determine $\omega_i(t)$ at each time step by using recursive least squares algorithm (Gams et al., 2009):

$$\begin{cases} \omega_i(t+1) = \omega_i(t) + \Psi_i P_i(t+1)(f - \omega_i(t)) \\ P_i(t+1) = P_i(t)/(\lambda + P_i(t)\Psi_i) \end{cases} \tag{10}$$

where $P_i(i = 1 \cdots N)$ denotes an inverse covariance matrix (Kumar, 1985). The initial conditions of this recursion algorithm are $\omega_i(0) = 0$ and $P_i(0) = 1$. The forgetting factor is chosen as $\lambda = 0.95$ in this paper.

When ω_i is determined, the input trajectory has been learned successfully. And then the future trajectory ($[\hat{y}_d(t), \hat{\dot{y}}_d(t), \hat{\ddot{y}}_d(t)]$) can be predicted by (4) with a predicted forcing term $\hat{f}$ which is given by (11).

$$\hat{f} = \frac{\sum_{i=1}^{N} \Psi_{pi} \omega_i}{\sum_{i=1}^{N} \Psi_{pi}} \tag{11}$$

Insert (11) into (4) we can get the predicted trajectory:

$$\begin{cases} \hat{\dot{z}} = \Omega(\alpha_z(\beta_z(g - \hat{y}) - \hat{z}) + \hat{f}) \\ \hat{\dot{y}} = \Omega \hat{z} \\ \hat{\ddot{y}} = \Omega \hat{\dot{z}} \end{cases} \tag{12}$$

In summary, the first step of FADMPs is to estimate the frequency (F_m) of input trajectory by using zero crossing detection algorithm. The second step of FADMPs is to learn the shape of input trajectory by using recursive least squares algorithm which includes (7)–(10). The final step of FADMPs is to predict the future trajectory by (11) and (12) according to a given phase lead Δ_ϕ which is a positive value defined by user. And it should be noticed that if $\Delta_\phi = 0$ the predicted trajectory will keep a same phase with the input trajectory. However, if $\Delta_\phi < 0$ the predicted trajectory will have a phase delay compared with the input trajectory. Hence, the most important feature of our FADMPs algorithm is that it can online arbitrarily adjust the phase of predicted trajectory by changing Δ_ϕ. And the shape of predicted trajectory will remain almost the same as the input trajectory when Δ_ϕ is changed.

3.2. Exoskeleton Active Walking Assistance Control Framework Based on FADMPs

Active walking assistance needs exoskeleton offer assistance force or torque on human body to reduce the metabolic cost of human body during walking. To reach this goal, the exoskeleton assistance torque acting on the human body must coincide with the human walking intention. Human walking intention can be estimated by the human joint torque (Li et al., 2018), which shows the strength and direction of human motion. However, the estimation result of human joint torque is sensitive to the signal noise. Hence, in a real control system, the measured human joint trajectory must be filtered by a low-pass filter before it is sent to the human inverse dynamics model to estimate joint torque. But the filtered signal will have an unavoidable phase delay compared with the original signal. The phase delay will lead to the conflict between exoskeleton and human intention. To compensate the phase delay, FADMPs algorithm is applied to online compensate the phase delay caused by the low-pass filter. As mentioned above, FADMPs algorithm is able to online change the phase of the predicted trajectory by adjusting Δ_ϕ. Hence, we can chose a proper Δ_ϕ to compensate phase delay according to the frequency response of the low-pass filter. In this paper, an exoskeleton active walking assistance control framework based on FADMPs is proposed to realize a walking assistance without phase delay.

As shown in **Figure 1**, there are mainly three parts in the exoskeleton active walking assistance control framework. The first part is the exoskeleton wearer who is the center of human-exoskeleton system and responsible for decision making and motion control. Under the stimulation of electroneurographic signal generated from human central nervous system, human muscle-skeleton system will generate limb joint actuation torque τ_h to generate joint motion trajectory $\boldsymbol{\theta_h}$. The second part of our control framework is human intention estimation which is the most important part of active assistance control. FADMPs algorithm is applied in this part to online learn the joint trajectory and compensate the phase delay caused by low-pass filter. And then, a torque estimator (Li et al., 2018) based on human inverse dynamics is used to estimate the human joint torque $\hat{\tau}_h$ during walking. The third part of our control framework is the low layer exoskeleton assistance torque feedback control system. The input of this system is the estimated human joint torque $\hat{\tau}_h$ multiplied by assistance ratio (AR) α. The feedback of this system is the human-exoskeleton interaction torque τ_{he} measured by load cell. The input of exoskeleton actuator τ_e is given by (13). The assistance torque feedback PI controller is (14). τ_c is the exoskeleton dynamics compensation torque which is used to compensate the inertial of exoskeleton. The mass and inertial of exoskeleton are denoted as m and J, respectively. The distance between the mass center of the output rod of exoskeleton and it's center of rotation is denoted as l. The above inertial and structure parameters of exoskeleton are obtained from SolidWorks (Dassault Systemes, USA).

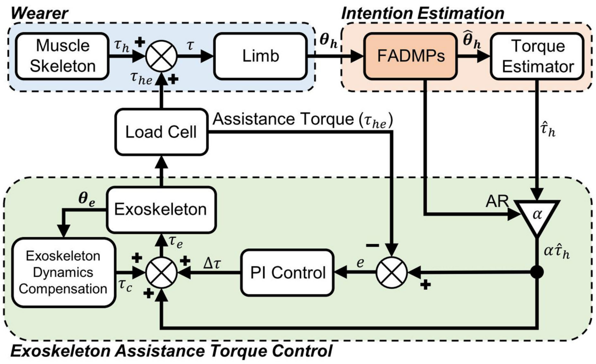


FIGURE 1 | Exoskeleton active walking assistance control framework based on FADMPs. τ_h is the human joint torque generated by muscle-skeleton system. τ is the total torque acting on the human joint. $\boldsymbol{\theta_h}$ is the actual measured and filtered human joint trajectory. $\hat{\boldsymbol{\theta}}_h$ is the predicted trajectory given by FADMPs algorithm. $\hat{\tau}_h$ is the estimated human joint torque. α is an assistance ratio (AR) which determines the strength of exoskeleton assistance. τ_{he} is the feedback of assistance torque. $\boldsymbol{\theta_e}$ is the output trajectory of exoskeleton. τ_c is the exoskeleton dynamics compensation torque. τ_e is the total input torque of exoskeleton.

$$\tau_e = \alpha\hat{\tau}_h + \tau_c + \Delta\tau \tag{13}$$

$$\begin{cases} \Delta\tau = K_P e(t) + K_I \displaystyle\int e(t)dt \\ e(t) = \alpha\hat{\tau}_h - \tau_{he} \end{cases} \tag{14}$$

$$\tau_c = mglsin(\theta_e) + J\ddot{\theta}_e \tag{15}$$

The step frequency is not always consistent during actual walking. It can be changed due to the environment or disturbance. And people may adjust walking speed at any time. In this paper, the step frequency is estimated by zero crossing detection method which only update the frequency when the limb joint trajectory pass through zero. Hence, there is a time-delay in frequency estimation when step frequency suddenly changes. As as shown in **Figure 2C**, the initial walking frequency is 1.5 Hz and it begins to decrease at 1 s. And it finally goes down to 1 Hz at 3 s. It is obvious that there is a time-delay exists in

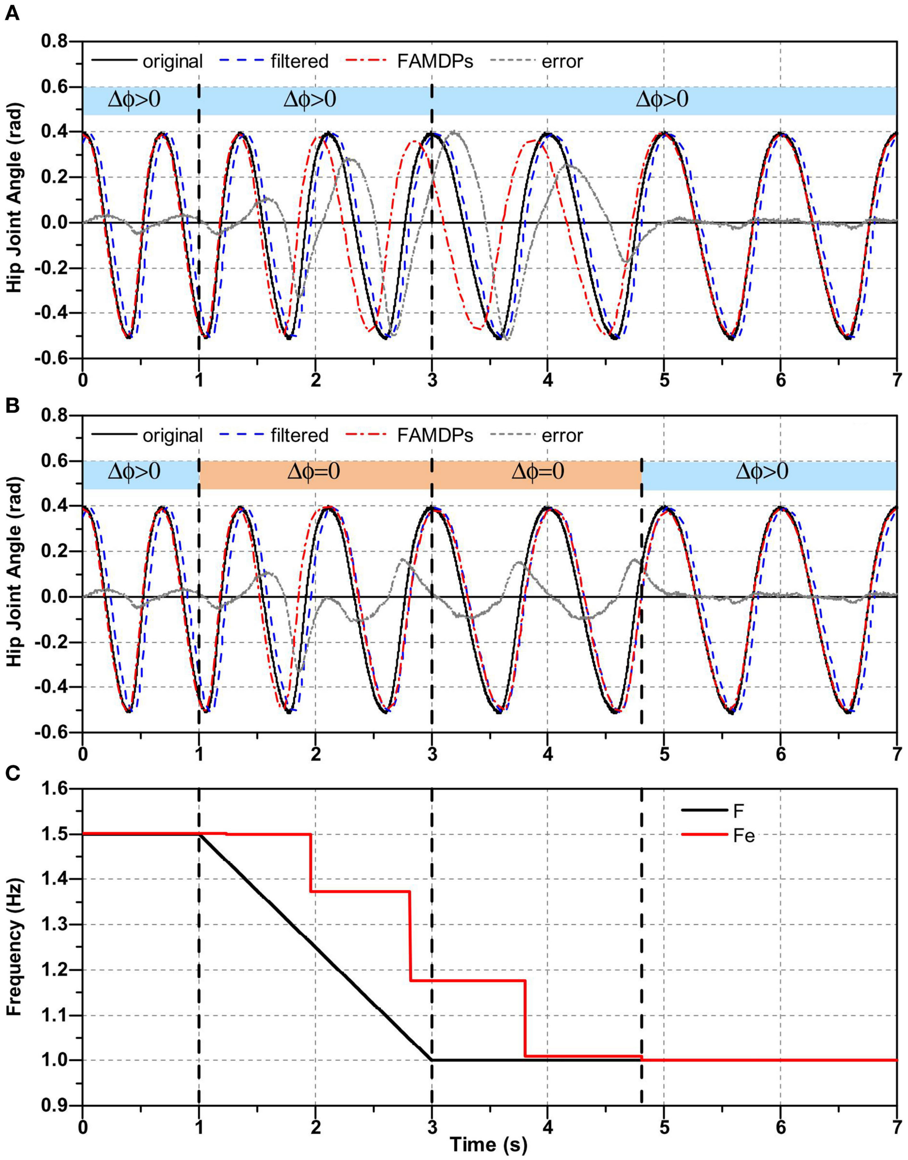


FIGURE 2 | Time-delay of frequency estimation and its effect on the trajectory prediction. In **(A,B)**, black line denotes the original trajectory containing white noise. The blue dash line denotes the filtered trajectory. The red dash dot line represents the predicted trajectory of FADMPs. The gray short dash line represents the error between the original trajectory and the predicted trajectory. In **(C)**, black solid line denotes the frequency of original trajectory and the red solid line represents the estimated frequency.

the result of frequency estimation. And as shown in **Figure 2A**, the time-delay in frequency estimation leads to a relatively large prediction error (maximum absolute error is about 0.52 rad, root mean square error (RMSE) is about 0.17 rad) of FADMPs when the walking frequency is changing.

Therefore, in this paper, we regulate that the phase compensation should be executed only when the step frequency is stable. If the step frequency error between two steps is <0.1 Hz ($|F_m - F_{m-1}| <= 0.1Hz$) the human walking is considered to be stable and the phase compensation of FADMPs is executed ($\Delta_\phi > 0$). On the contrary, if human walking is unstable FADMPs will not compensate the phase delay ($\Delta_\phi = 0$). Furthermore, if walking is unstable the AR will be zero ($\alpha = 0$) which means the exoskeleton is working in transparent mode (Qiu et al., 2020). And if walking is stable the AR will be set as a positive value ($0 < \alpha < 1$) which means exoskeleton will work on active assistance mode (Qiu et al., 2020) and provide an assistance torque on human joint during walking. The above regulation can be summarized as the rule (16).

$$|F_m - F_{m-1}| \begin{cases} <= 0.1 : \Delta_\phi > 0, 0 < \alpha < 1 \\ > 0.1 : \Delta_\phi = 0, \alpha = 0 \end{cases} \qquad (16)$$

TABLE 1 | The initial parameters of FADMPs and AO.

Algorithm	Initial parameters
FADMPs	$N = 50$, $h = 2.5N$, $\alpha_z = 25$, $g = 0$, $\dot{z}(0) = 0$, $\dot{y}(0) = 0$, $\ddot{y}(0) = 0$
AO	$N = 50$, $h = 2.5N$, $K = 6$, $\upsilon = 6$, $\eta = 0.25$, $\alpha_i(0) = 0$, $\omega(0) = 2\pi$, $\phi_0(0) = 0$
AO	$N = 50$, $h = 2.5N$, $K = 6$, $\upsilon = 6$, $\eta = 0.25$, $\alpha_i(0) = 0$, $\omega(0) = \pi$, $\phi_0(0) = \pi/2$

Figure 2B shows the trajectory prediction results of FADMPs based on the rule (16). It is obvious that, compared with **Figure 2A**, the prediction error is significantly reduced (maximum absolute error is about 0.33 rad, RMSE is about 0.073 rad) when the walking frequency changes. The reason for this phenomenon is that FADMPs no longer changes the phase of predicted trajectory ($\Delta_\phi = 0$) when the walking frequency changes. And for this reason, the predicted trajectory of FADMPs (red solid dot line) will coincide with the filtered trajectory (blue dash line). Therefore, the trajectory error between original trajectory and the predicted trajectory can still be reduced by executing the regulation (16) even if the walking frequency suddenly changes.

3.3. Comparison Between FADMPs and AO

In this part, a simulation are carried out to further investigate the performance of FADMPs when the walking frequency and amplitude are suddenly changed. And to show the advantages of FADMPs, the performance of AO and FADMPs are compared with each other. The human hip joint trajectory during walking is generated by the DMPs algorithm proposed in Schaal (2006). Before simulation the actual hip joint trajectory during walking is learned by DMPs, and then DMPs was applied to generate the periodical hip joint trajectory of any frequency and amplitude while keeping the shape of trajectory.

The simulation protocols are set as follow. The total simulation time is 15 s. The initial frequency of hip joint trajectory is 2 Hz and the frequency is changing to 3 Hz at 5 s and changing back to 2 Hz at 9 s. The initial amplitude of hip joint trajectory is about 0.81 rad and the amplitude is changing to 2.46 rad at 5 s and changing back to 0.81 rad at 10 s. The initial parameters of the FADMPs and AO are shown in **Table 1**. The main parameters of FADMPs include the number (N) and width (h) of Gaussian-like kernel functions, the open-loop gain of FADMPs (α_z), the oscillation baseline of the output of FADMPs ($g = 0$) and the initial state of FADMPs $[\dot{z}(0), \dot{y}(0), \ddot{y}(0)]$. The main parameters of AO include the number (N) and width (h) of Gaussian-like kernel functions, the number of oscillators (K), the integrator gain of each oscillator (υ, η), the initial frequency of oscillator $[\omega(0)]$ and the initial phase of the basic oscillator $[\phi_0(0)]$. To show the influence of initial parameters of AO on the trajectory learning and prediction, as shown in **Table 1**, there are two sets of initial parameters of AO used in simulation.

The simulation results are shown in **Figure 3**. **Figure 3A** shows the hip joint trajectory online prediction results of FADMPs and AO algorithm. The original hip joint trajectory was filtered by a Butterworth low-pass filter (4th-order, cutoff frequency 10 Hz). To compensate the phase delay caused by the low-pass filter, both FADMPs and AO can use a kernel-based non-linear filter shown in Ronsse et al. (2011) to adjust the phase lead (Δ_ϕ) of predicted trajectory. In this simulation, Δ_ϕ of AO and FADMPs are set according to (17). To show the performance of phase compensation, the absolute error between the original trajectory and the predicted trajectory is calculated and shown in **Figure 3C**. The frequency estimation results of FADMPs and AO are shown in **Figure 3B**.

As mentioned before, FADMPs only compensate the phase delay when the walking frequency is stable. Hence, as shown in (C1) and (B1), the trajectory error is significantly reduced when the frequency error between two steps satisfies the condition shown in (16), and this phenomenon indicates that FADMPs successfully compensates the phase delay in the filtered trajectory. But the phase compensation performance of AO is not as good as FADMPs. As shown in (C2), the absolute error of AO is significantly higher than the one of FADMPs, especially after the walking frequency and amplitude have changed. The main reason for this phenomenon is that the fundamental frequency estimation results of AO fluctuate around it's true value, as shown in (B2). And the changing of walking frequency and motion amplitude have a great effect on the frequency estimation results of AO. Comparing with (B1), it is obvious that the frequency estimation result of FADMPs is more stable than AO. For this reason, the trajectory prediction performance of FADMPs is better than AO.

Furthermore, the results shown in (B3) and (C3) indicate that the initial value of AO have significant effect on the performance of frequency estimation and trajectory prediction. It is obvious that the convergence speed of the fundamental frequency $\omega(t)$ of AO becomes slower and $\omega(t)$ even becomes negative at about 10.7 s when the initial frequency and phase of AO are respectively changed as $\omega(0) = \pi$ and $\phi_0(t) = \pi/2$. Due to the length limitation of this paper, we only discussed the influence of $\omega(0)$ and $\phi_0(t)$, but the results in Chinimilli et al. (2019) have shown that the initial amplitude of each oscillator $[\alpha_i(0)]$ will also has significant effect on the frequency estimation and trajectory prediction of AO. Above simulation results show that the initial values of AO have a significant effect on the performance of AO. However, to the best of our knowledge, there is still no suitable theory to guide how to choose proper initial values of AO.

On the contrary, the performance of frequency estimation and trajectory prediction of FADMPs is barely affected by its initial values. On one hand, the frequency estimation results of FADMPs is more stable than AO and the performance of trajectory prediction of FADMPs is also better than AO. On the other hand, as shown in (D), the predicted velocity trajectory of FADMPs is smoother than the one of AO and the velocity oscillation of FADMPs is also smaller than AO when walking frequency and amplitude are changing. Moreover, comparing with (D2) and (D3), we can find that the initial value of AO can also affect the smoothness of trajectory prediction of AO. A smooth trajectory prediction is important for exoskeleton to provide a stable and comfortable assistance. Therefore, the above

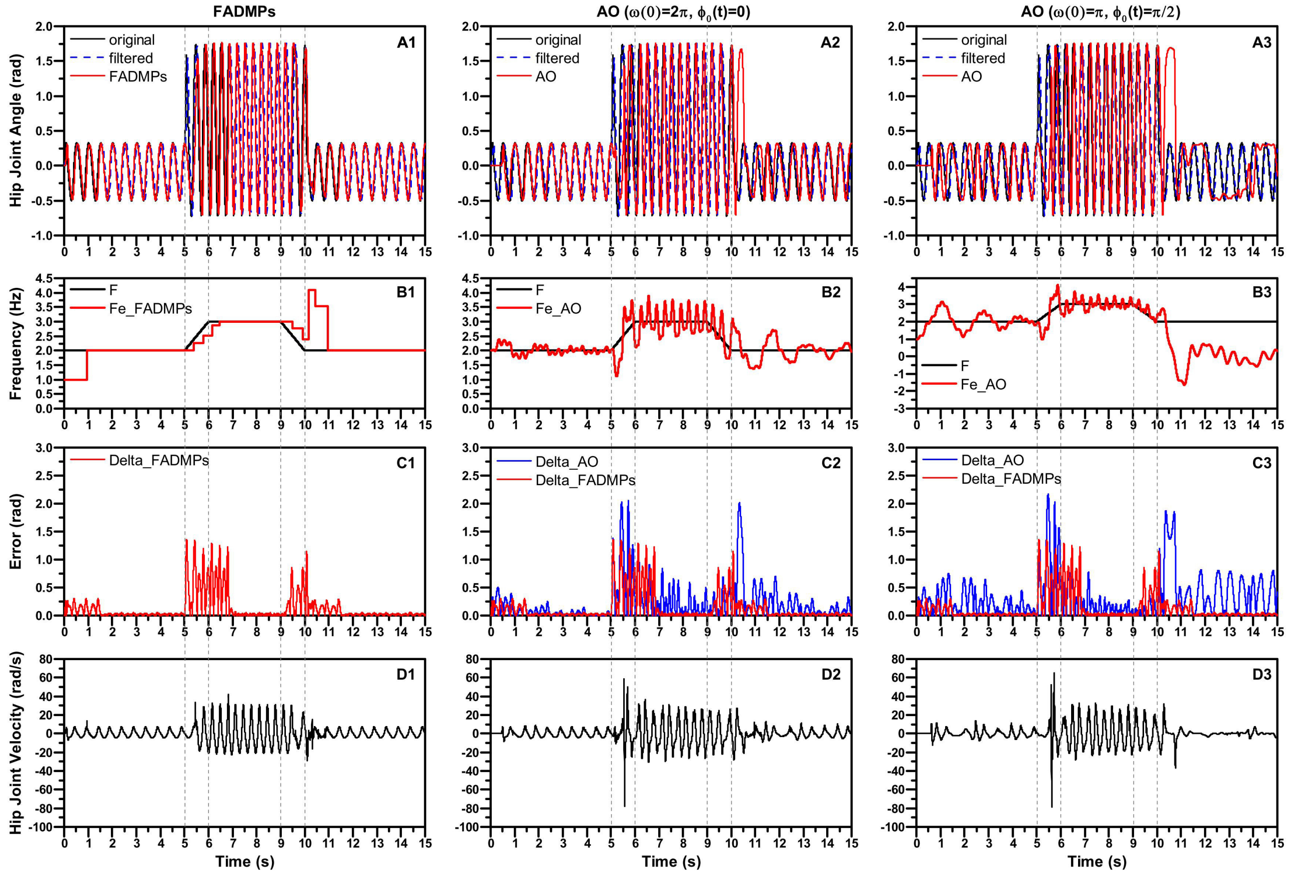


FIGURE 3 | Simulation of FADMPs and AO. The black solid line, blue dash line and red solid line shown in **(A)** represent the original trajectory containing noise, the low-pass filtered trajectory and the prediction trajectory of FADMPs or OA, respectively. The black line and red line shown in **(B)** represent the true frequency of trajectory and the estimated frequency, respectively. The red line and blue line shown in **(C)** denote the prediction error of FADMPs and AO, respectively. The black line shown in **(D)** denotes the predicted joint velocity.

simulation results indicate that the FADMPs algorithm proposed in this paper can avoid the drawbacks of AO algorithm and have a better performance on the frequency estimation and trajectory prediction than AO algorithm.

4. SIMULATIONS

In this section, the exoskeleton active walking assistance control framework based on FADMPs will be tested by simulation. The simulation framework of active walking assistance control based on FADMPs is shown in **Figure 4A**.

The simulation control system is established by using Matlab/Simulink. A human model (height: 1.7 m, weight: 65 kg) shown in **Figure 4B** is built by using Matlab/Simscape which is a powerful multidomain physical simulation system. As shown in **Figure 4C**, the motion of human model is limited in the sagittal plane. The input of human model is the lower limb joint trajectory (q_{in}) and joint assistance torque ($\alpha\hat{\tau}_h$). q_{in} is generated by a DMPs based motion generator (Schaal, 2006) shown in **Figure 4A**. Before simulation, the real human joint trajectory during walking is learned by DMPs and then we can use the DMPs to generate target joint trajectory q_{in} of human model. The advantage of using DMPs to generate q_{in} is that the amplitude and frequency of the generated trajectory can be easily adjusted only by changing the scale parameters of DMPs (Schaal, 2006).

The hip and knee joint of the human model are assisted by a massless ideal exoskeleton torque actuator which is able to generate any assistance torque profile on human joint. And the joint motion of human model will remain the same when the joint is assisted by exoskeleton. Hence, our active assistance control framework can prove to be effective if the energy consumption of human joints are reduced.

To verify the performance of our active assistance control framework based on FADMPs, two simulation experiments are carried out. In the first simulation experiment, the frequency of q_{in} will remain at 0.5, 1.0, 2.0, and 3.0 Hz, respectively. And ten different α (0.1–1.0) will be chosen to show the influence of AR. In the second simulation, the frequency of q_{in} will firstly increase from 1.0 to 3.0 Hz and then decrease from 3.0 to 1.0 Hz. And the AR will keep at 0.3 during the whole simulation.

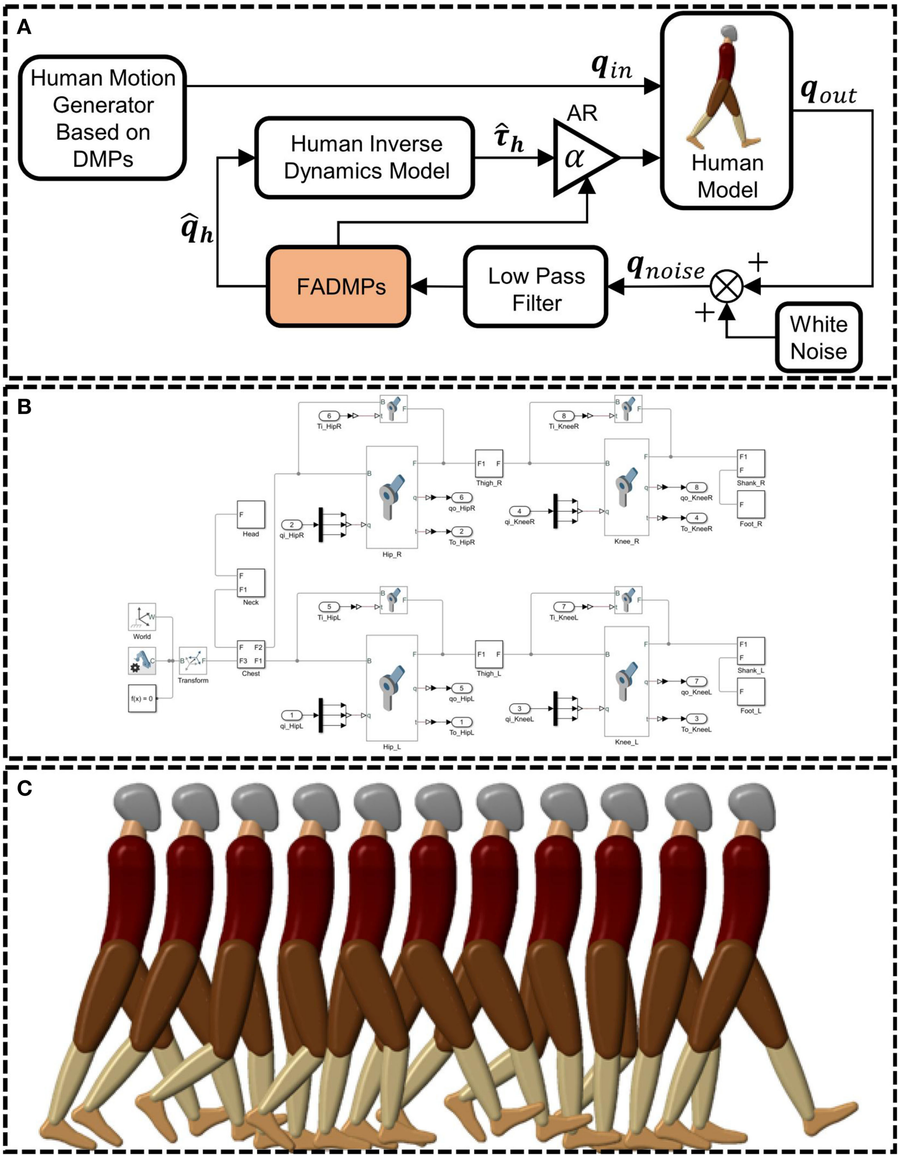


FIGURE 4 | Simulation framework of active walking assistance control based on FADMPs. q_{in} is the target hip and knee joint trajectory of human model. q_{out} is the output joint trajectory of human model. q_{noise} is the measurement of joint trajectory. $\hat{q}_h$ is the predicted trajectory of FADMPs. $\hat{\tau}_h$ is the estimation of human joint torque based on the human inverse dynamics model. **(A)** Is the simulation framework of active assistance based on FADMPs. **(B)** Is the human model established by Matlab/Simscape. **(C)** Shows the snapshot of active walking assistance simulation.

4.1. Active Assistance Control Simulation on Constant Frequency Walking

The FADMPs algorithm proposed in this paper is able to online trace the frequency and learn the shape of input trajectory. **Figure 5** shows the online learning and prediction results of the hip and knee joint trajectory based on FADMPs.

Figures 5A,B shows the online prediction results without phase compensation ($\Delta_\phi = 0$). It is obvious that the red dash line coincides with the blue solid line, which means the predicted trajectory keeps the same phase with the filtered trajectory. Hence, there is a phase delay exists in the predicted trajectory compared with the original trajectory. And due to the phase delay, as shown in **Figures 5A,B**, there is a large error between the original trajectory and the predicted trajectory. The RMSE of the predicted hip joint trajectory is about 0.16 rad and the RMSE of the predicted knee joint trajectory is about 0.22 rad. As mentioned before, to reduce the error caused by phase delay, the phase delay of the predicted trajectory can be compensated by FADMPs algorithm if we set Δ_ϕ as a positive value which is determined by the frequency response of low-pass filter. In this paper, Δ_ϕ is given by (17) which is a linear approximate of the phase-frequency characteristic of the Butterworth low-pass filter

(4th-order, cutoff frequency 10 Hz) in the range of 0–5 Hz.

$$\Delta_\phi = 15F_m\pi/180 \tag{17}$$

The results of phase compensation are shown in **Figures 5C,D**. According to (16), the human walking gait is considered to be stable when the error between two adjacent steps is <0.1 Hz. And FADMPs will execute phase compensation when human walk in a stable gait. Therefore, as shown in **Figures 5C,E**, FADMPs starts compensate the phase delay of hip joint trajectory at 1.45 s. Similarly as shown in **Figures 5D,F**, FADMPs starts compensate the phase delay of knee joint trajectory at 1.51 s. Finally when the phase delay has been compensated, the error between the original trajectory and the predicted trajectory of FADMPs is significantly reduced. The RMSE of the predicted hip and knee joint trajectory are about 0.0088 rad (reduce 94.5%) and 0.0106 rad (reduce 95.2%), respectively. Therefore, this simulation shows that the trajectory error between the original trajectory and predicted trajectory of FADMPs can be significantly reduced by compensating the phase delay caused by the low-pass filter.

Figure 5 shows the influence of phase delay on the hip joint torque and work when walking frequency is 1 Hz and AR is set as 0.5. As we can see from **Figure 5G**, the phase of exoskeleton assistance torque profile $\alpha\hat{\tau}_h$ ($T1$) is not coincide with the original human joint torque ($T0$) due to the phase delay in the predicted joint trajectory $\hat{q}_h$. For this reason, the direction of exoskeleton assistance torque $\alpha\hat{\tau}_h$ is not always keep the same with the original human joint torque. And if $\alpha\hat{\tau}_h$ is in an opposite direction of the original human joint torque, exoskeleton will impede the human motion (Li et al., 2018; Qiu et al., 2020). Hence, as shown in **Figure 5G**, the human joint torque (blue dash line) becomes larger than the original hip joint torque (black solid line) when the assistance torque (red solid line) is in an opposite direction of original human joint torque (black solid line). On the contrary, as shown in **Figure 5H**, the phase of exoskeleton assistance torque $\alpha\hat{\tau}_h$ is almost coincide with the original human joint torque because the phase delay of the predicted joint trajectory is compensated by FADMPs. Therefore, the human joint torque (green short dash line) significantly becomes smaller than the original human joint torque (black solid line). Furthermore, the work of human joint torque during 40 s walking is shown in **Figure 5I** from which we can see the significant difference before and after the phase delay is compensated. The normalized work of hip joint torque after 40 s walking without exoskeleton assistance ($W0$) is about 24.74 J/kg. When the phase delay of assistance torque is not compensated, the normalized work of hip joint torque after 40 s walking with exoskeleton assistance ($W1$) is about 19.75 J/kg (reduce 20.17%). However, when the phase delay is compensated by FADMPs, the normalized work of hip joint torque becomes about 16.67 J/kg (reduce 32.62%). Therefore, above simulation results indicate that the performance of exoskeleton active assistance can be significantly improved by compensating the phase delay of the filtered joint trajectory.

Figure 6 shows the change of energy consumption of hip joint after 40 s walking at different frequency (0.5, 1, 2, and 3 Hz). First of all, from **Figures 6A–D**, we can see that for a same AR the hip joint can reduce more energy consumption with a phase compensated assistance of exoskeleton. Moreover, with the increase of walking frequency, the performance of assistance gradually becomes worse if the phase delay is not compensated. Hence, these simulation results further demonstrate that compensating the phase delay is necessary for improving the performance of exoskeleton active assistance, especially for a high AR and high walking frequency assistance.

4.2. Active Assistance Control Simulation on Variable Frequency Walking

Previous simulations only focus on the assistance of a constant frequency walking. In this section, we will investigate the assistance performance of FADMPs on the variable frequency walking. As shown in **Figure 7C**, the walking frequency firstly increases from 1.0 to 3.0 Hz and then decreases from 3.0 to 1.0 Hz. On one hand, according to (16), FADMPs compensates the phase delay only when the walking frequency is stable. Hence, as shown in **Figures 7A,B**, the prediction error will be significantly reduced by phase compensation when walking frequency is stable and it becomes larger when the walking frequency is changing. On the other hand, α is simultaneously set as 0 when the walking frequency is unstable, which means the exoskeleton is working in transparent mode when walking frequency is unstable. And if walking frequency is stable, α will be set as a positive value and exoskeleton will work in active assistance mode.

Figure 6E shows the energy consumption of hip joint after 40 s variable frequency walking with the assistance of exoskeleton ($\alpha = 0.3$). It is obvious that, comparing with $W0$, the performance of active assistance with phase compensation (173.07 J/kg, reduce 19.41%) is much better than the one without phase compensation (206.42 J/kg, reduce 3.91%). **Figure 6F** shows the influence of AR on the performance of active assistance of variable frequency walking. After the phase compensation, the hip joint work decreases linearly with the increase of AR. However, if the phase delay is not compensated, the reduction rate of hip joint work is much lower and when AR exceeds 0.5 the hip joint work even surpasses the hip joint work (214.83 J/kg) without the assistance of exoskeleton.

Above simulation results show that compensating the phase delay can significantly improve the performance of active assistance of exoskeleton when the walking frequency is variable. Therefore, the simulation results shown in **Figure 6** demonstrate that the active assistance control based on FADMPs is able to significantly improve the performance of active assistance of exoskeleton by online compensating the phase delay of the filtered joint trajectory. Especially for a high walking frequency and high AR assistance, FADMPs can significantly improve the assistance performance of exoskeleton.

5. EXPERIMENTS

5.1. Participants

Nine healthy volunteers (age: 32 ± 3.61, weight: 66.89 ± 7.03 kg, height: 1.70 ± 0.03 m) participated in the active walking assistance experiments shown in **Figure 8**. All participants have no musculoskeletal injuries or cardiovascular disease and

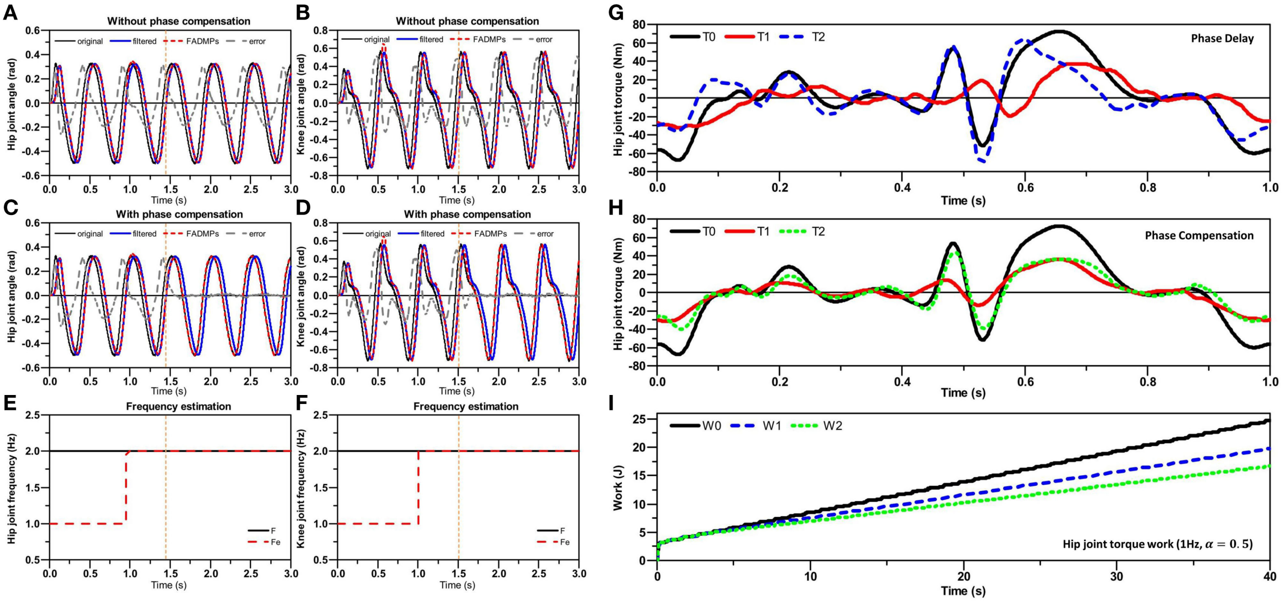


FIGURE 5 | Simulation results of FADMPs. In **(A–D)**, the black solid line represents the original input joint trajectory containing measurement noise. The blue solid line represents the filtered trajectory (Butterworth low-pass, 4th order, cutoff frequency 10 Hz). The red short dash line represents the prediction results of FADMPs. The gray dash line represents the error between the original trajectory and the prediction trajectory. In **(E,F)**, the black solid line denotes the frequency of original trajectory and the red dash line denotes the results of frequency estimation of FADMPs. The profile of active assistance torque (1 Hz, $\alpha = 0.5$) and the work of hip joint torque are shown in **(G–I)**. In **(G,H)**, the black solid line $T0$ is the human joint torque during walking without exoskeleton assistance. The red solid line $T1$ is the assistance torque profile of exoskeleton ($\alpha = 0.5$). $T2$ is the human joint torque during walking with the assistance of exoskeleton. In **(I)**, the black line $W0$ is the work of human joint torque during 40 s walking without the exoskeleton assistance. The blue dash line $W1$ is the work of human joint torque with a phase delay assistance of exoskeleton ($\Delta_\phi = 0$). The green short dash line $W2$ is the work of human joint torque with a phase compensated assistance of exoskeleton ($\Delta_\phi > 0$).

they gave their informed consent before participating in the experiments which were approved by the local ethical committee.

5.2. Hip Exoskeleton System

The hip exoskeleton system used in this paper is shown in **Figure 8**. The hip exoskeleton designed by our team contains two motor-driven joints which are made up of a 24 V brushless DC motor and a planetary gearbox ($i = 8$). There are two load cells embedded in the output shafts of exoskeleton to measure the human-exoskeleton interaction force. The human lower limb joints' motion trajectories are measured by five 9-axis inertial measurement unit (IMU) which are placed on the waist (IMU × 1), thigh (IMU × 2) and shank (IMU × 2), respectively. The feet ground reaction forces are measured by two pressure insoles. The total mass of the hip exoskeleton is about 11 kg and the output rated torque of actuator is 40 Nm.

5.3. Experimental Protocol

In the active walking assistance experiments, as shown in **Figure 8**, all of the participants were wearing Metabolic system, EMG sensors and IMU while walking on a treadmill. Four walking conditions were evaluated in our experiments: *Free*, *OFF*, *TRA*, and *ASS*. In the *Free* condition, as shown in **Figure 8A**, participants were walking on the treadmill without wearing exoskeleton. But it should be noticed that in order to compare the difference of human metabolic cost before and after wearing exoskeleton the weight of exoskeleton should be considered in the *Free* condition. Hence, a heavy load (11 kg) which has an equivalent weight of the hip exoskeleton should be carrying during the free walking experiment. In the *OFF* condition, participants were wearing the hip exoskeleton and walking on the treadmill. But the exoskeleton was power off in this condition. Hence, exoskeleton was passively moving with human body. In the *TRA* condition, exoskeleton was working in a transparent mode which was neither impeding nor assisting human walking. On the contrary, in the *ASS* condition, exoskeleton was working in active assisting mode to assist human walking. The active control framework shown in **Figure 1** was applied in the active assisting mode.

There are two walking speed modes in active walking assistance experiments: constant speed walking and variable speed walking. In the constant speed walking experiments, as shown in **Figure 9A**, participants were walking on the treadmill and the velocity of treadmill was respectively set as 2, 4, and 6 km/h. In the variable speed walking experiments, as shown in **Figure 9B**, the target velocity of the treadmill was firstly increasing from 2 to 6 km/h and then decreasing from 6 to 2 km/h.

The experiment process for each participant is shown in **Figure 9C**. At the beginning of experiment, participants were walking in the *Free* condition for 11 min. And then participants had a 10 min rest. Next, participants put on the hip exoskeleton

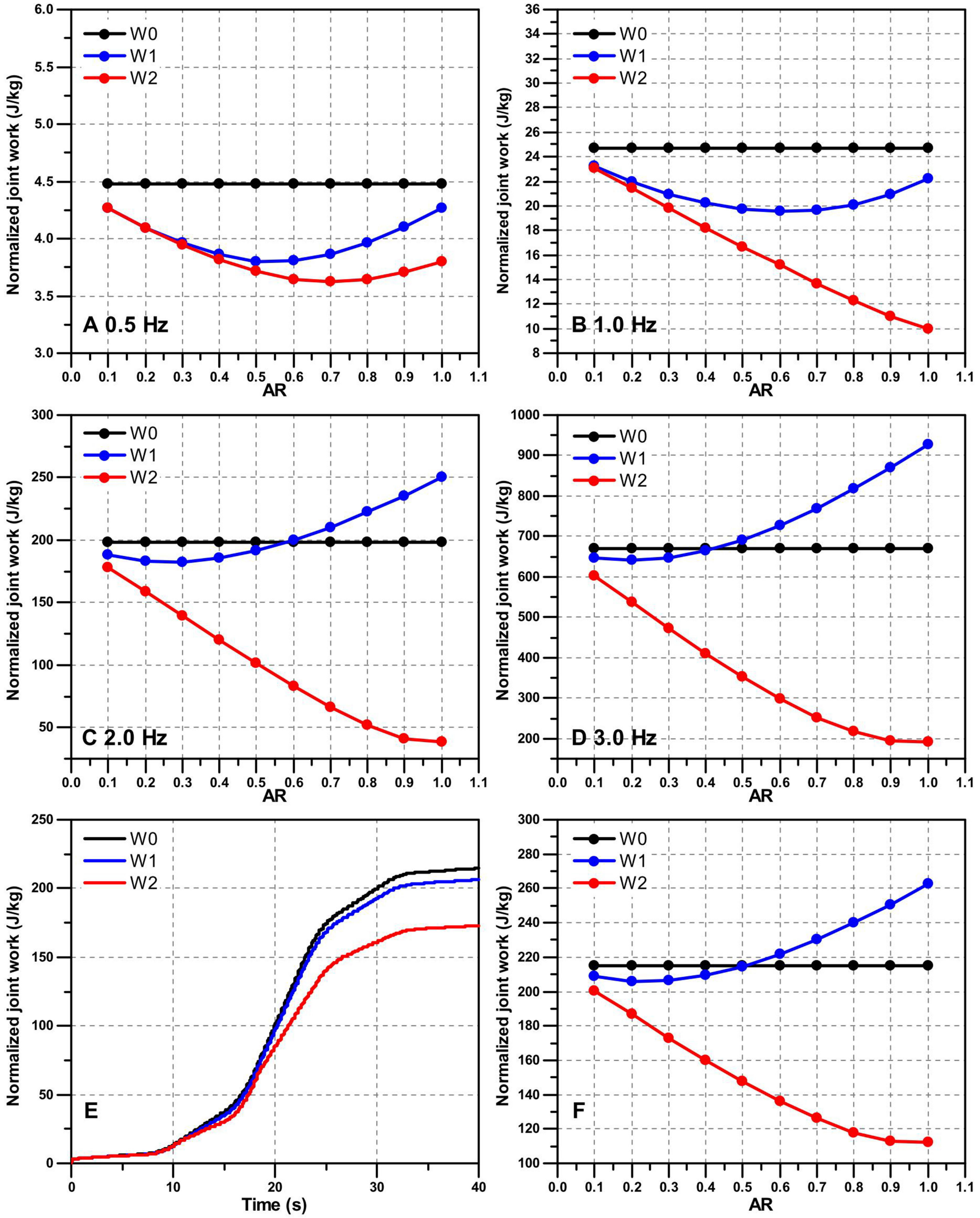


FIGURE 6 | The performance of active assistance on the hip joint work (after 40 s walking) under different walking frequency and AR. The black dot line $W0$ is the hip joint work without the assistance of exoskeleton. The blue dot line $W1$ is the hip joint work with a phase delay assistance of exoskeleton ($\Delta_\phi = 0$). The red dot line $W2$

(Continued)

FIGURE 6 | is the hip joint work with a phase compensated assistance of exoskeleton ($\Delta_\phi > 0$). **(A–D)** Show normalized hip joint work with the active assistance of exoskeleton after 40 s constant frequency walking. In this paper, the hip joint work is normalized by the weight of each subject. **(E,F)** Show the performance of active assistance on the 40 s variable frequency walking. **(E)** Shows the normalized hip joint work with the active assistance of exoskeleton ($\alpha = 0.3$). **(F)** Shows the influence of different AR on the performance of active assistance during variable frequency walking.

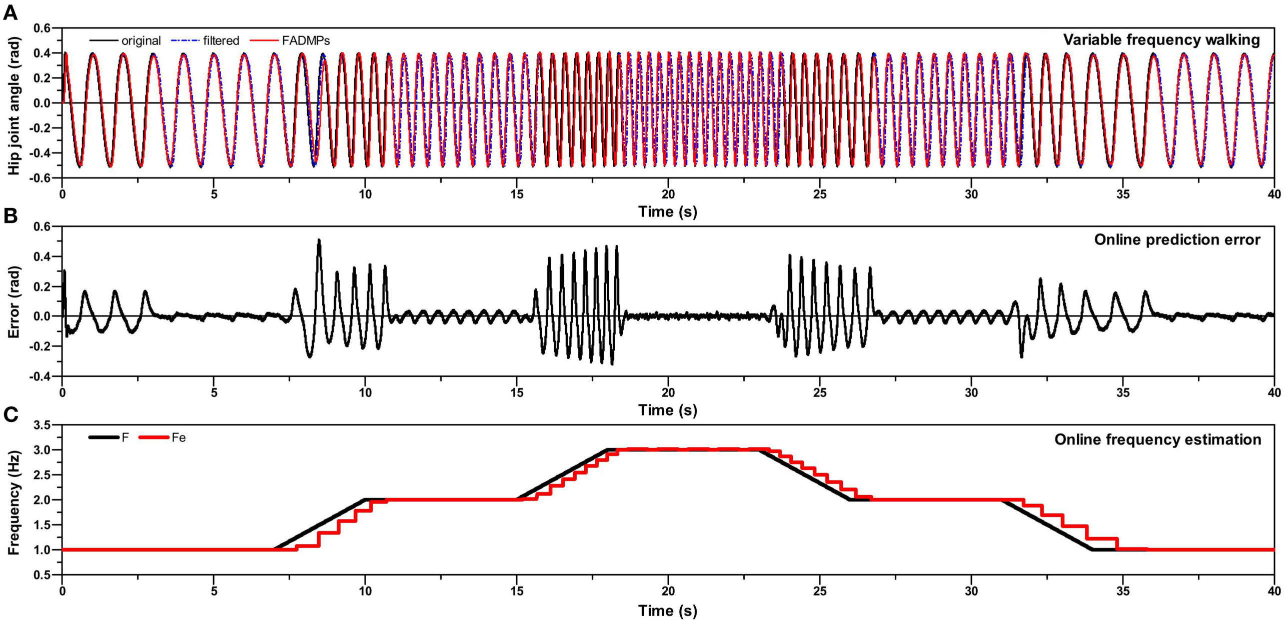


FIGURE 7 | Simulation of online variable frequency hip joint trajectory learning and prediction based on FADMPs. **(A)** Is the human hip joint angle during variable frequency walking. **(B)** Shows the error between the original trajectory and the predicted trajectory of FADMPs. **(C)** Shows the frequency of original trajectory (black solid line) and estimated frequency (red solid line).

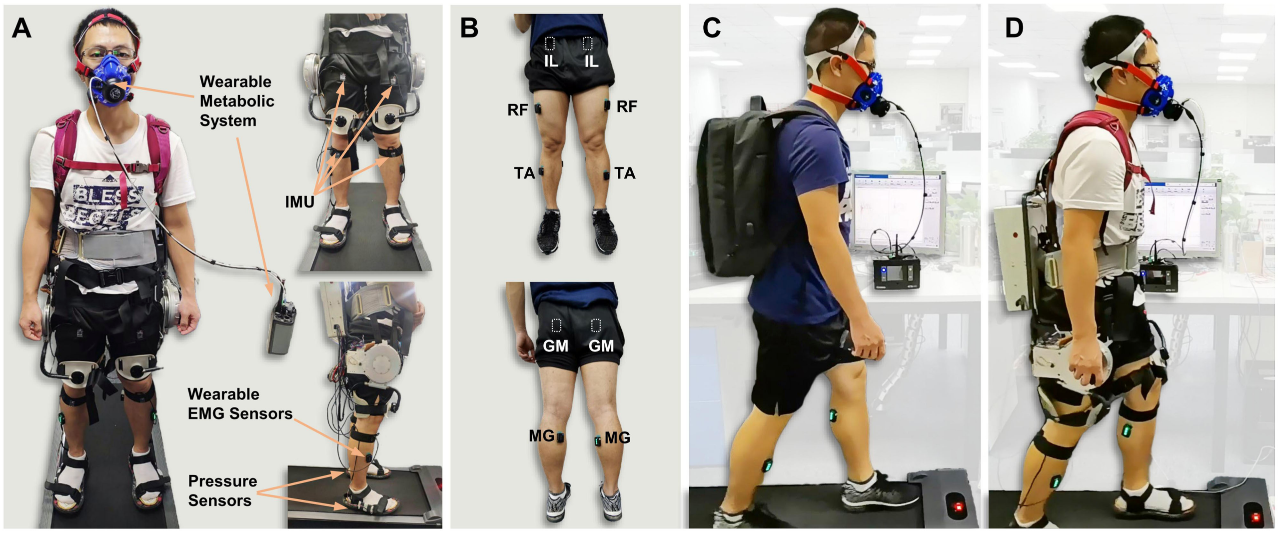


FIGURE 8 | Experimental method. **(A)** The front and side view of an instrumented participant. The participant's motion is measured by the IMU wearing on the limb. The muscle activity during walking is record by surface electrodes (Delsys Trigno Avanti, USA). The participant's metabolic cost during walking is measured by a wearable metabolic system (K5, COSMED, Italy). **(B)** Placement of surface electrodes: iliopsoas (IL), rectus femoris (RF), tibialis anterior (TA), gluteus maximus (GM), and medial gastrocnemius (MG). **(C,D)** Show the active walking assistance experiments. **(C)** Participant walking with a heavy load on back. The weight of the load is equal to the hip exoskeleton. **(D)** Participant walking with the hip exoskeleton.

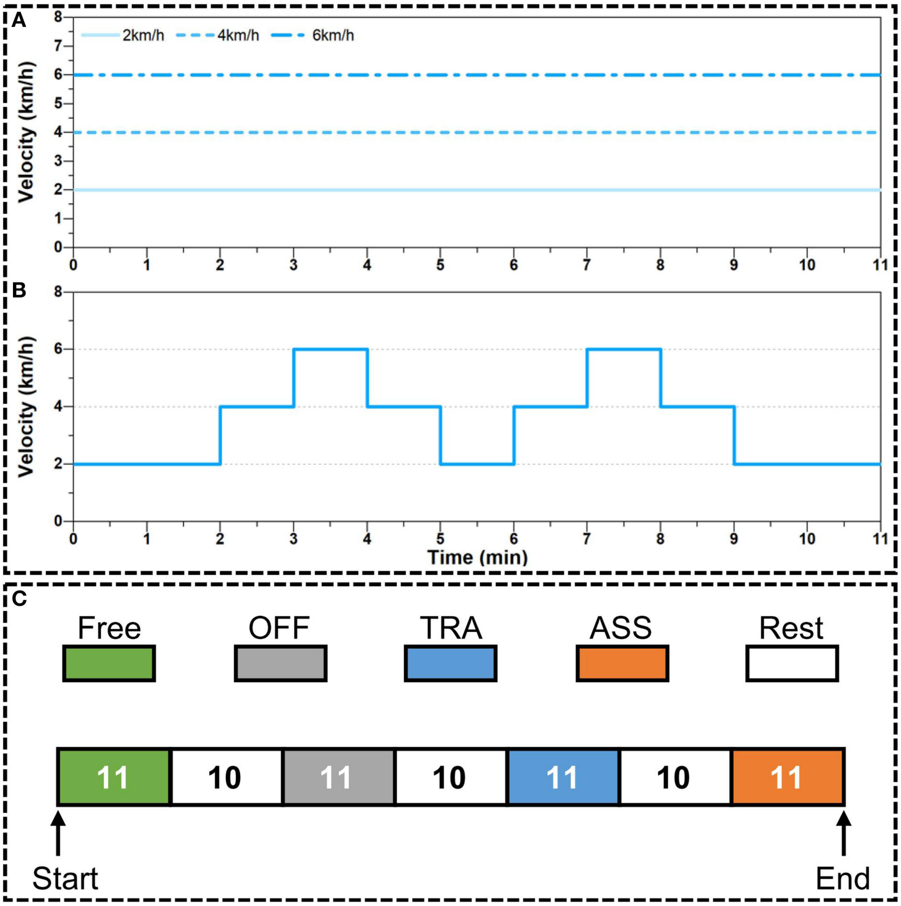


FIGURE 9 | Experiment protocol of active walking assistance control. **(A)** The target velocity of treadmill during constant speed walking experiments. **(B)** The target velocity of treadmill during variable speed walking experiments. **(C)** The process of walking assistance experiments. The number in each block denotes the duration time (units: min) of each walking mode: free walking (Free), exoskeleton working in power-off mode (OFF), exoskeleton working in transparent mode (TRA), and exoskeleton working in active assistance mode (ASS).

and walked in the *OFF* condition for 11 min. After that, participants rested for 10 min and then continued to walk in *TRA* condition for 11 min. And finally, after 10 min rest, participants were walking in *ASS* condition for the last 11 min. The experiment data, including joint trajectory, EMG signal and metabolic cost, were only recorded when participants were walking in the *Free*, *OFF*, *TRA*, and *ASS* conditions.

5.4. Results

To further investigate the performance of the active walking assistance control method proposed in this paper, the active walking assistance experiments shown in **Figures 8C,D** were carried out. And this paper will evaluate the performance of the active walking assistance control from the following three aspects: joint trajectory, muscle activity and human metabolic cost. There are three situations in which exoskeleton can be considered to provide effective active assistance on human body (Nagarajan et al., 2016). First, the amplitude of human motion increases while the muscle activity remains the same. Second, the muscle activity decreases while the amplitude of human motion keeps the same. Third, not only the amplitude of human motion increases, but also the muscle activity decreases.

5.4.1. Effect of Active Assistance on the Limb Joint Trajectory

Figure 10 shows the online learning and prediction results of FADMPs algorithm. In **Figures 10A,B**, the black solid line denotes the original measured joint trajectory which is recorded by the IMU system shown in **Figure 8B**. The original trajectory is not smooth enough because the update rate of IMU is 100 Hz which is less than the sample rate of the exoskeleton control system (1 kHz). Hence, the original trajectory must be filtered by a low pass filter (4th-order, Butterworth, cut-off frequency 10 Hz). However, as shown in **Figure 10**, the filtered trajectory (blue solid line) has a phase delay compared with the original trajectory. To compensate the phase delay, a FADMPs algorithm is proposed in this paper to online compensate the phase delay in the filtered trajectory. As shown in **Figure 10**, the phase of the trajectory predicted by FADMPs (red line) is almost coincide with the phase of original trajectory. The RMSE between the original

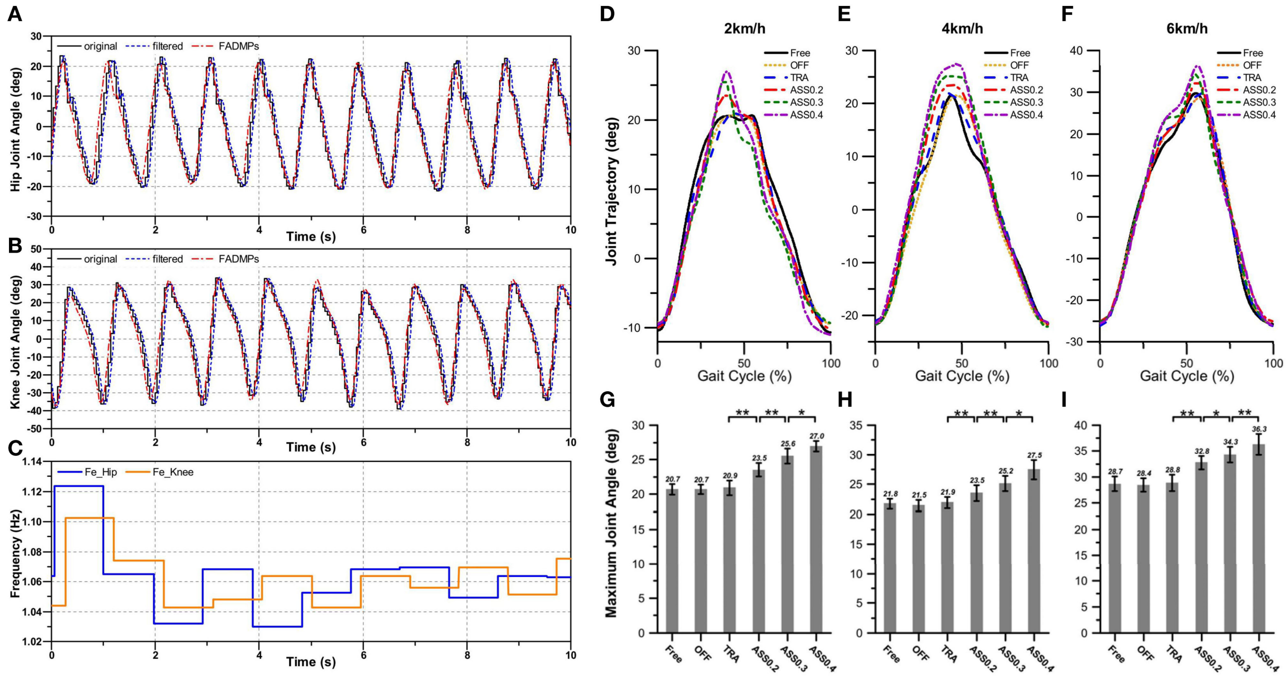


FIGURE 10 | The results of active walking assistance experiments. The online learning and prediction results of hip and knee trajectory during active walking assistance experiments with a constant walking speed of 4 km/h are shown in **(A,B)**, respectively. **(C)** Shows the frequency estimation result of hip and knee joint trajectory, respectively. The changing of the hip joint trajectory during active walking assistance is shown in **(D–I)**. The braces with $**$ denote statistical significant difference between two condition ($p < 0.01$). And $*$ represents a significant difference among trials ($p < 0.05$). *OFF* represents exoskeleton is working in power-off mode. *TRA* represents exoskeleton is working in transparent mode. *ASS*0.2 denotes exoskeleton is working in active assistance mode with $\alpha = 0.2$. Similarly, *ASS*0.3 and *ASS*0.4 represent active assistance mode with $\alpha = 0.3$ and $\alpha = 0.4$, respectively.

TABLE 2 | The RMSE between the original trajectory and the prediction trajectory of FADMPs.

Speed	RMES of left hip trajectory			RMES of left knee trajectory		
	Filtered	**FADMPs**	**Rate (%)**	**Filtered**	**FADMPs**	**Rate (%)**
2 km/h	2.99°	1.63°	45.48	5.27°	2.57°	51.23
4 km/h	4.49°	2.76°	38.53	6.93°	3.42°	50.65
6 km/h	6.37°	4.58°	28.10	8.42°	4.90°	41.81
Speed	**RMES of right hip trajectory**			**RMES of right knee trajectory**		
	Filtered	**FADMPs**	**Rate (%)**	**Filtered**	**FADMPs**	**Rate (%)**
2 km/h	2.92°	1.64°	43.84	5.57°	3.14°	43.63
4 km/h	4.66°	2.85°	38.84	7.73°	4.42°	42.82
6 km/h	6.27°	4.31°	31.26	8.95°	4.36°	51.28

trajectory and the prediction trajectory of FADMPs is calculated in this paper. As shown in **Table 2**, the RMSE between the original trajectory and the prediction trajectory is significantly reduced when the phase delay is compensated by FADMPs. The *Rate* in **Table 2** means the reduction rate of RMSE of the phase compensated trajectory ($\Delta_\phi > 0$) comparing with the RMSE of the filtered trajectory. It is obvious that the reduction rate of RMSE is relatively high, which demonstrates that the phase delay is successfully compensated by FADMPs.

Statistical significance of the changing of hip joint trajectory was evaluated by using one-way repeated measures analysis of variance (ANOVA). The changing of hip joint trajectory in different walking conditions and speeds are shown in **Figures 10D–I**, which show that there is no significant different in hip joint trajectory when participants walking in *Free*, *OFF*, and *TRA* mode. This phenomenon means that the hip exoskeleton has no effects on the human motion when it offers no assistance on human body. But when hip exoskeleton is walking on *ASS* mode, the hip joint trajectory has a significant changing and the maximum hip joint angle increases with the rising of the AR. These results indicate that the motion range of human hip joint during walking can be

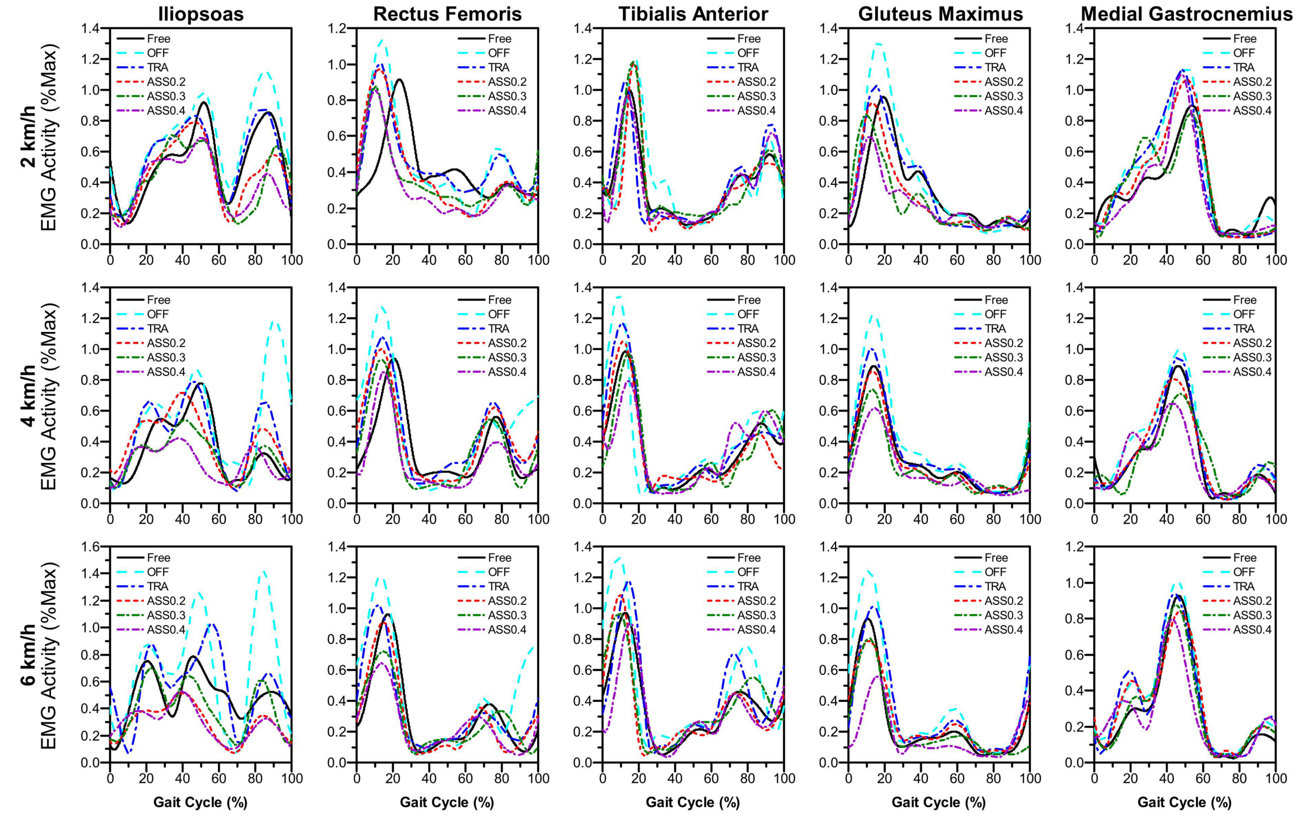


FIGURE 11 | Normalized average muscle EMG envelope of all participants during the active walking assistance experiments. The normalized EMG envelope as a percent of gait cycle (heel-strike to heel-strike) for the five muscles investigated.

significantly improved by exoskeleton when it working in *ASS* mode.

5.4.2. Effect of Active Assistance on the Lower Limb Muscle Activity

As mentioned before, in order to evaluate the performance of active walking assistance control, we need to further investigate the muscle activity of the participant during walking. According to the experiment protocol proposed in section 5.3, EMG signals from five lower limb muscles were simultaneously record by a wireless EMG measurement system shown in **Figure 8**. The lower limb muscles investigated in this paper were: illiopsoas (IL), rectus femoris (RF), tibialis anterior (TA), gluteus maximus (GM), and medial gastrocnemius (MG). The EMG signals were band-pass filtered (2th order Butterworth, cut-off 100–400 Hz). And the EMG linear envelope was estimated by using a moving RMS window (window length: 0.125 s, window overlap: 0.0625 s). To compare the muscle activity of different walking modes, for each participant and for each muscle, the EMG linear envelope was normalized to the average peak value (averaged from 2 to 10 min) measured during the *Free* mode. **Figure 11** shows the normalized average muscle EMG linear envelope of all participants during experiments at different walking speeds and modes.

Figure 11 shows that not only the amplitude of the lower limb muscle EMG envelope are significantly affected by the hip exoskeleton. Comparing with the EMG envelope during *OFF* mode, the amplitude of average EMG envelope is reduced when the hip exoskeleton is working on *ASS* mode ($\Delta_{\phi} > 0$). And the shape of average EMG envelope changes obviously with the assistance of hip exoskeleton. In order to further quantify the changes in muscle activity, the integral of the normalized average EMG envelope (*iEMG*) is given by:

$$iEMG = \Delta t \cdot \sum_{n=1}^{N} Y_n \tag{18}$$

where Y_n is the n_{th} sample of the normalized average EMG envelope, N is the total number of EMG samples and Δt is the integration step time.

As shown in **Figure 12**, *iEMG* is computed to quantitatively evaluate the effect of the active assistance of hip exoskeleton on the lower limb muscle activity. In this paper, one-way ANOVA is used for evaluating the significant changing of the *iEMG*. As shown in **Figure 12**, *iEMG* of five muscle groups significantly increase from the *Free* to the *OFF* condition. This phenomenon indicates that human muscle activity is enlarged when exoskeleton is working on power-off mode. The main

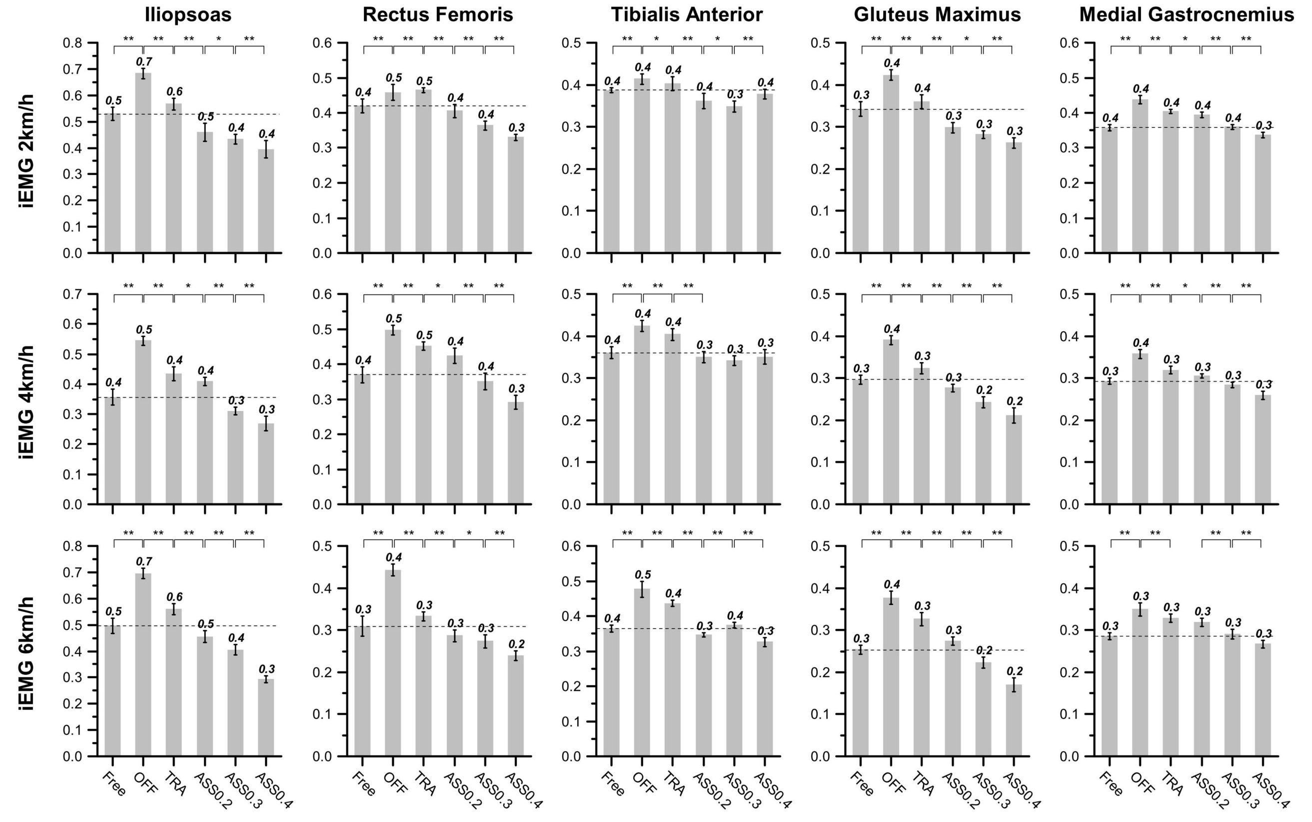


FIGURE 12 | The integral of the EMG activity of all participants at different walking conditions and speeds. The braces with $**$ denote statistical significant difference between two condition ($p < 0.01$). And $*$ represents a significant difference among trials ($p < 0.05$).

reason for this phenomenon is that the inertia of exoskeleton are completely compensated by human body. When exoskeleton is working on *TRA* mode, exoskeleton compensates most part of it's inertial. Hence, comparing with *OFF* mode, *iEMG* in *TRA* mode is significantly reduced. When hip exoskeleton is working on *ASS* mode *iEMG* is further reduced, especially for the IL, GM, and RF. This phenomenon indicates that hip exoskeleton has a greater influence on the IL, GM, and RF. Furthermore, the reduction rate of *iEMG* of IL, GM, and RF are rising with the increase of AR. But for the *iEMG* of TA and MG, their reduction rate is smaller than IL, GM, and RF. And TA is not significantly affected by the rising of AR.

Combining above experiment results shown in **Figures 10**, **11**, we can see that the motion amplitude of hip joint is increased and the muscle activity of the main extensor (GM and RF) and main flexor (IL) of the hip joint are simultaneously reduced while exoskeleton is working on ASS mode. And with the rising of AR, the motion amplitude of hip joint will also increase and the muscle activity of the main muscle groups of hip joint will reduce too. Therefore, these results demonstrate that the participants' hip joints are successfully assisted when hip exoskeleton is working on *ASS* mode ($\Delta_\phi > 0$).

5.4.3. Effect of Active Assistance on the Metabolic Cost

Muscle activity only reflects the energy change of the local muscle of the human body, but not the energy consumption of the whole body. Hence, the next part of this section will focus on how hip exoskeleton affects the human metabolic cost during active walking assistance experiments.

The human metabolic cost during constant speed walking and variable speed walking are recorded by a wearable metabolic system (K5, COSMED, Italy) shown in **Figure 8**. According to the experiment process shown in **Figure 10C**, the metabolic cost is only recorded in the following experiment process: *Free*, *OFF*, *TRA*, and *ASS*. And for each process, the human metabolic cost, including carbon dioxide rate and oxygen rate, is measured for 11 min by using breath-by-breath method. The wearable metabolic system used in our walking experiments is able to automatically calculate the metabolic power. To compare the differences in energy power between different participants, the metabolic power is normalized by the wight of each participant. And the normalized metabolic power is averaged from the 2 to 10 min of each experiment process. The normalized average metabolic power of all participants during the active walking experiment is shown in **Figure 13**.

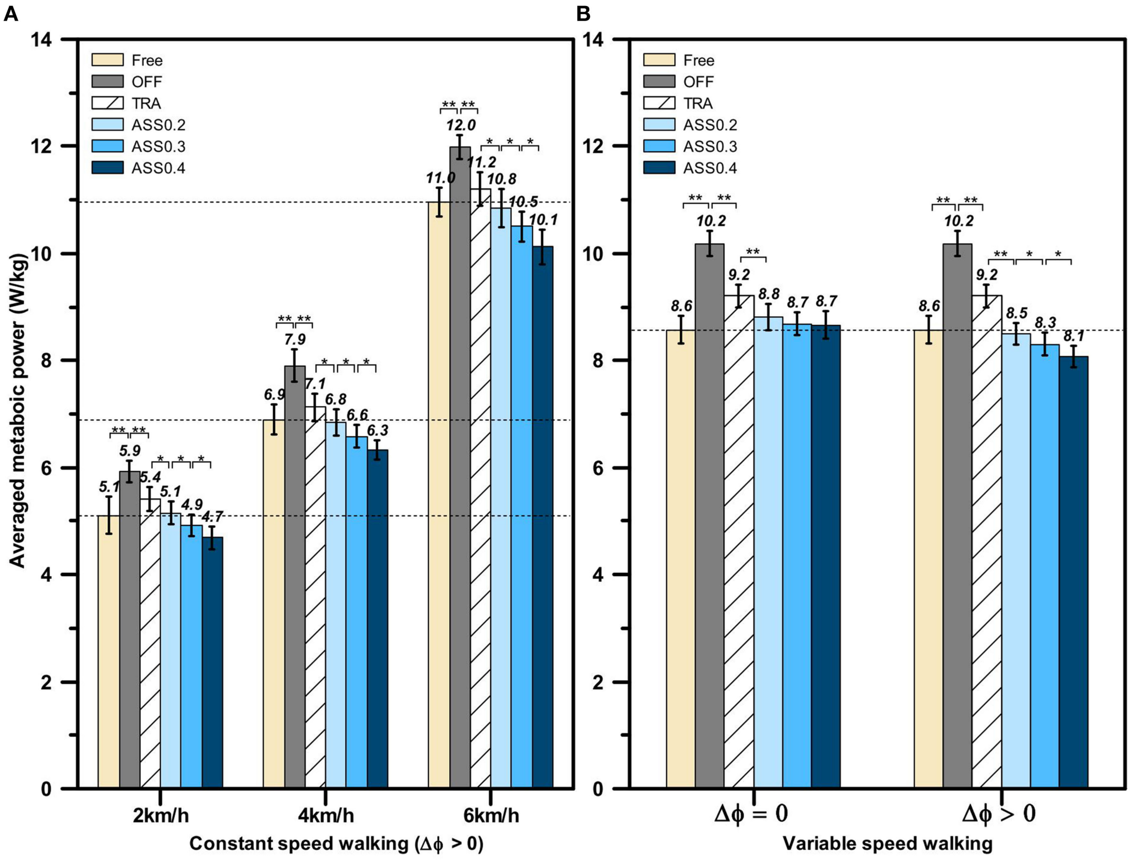


FIGURE 13 | The normalized average metabolic power of all participants during active walking experiments. **(A)** The normalized average metabolic power during constant speed walking assistant experiments. **(B)** The normalized average metabolic power during variable speed walking assistance experiments. The braces with $**$ denote statistical significant difference between two condition ($p < 0.01$). And $*$ represents a significant difference among trials ($p < 0.05$).

The averaged metabolic power in constant speed walking experiment is shown in **Figure 13A**, from which we can find that the hip exoskeleton is able to significantly affect the human metabolic power. The statistical significance difference between each walking mode is evaluated by using one-way ANOVA. First of all, comparing the averaged metabolic cost in *Free* walking mode at 2, 4, and 6 km/h (5.11 ± 0.34, 6.89 ± 0.28, 10.95 ± 0.26 W/kg), the human metabolic power is significantly increased ($p < 0.01$) when exoskeleton is working on *OFF* mode (5.93 ± 0.21, 7.90 ± 0.30, 11.99 ± 0.23 W/kg). And comparing with *OFF* mode, the human averaged metabolic power is significantly reduced ($p < 0.01$) when exoskeleton is working on *TRA* mode (5.41 ± 0.22, 7.13 ± 0.26, 11.21 ± 0.31 W/kg). The main reason for this phenomenon is that exoskeleton compensates most of it's inertial while working on *TRA* mode. Therefore, the impedance between human and exoskeleton is reduced. When exoskeleton is working on *ASS* mode and the phase delay is compensated by FADMPs ($\Delta_\phi > 0$), the averaged metabolic power is further reduced and the reduction rate of the averaged metabolic power is increased with the rising of AR. The above experiment results demonstrate that the active walking assistance control method based on FADMPs ($\Delta_\phi > 0$) is able to reduce the human metabolic power during constant speed walking.

To further investigate the performance of our active assistant control method during variable speed walking and investigate the effect of phase compensation on the performance of active assistance, the variable speed walking assistance experiment is carried out according to the experiment protocol in section 5.3. The target velocity of treadmill during variable speed walking assistance experiment is set as the value shown in **Figure 10B**. The normalized average metabolic power in variable speed walking is shown in **Figure 13B**. Comparing with the averaged metabolic power in *TRA* mode (9.20 ± 0.22 W/kg), the metabolic power in *ASS* mode is significantly reduced when $\Delta_\phi = 0$ and $\Delta_\phi > 0$. However, when $\Delta_\phi = 0$, with the rising of AR, there is no significant changing in the reduction rate of the averaged metabolic power and the averaged metabolic powers in *ASS*0.2, *ASS*0.3, and *ASS*0.4 walking modes (8.81 ± 0.25, 8.69 ± 0.21, 8.66 ± 0.26 W/kg) are higher than the averaged metabolic power in *Free* condition (8.57 ± 0.26 W/kg). But when $\Delta_\phi > 0$, on the contrary, the averaged metabolic power is significantly reduced with the rising of the AR and all of the averaged metabolic powers in *ASS* walking mode (8.49 ± 0.20, 8.31 ± 0.22, 8.07 ± 0.21 W/kg) are lower than the one in *Free* walking mode (8.57 ± 0.26 W/kg). Hence, the above experiments results indicate that the active walking assistance control method proposed in this paper is effective for both constant speed walking and variable speed walking assistance, and the performance of active assistance becomes better when the phase delay in the filtered joint trajectory is compensated by the FADMPs algorithm proposed in this paper.

6. CONCLUSIONS

This paper introduces an exoskeleton active walking assistance control framework based on FADMPs. FADMPs is an online learning algorithm which is able to online learning and prediction the human joint trajectory during walking. Comparing with AO algorithm, FADMPs has three main advantages: (1) The initial parameters of FADMPs have almost no effect on the performance of frequency estimation and trajectory prediction. (2) The sudden change of walking frequency and motion amplitude have no effect on the performance of frequency estimation and trajectory prediction. (3) FADMPs can online predict a smoother motion trajectory by only adjust the phase lead Δ_{ϕ}. Based on these advantages, the inevitable phase delay in the lowpass filtered joint trajectory can be online compensated by FADMPs. Therefore, the active walking assistance control framework based on FADMPs is able to provide a no-phase-delay assistance to the human joint during walking.

The active walking assistance control framework proposed in this paper is suitable for both constant speed walking assistance and variable speed walking assistance. The simulation results of active walking assistance indicate that the phase delay existed in the filtered trajectory is not beneficial to improve the performance of active assistance. The phase delay of the filtered trajectory will make the performance of active assistance become worse with the increase of walking frequency and AR. The performance of active assistance will be significantly improved when the phase delay is compensated by FADMPs, especially for the variable speed walking assistance. The effectiveness of the proposed active walking assistance control framework based on FADMPs is further demonstrated by the active walking assistance experiments. The experiment results show that the proposed control framework can improve the range of joint motion, reduce the related low limb muscle activity and cut down the metabolic cost during walking. And the reduction rate of human metabolic cost during variable speed walking is significantly increased when the phase delay is compensated by the FADMPs. Hence, both simulation and experiment results show that the active walking assistance control framework based on FADMPs is benefit for improving the performance of active walking assistance, especially for a high speed or variable speed walking assistance with a high AR.

The main limitation of the active walking assistance control framework based on FADMPs is that the control frame provide assistance to human body only when walking frequency is stable. However, in daily life, walking frequency may change due to environmental changes or external interference. When human suffers from an unexpected external inference, human may lose balance or even fall (Guo et al., 2019). To improve the stability of human walking, it is necessary for exoskeleton to active assist human motion when walking frequency is unstable to make human regain balance faster. Hence, in the future research, we will investigate an active control framework that can provide active assistance in both stable and unstable gaits. And we will try to develop a robot learning system (Bing et al., 2018; Yang et al., 2018) which enables the exoskeleton to recognize the type of environment and to choose the optimal assist strategy according to the different environment (Krausz and Hargrove, 2019).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Harbin Institute of Technology Ethical Review Board. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

SQ developed the FADMPs algorithm and responsible for data collection and processing. SQ and JD established the exoskeleton prototype and implemented the controller. XW designed the experiment protocol. SQ and FZ participated in the design and drafting of the manuscript. WG and FZ were involved in the results interpretation and critical revision of the study. All authors read and approved the final manuscript.

FUNDING

This work was supported by National Natural Science Foundation of China (Grant no. 61773139), National Natural Science Foundation of China (Grant no.51521003), National Natural Science Foundation of China (Grant no. U2013602), Shenzhen Science and Technology Program (Grant no. KQTD2016112515134654), Shenzhen Science and Technology Research and Development Foundation (Grant no. JCYJ20190813171009236).

ACKNOWLEDGMENTS

The authors thank Xiwen Meng and Shengguang Li for their helps in placing EMG electrodes for muscular activity measurements.

REFERENCES

Al-Quraishi, M. S., Elamvazuthi, I., Daud, S. A., Parasuraman, S., and Borboni, A. (2018). EEG-based control for upper and lower limb exoskeletons and prostheses: a systematic review. *Sensors* 18:3342. doi: 10.3390/s18103342

Asbeck, A. T., Rossi, S. M. M. D., Holt, K. G., and Walsh, C. J. (2015). A biologically inspired soft exosuit for walking assistance. *Int. J. Robot. Res*. 34, 744–762. doi: 10.1177/0278364914562476

Bing, Z., Meschede, C., Röhrbein, F., Huang, K., and Knoll, A. C. (2018). A survey of robotics control based on learning-inspired spiking neural networks. *Front. Neurorobot*. 12:35. doi: 10.3389/fnbot.2018.00035

Chao, Z., Yang, C., Chen, Z., and Dai, S. L. (2018). Robot learning human stiffness regulation for hybrid manufacture. *Assembly Autom*. 38, 539–547. doi: 10.1108/AA-02-2018-019

Chinimilli, P. T., Subramanian, S. C., Redkar, S., and Sugar, T. (2019). "Human locomotion assistance using two-dimensional features based adaptive oscillator," in *2019 Wearable Robotics Association Conference (WearRAcon)* (Scottsdale, AZ: IEEE), 92–98. doi: 10.1109/WEARRACON.2019.8719628

Ding, M., Nagashima, M., Cho, S. G., Takamatsu, J., and Ogasawara, T. (2020). Control of walking assist exoskeleton with time-delay based on the prediction of plantar force. *IEEE Access* 8, 138642–138651. doi: 10.1109/ACCESS.2020.3010644

Esquenazi, A., Talaty, M., Packel, A., and Saulino, M. (2012). The rewalk powered exoskeleton to restore ambulatory function to individuals with thoracic-level motor-complete spinal cord injury. *Am. J. Phys. Med. Rehabil*. 91, 911–921. doi: 10.1097/PHM.0b013e318269d9a3

Fontana, M., Vertechy, R., Marcheschi, S., Salsedo, F., and Bergamasco, M. (2014). The body extender: a full-body exoskeleton for the transport and handling of heavy loads. *IEEE Robot. Autom. Mag*. 21, 34–44. doi: 10.1109/MRA.2014.2360287

Gams, A., Ijspeert, A. J., Schaal, S., and Lenarčič, J. (2009). On-line learning and modulation of periodic movements with nonlinear dynamical systems. *Auton. Robots* 27, 3–23. doi: 10.1007/s10514-009-9118-y

Guo, W., Qiu, S., Zha, F., Deng, J., Wang, X., and Chen, F. (2019). A novel active balance assistive control strategy based on virtual stiffness model of xcom. *Assembly Autom*. 40, 132–142. doi: 10.1108/AA-10-2018-0159

Huang, R., Cheng, H., Guo, H., Lin, X., and Zhang, J. (2018). Hierarchical learning control with physical human-exoskeleton interaction. *Inform. Sci*. 432, 584–595. doi: 10.1016/j.ins.2017.09.068

Huang, R., Cheng, H., Qiu, J., and Zhang, J. (2019). Learning physical human-robot interaction with coupled cooperative primitives for a lower exoskeleton. *IEEE Trans. Autom. Sci. Eng*. 16, 1566–1574. doi: 10.1109/TASE.2018.2886376

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., and Schaal, S. (2013). Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput*. 25, 328–373. doi: 10.1162/NECO_a_00393

Jamwal, P. K., Hussain, S., and Ghayesh, M. H. (2020). Robotic orthoses for gait rehabilitation: an overview of mechanical design and control strategies. *Proc. Inst. Mech. Eng. H* 234, 444–457. doi: 10.1177/0954411919898293

Kalita, B., Narayan, J., and Dwivedy, S. K. (2020). Development of active lower limb robotic-based orthosis and exoskeleton devices: a systematic review. *Int. J. Soc. Robot*. doi: 10.1007/s12369-020-00662-9. [Epub ahead of print].

Kazerooni, H., Racine, J. L., Huang, L., and Steger, R. (2005). "On the control of the Berkeley lower extremity exoskeleton (BLEEX)," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation* (Barcelona: IEEE), 4353–4360. doi: 10.1109/ROBOT.2005.1570790

Kim, K., Yu, C. H., Jeong, G. Y., Heo, M., and Kwon, T. K. (2013). Analysis of the assistance characteristics for the knee extension motion of knee orthosis using muscular stiffness force feedback. *J. Mech. Sci. Technol*. 27, 3161–3169. doi: 10.1007/s12206-013-0837-9

Krausz, N. E., and Hargrove, L. J. (2019). A survey of teleceptive sensing for wearable assistive robotic devices. *Sensors* 19:5238. doi: 10.3390/s19235238

Kumar, P. (1985). Theory and practice of recursive identification. *IEEE Trans. Autom. Control* 30, 1054–1056. doi: 10.1109/TAC.1985.1103802

Li, M., Deng, J., Zha, F., Qiu, S., Wang, X., and Chen, F. (2018). Towards online estimation of human joint muscular torque with a lower limb exoskeleton robot. *Appl. Sci*. 8:1610. doi: 10.3390/app8091610

Liang, C., and Hsiao, T. (2020). Admittance control of powered exoskeletons based on joint torque estimation. *IEEE Access* 8, 94404–94414. doi: 10.1109/ACCESS.2020.2995372

Lorenzo, G., Simona, C., Andrea, P., Raffaele, M. L., Silvestro, M., and Nicola, V. (2018). Gastrocnemius myoelectric control of a robotic hip exoskeleton can reduce the user's lower-limb muscle activities at push off. *Front. Neurosci*. 12:71. doi: 10.3389/fnins.2018.00071

Lu, R., Li, Z., Su, C. Y., and Xue, A. (2014). Development and learning control of a human limb with a rehabilitation exoskeleton. *IEEE Trans. Ind. Electron*. 61, 3776–3785. doi: 10.1109/TIE.2013.2275903

Mosher, R. S. (1967). "Handyman to Hardiman," in *1967 Automotive Engineering Congress and Exposition* (Michigan). doi: 10.4271/670088

Nagarajan, U., Aguirre-Ollinger, G., and Goswami, A. (2016). Integral admittance shaping: a unified framework for active exoskeleton control. *Robot. Auton. Syst*. 75, 310–324. doi: 10.1016/j.robot.2015.09.015

Oh, S., Baek, E., Song, S. K., Mohammed, S., Jeon, D., and Kong, K. (2015). A generalized control framework of assistive controllers and its application to lower limb exoskeletons. *Robot. Auton. Syst*. 73, 68–77. doi: 10.1016/j.robot.2014.10.001

Ortiz, M., Iáñez, E., Contreras-Vidal, J. L., and Azorín, J. M. (2020). Analysis of the EEG rhythms based on the empirical mode decomposition during motor imagery when using a lower-limb exoskeleton. A case of study. *Front. Neurorobot*. 14:48. doi: 10.3389/fnbot.2020.00048

Qiu, S., Guo, W., Zha, F., Wang, X., Sheng, W., Chen, F., et al. (2020). "Conditions for active assistance control of exoskeleton robot," in *2020 5th International Conference on Advanced Robotics and Mechatronics (ICARM)* (Shenzhen: IEEE), 220–227. doi: 10.1109/ICARM49381.2020.9195381

Quintero, H., Farris, R., Hartigan, C., Clesson, I., and Goldfarb, M. (2011). A powered lower limb orthosis for providing legged mobility in paraplegic individuals. *Top. Spinal Cord Inj. Rehabil*. 17, 25–33. doi: 10.1310/sci1701-25

Righetti, L., Buchli, J., and Ijspeert, A. J. (2006). Dynamic hebbian learning in adaptive frequency oscillators. *Phys. D Nonlin. Phenom*. 216, 269–281. doi: 10.1016/j.physd.2006.02.009

Ronsse, R., Lenzi, T., Vitiello, N., Koopman, B., Asseldonk, E. V., Rossi, S. M. M. D., et al. (2011). Oscillator-based assistance of cyclical movements: model-based and model-free approaches. *Med. Biol. Eng. Comput*. 49:1173. doi: 10.1007/s11517-011-0816-1

Ronsse, R., Vitiello, N., Lenzi, T., Van Den Kieboom, J., Carrozza, M. C., and Ijspeert, A. J. (2010). "Adaptive oscillators with human-in-the-loop: proof of concept for assistance and rehabilitation," in *2010 3rd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics* (Tokyo: IEEE), 668–674. doi: 10.1109/BIOROB.2010.5628021

Ruiz Garate, V., Parri, A., Yan, T., Munih, M., Molino Lova, R., Vitiello, N., et al. (2017). Experimental validation of motor primitive-based control for leg exoskeletons during continuous multi-locomotion tasks. *Front. Neurorobot*. 11:15. doi: 10.3389/fnbot.2017.00015

Schaal, S. (2006). "Dynamic movement primitives-a framework for motor control in humans and humanoid robotics," in *Adaptive Motion of Animals and Machines*, eds H. Kimura, K. Tsuchiya, A. Ishiguro, and H. Witte (Tokyo: Springer), 261–280. doi: 10.1007/4-431-31381-8_23

Seo, K., Hyung, S., Choi, B. K., Lee, Y., and Shim, Y. (2015). "A new adaptive frequency oscillator for gait assistance," in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA: IEEE), 5565–5571.

Seo, K., Kim, K., Park, Y. J., Cho, J. K., Lee, J., Choi, B., et al. (2018). "Adaptive oscillator-based control for active lower-limb exoskeleton and its metabolic impact," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Prague: IEEE), 6752–6758. doi: 10.1109/ICRA.2018.8460841

Witte, K. A., Fiers, P., Sheets-Singer, A. L., and Collins, S. H. (2020). Improving the energy economy of human running with powered and unpowered ankle exoskeleton assistance. *Sci. Robot*. 5:eaay9108. doi: 10.1126/scirobotics.aay9108

Yang, C., Chen, C., He, W., Cui, R., and Li, Z. (2018). Robot learning system based on adaptive neural control and dynamic movement primitives. *IEEE Trans. Neural Netw. Learn. Syst*. 30, 777–787. doi: 10.1109/TNNLS.2018.2852711

Young, A. J., and Ferris, D. P. (2017). State of the art and future directions for lower limb robotic exoskeletons. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 171–182. doi: 10.1109/TNSRE.2016.2521160

Young, A. J., Hannah, G., and Ferris, D. P. (2017). A biomechanical comparison of proportional electromyography control to biological torque control using a powered hip exoskeleton. *Front. Bioeng. Biotechnol.* 5:37. doi: 10.3389/fbioe.2017.00037

Zeng, C., Yang, C., Cheng, H., Li, Y., and Dai, S. L. (2020). Simultaneously encoding movement and sEMG-based stiffness for robotic skill learning. *IEEE Trans. Ind. Inform*. 17, 1244–1252. doi: 10.1109/TII.2020.2984482

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

ORIGINAL RESEARCH
published: 21 May 2021
doi: 10.3389/fnbot.2021.651177

Development of an Active Cable-Driven, Force-Controlled Robotic System for Walking Rehabilitation

Juan Fang[1]*, *Michael Haldimann*[1], *Laura Marchal-Crespo*[2,3] *and Kenneth J. Hunt*[1]

[1] *Division of Mechanical Engineering, Department of Engineering and Information Technology, Institute for Rehabilitation and Performance Technology, Bern University of Applied Sciences, Burgdorf, Switzerland,* [2] *Department of Cognitive Robotics, Delft University of Technology, Delft, Netherlands,* [3] *Motor Learning and Neurorehabilitation Laboratory, ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*

Edited by:
Shuai Li,
Swansea University, United Kingdom

Reviewed by:
Wenbin Chen,
Huazhong University of Science and Technology, China
Xikai Tu,
North Carolina State University, United States

***Correspondence:**
Juan Fang
juan.fang@bfh.ch

Received: *08 January 2021*
Accepted: *08 April 2021*
Published: *21 May 2021*

Citation:
Fang J, Haldimann M, Marchal-Crespo L and Hunt KJ (2021) Development of an Active Cable-Driven, Force-Controlled Robotic System for Walking Rehabilitation.
Front. Neurorobot. 15:651177.
doi: 10.3389/fnbot.2021.651177

In a parallel development to traditional rigid rehabilitation robotic systems, cable-driven systems are becoming popular. The robowalk expander product uses passive elastic bands in the training of the lower limbs. However, a well-controlled assistance or resistance is desirable for effective walking relearning and muscle training. To achieve well-controlled force during locomotion training with the robowalk expander, we replaced the elastic bands with actuator-driven cables and implemented force control algorithms for regulation of cable tensions. The aim of this work was to develop an active cable-driven robotic system, and to evaluate force control strategies for walking rehabilitation using frequency-domain analysis. The system parameters were determined through experiment-assisted simulation. Then force-feedback lead controllers were developed for static force tracking, and velocity-feedforward lead compensators were implemented to reduce velocity-related disturbances during walking. The technical evaluation of the active cable-driven robotic system showed that force-feedback lead controllers produced satisfactory force tracking in the static tests with a mean error of 5.5%, but in the dynamic tests, a mean error of 13.2% was observed. Further implementation of the velocity-feedforward lead compensators reduced the force tracking error to 9% in dynamic tests. With the combined control algorithms, the active cable-driven robotic system produced constant force within the four cables during walking on the treadmill, with a mean force-tracking error of 10.3%. This study demonstrates that the force control algorithms are technically feasible. The active cable-driven, force-controlled robotic system has the potential to produce user-defined assistance or resistance in rehabilitation and fitness training.

Keywords: cable-driven robots, force control, dynamic modeling, frequency-domain analysis, velocity compensation, rehabilitation robotic systems

INTRODUCTION

Diseases of or injuries to the cerebrovascular system often result in impaired sensorimotor function. Neural plasticity of the central nervous system supports that new neural connections can be developed through intensive, repetitive, and task-specific training (Maier et al., 2019). Advances in robotic technology have brought several rehabilitation robotic systems to the market. Most rehabilitation robotic systems use rigid mechanisms to assist training for the upper or lower limbs (Hesse et al., 2003; Hidler et al., 2005), including exoskeleton-based systems (e.g., Armeo Power and Lokomat from Hocoma, Switzerland) and end-effector based systems (e.g., MIT-MANUS from Interactive Motion Technologies Inc., Cambridge, Massachusetts, USA, and the G-EO system from Reha Technology, Switzerland). Along with these rigid rehabilitation robotic systems, cable-aided rehabilitation therapy has emerged in recent years with advantages including low moment of inertia, high power output, and soft human-machine contact (Rosati et al., 2017). Cable-driven rehabilitation systems available on the market include Diego (Tyromotion GmbH, Austria) for upper limb rehabilitation (Aprile et al., 2019), and the robowalk expander system (h/p/cosmos Sports & Medical GmbH, Germany) for lower limb training (Schulze et al., 2019).

The robowalk expander product is a system using passive elastic bands in the training of the lower limbs. It includes two supporting frames which can be fixed on a treadmill, and four rubber bands which can be attached to the thigh, shank, or ankle segments via cuffs. Depending on the band setup and pre-tension adjustment, the robowalk expander system can provide assistance or resistance to the legs during walking/running on a treadmill (Schulze et al., 2019). For patients who have walking impairments, assistance from the elastic bands can help them to walk on the treadmill with improved speed and stride length (Beckmann-Hemmers, 2012; Jöllenbeck and Pietschmann, 2018). However, the passive assistance provided by the unactuated elastic bands is dependent on the leg position. Due to the factors of different temperatures and varying rubber elasticity after long-term usage (Beckmann-Hemmers, 2012), the force provided by the elastic band might change non-linearly with the leg position. A well-controlled assistance or resistance, which allows patients to learn how to exert force to walk correctly and thereby to bring improved walking skills (Wu et al., 2011), is believed to be desirable for effective walking relearning and muscle training.

In order to provide well-controlled guidance, several cable-driven rehabilitation robotic systems were developed, with various control algorithms investigated. The cable-driven locomotor trainer (CaLT) used four cable-driven units to control the end-effector ankle segments, where two units were connected jointly on one ankle segment. Using a proportional-derivative (PD) controller for position control, CaLT produced assistance and resistance during walking without changing the ankle trajectory (Wu et al., 2011). Tests on patients with incomplete spinal cord injury showed that increased overground walking speed was obtained after 8 weeks of training using CaLT. Agrawal et al. designed a cable-driven active leg exoskeleton (C-ALEX) for walking rehabilitation (Jin et al., 2015). Cable-driven active leg exoskeleton included two cable-driven units for the right thigh and another two cable-driven units for the right shank segment. Using proportional–integral–derivative (PID) algorithms, friction compensation and a feedforward element, C-ALEX produced a well-controlled ankle trajectory in the right leg during walking. The Robotic Physical Exercise and System (ROPES) included two spring-assisted cable-driven units (one connected to the left thigh, and the other to the left shank) and two cable-driven units (both connected to the left foot) (Alamdari and Krovi, 2015a). Using PD algorithms, forward kinematics, and gravity compensation for position control, ROPES produced walking-like movement in the left hip, knee, and ankle joints.

Apart from trajectory tracking, several studies investigated control strategies for force regulation in cable-driven robotic systems, where motion-induced disturbances require additional compensation. Zou et al. developed a lower limb rehabilitation system with eight cable-driven units, where two units were used to control each thigh and each shank of both legs (Zou et al., 2019). Frequency-domain analysis revealed that the system structure including the moment of inertia and cable stiffness influenced the dynamics of the cable-driven robotic system significantly (Zou et al., 2019). The force control strategy comprised an integral term, a lead compensator and a feedforward correction element. Experimental results from static tests showed that the system produced forces that tracked sinusoidal force references with different frequencies. In contrast to force control in a static situation, force control during movement had further disturbance from the actuator velocity (Zou et al., 2018a, 2019). To compensate such velocity-related disturbances, a force compensator using the structure invariance principle was investigated (Zhang et al., 2017; Zou et al., 2018a,b; Wang et al., 2020), while disturbance observers for force control were also implemented (Jung and Bae, 2016; Wang et al., 2019). These force control algorithms, combined with disturbance compensation, often result in complex control strategies.

To achieve well-controlled force during locomotion training with the robowalk expander, we replaced the elastic bands with actuator-driven cables and implemented novel force control algorithms for regulation of cable tensions. It was observed here that movement-induced disturbances influenced the force-tracking performance significantly. Velocity disturbance is a common issue in force-controlled cable-driven robotic systems, but how to compensate the velocity disturbance is the key issue. The disturbance compensators described in the literature (Jung and Bae, 2016; Zhang et al., 2017; Zou et al., 2018a,b; Wang et al., 2019, 2020), due to different system dynamics, could not be easily applied in the current study. The current work presented a novel approach to address this issue. By analyzing the kinetic model of the cable-driven system and estimating the system parameters, this paper introduced lead control algorithms for force control and also for velocity compensation. The methods of how to determine the suitable parameters for the lead control algorithms were deduced in detail using frequency-domain analysis. A comprehensive search in the databases of PubMed, Web of Science and Google Scholar using search items of "force control AND cable-driven," "lead controller OR lead compensator," "velocity compensation OR speed compensation" failed to yield any similar control strategies developed in the current work. The lead control algorithms were straightforwardly developed and produced satisfactory force tracking with negligible velocity-related disturbances during walking. After being validated in the current study, this novel control strategy has potential to be generally applied in force-controlled cable-driven robotic systems.

The aim of this work was to develop an active cable-driven robotic system, and to evaluate force control strategies for walking rehabilitation using frequency-domain analysis. The active cable-driven, force-controlled robotic system has the potential to produce user-defined assistance or resistance in rehabilitation and fitness training.

METHODS

The Active Cable-Driven Robotic System

Two robowalk expander units were fixed on a treadmill (model Venus, h/p/cosmos Sports & Medical GmbH, Germany), with one at the front and one at the rear (**Figure 1A**). The treadmill speed was controlled by a computer via an RS-232 serial port. In order to achieve well-controlled active forces, four rigid cables were used and four motor-drum actuators were designed (**Figure 1B**) to control the cable tensions. Two cable-driven units were mounted on the front-right support frame of the expander unit, while the other two were mounted on the rear-right support frame (**Figures 1A,B**).

Two actuator assemblies (DC motor model RE50, gearhead with a ratio of 91, Encoder HEDL 5540; maxon motor, Switzerland) pulled the right thigh while moving forwards and backwards, and two actuator assemblies (DC motor model RE40, gearhead with a ratio of 113, Encoder HEDL 5540; maxon motor, Switzerland) pulled the right shank similarly. Two servocontrollers (EPOS4 70/15; maxon motor, Switzerland) sent torque commands to the thigh actuators, and two further servocontrollers (EPOS4 50/5; maxon motor, Switzerland) controlled the shank actuators. A shunt regulator (DSR 70/30; maxon motor, Switzerland) was used to buffer energy generated when the motors worked as generators. For safety, an emergency stop was fixed on the right horizontal handrail (see **Supplementary Material** "Walk with the active cable-driven robotic system. mp4"). The servocontrollers provide a Safe Torque Off (STO) function, which brings the actuator to a torque-free, safe condition when the stop command is received. To enhance safety, the wires for the inputs of the STO function were connected to the emergency stop of the treadmill. By this means, activating the emergency stop button stops the treadmill immediately, and simultaneously sets the four actuators into the zero-torque-output safe mode.

Force sensors with range −500 to 500 N were connected in series in each cable close to the leg fixation points (**Figure 1A**). The two force sensors for the thigh segment (Bengbu Sensor Company, China; resolution 0.00012 N) were recorded using an analog input card (ELM3144-0000, Beckhoff Automation GmbH & Co, Germany). The two force sensors for the shank segment (KD40S, Transmetra GmbH, Switzerland; resolution 0.00003 N) were recorded using an amplifier/converter interface (EL3356-0010, Beckhoff Automation GmbH & Co, Germany).

The real-time control program was developed using TwinCAT3 software (Beckhoff Automation GmbH & Co., Germany) running in a laptop at a sample rate of 1,000 Hz (**Figure 1C**). The control program sent speed commands to and received the speed information from the treadmill via a serial interface (KL6001, Beckhoff Automation GmbH & Co., Germany), which was installed on the EtherCAT Bus Coupler (BK1120, Beckhoff Automation GmbH & Co., Germany). The force information was transferred to the control program via an EtherCAT coupler (EK1100, Beckhoff Automation GmbH & Co., Germany). Four EtherCAT cards (Part No. 581245, maxon motor, Switzerland) were used to establish the communication between the servocontrollers and the TwinCAT program (**Figure 1C**). TC3 Controller Toolbox in TwinCAT3 was used to develop the control algorithms. A live-view Human Machine Interface (HMI) was developed using TC3 HMI Engineering toolbox, which allowed operation of the active cable-driven robotic system via a tablet. TC3 Measurement Toolbox was used to present, analyze, save, and export the experimental data, including the reference and actual force, actuator control signals, and velocities.

Control Development

Dynamics of the Active Cable-Driven Robotic System

A model of the active cable-driven robotic system was developed in Matlab/Simulink (The Mathworks, Inc., Natick, USA). The cable was considered as a spring with a stiffness of k_s, while the weight and damping of the cable were neglected. The dynamic model of one cable-actuator unit in the active cable-driven robotic system is depicted (**Figure 2**).

T_1 is the total torque that the actuator (the motor and the gearbox) generates. The servocontroller sent the torque command as a percentage of the maximal torque T_{mp}. The transfer function of the actuator linking the torque command T_{mp} to the output torque T_1, was expressed as a linear first-order transfer function:

$$T_{mp} \rightarrow T_1 : G_{T_1 T_{mp}}(s) = \frac{k_a}{\tau_a s + 1}, \tag{1}$$

where τ_a is the time constant and was determined by the setting of the servocontrollers, while k_a is the steady-state gain and was determined using the motor and gearbox specifications. T_2 is the torque used to compensate the viscous component of the actuator, which is a function of the final angular velocity of the actuator ω:

$$\omega \rightarrow T_2 : G_{T_2 \omega} = J_a D_f + B_a, \tag{2}$$

where J_a and B_a are the moment of inertia and damping of the actuator. D_f is the derivative term in combination with a low pass filter, which is:

$$D_f(s) = \frac{s}{\frac{1}{\omega_{lp}} s + 1}. \tag{3}$$

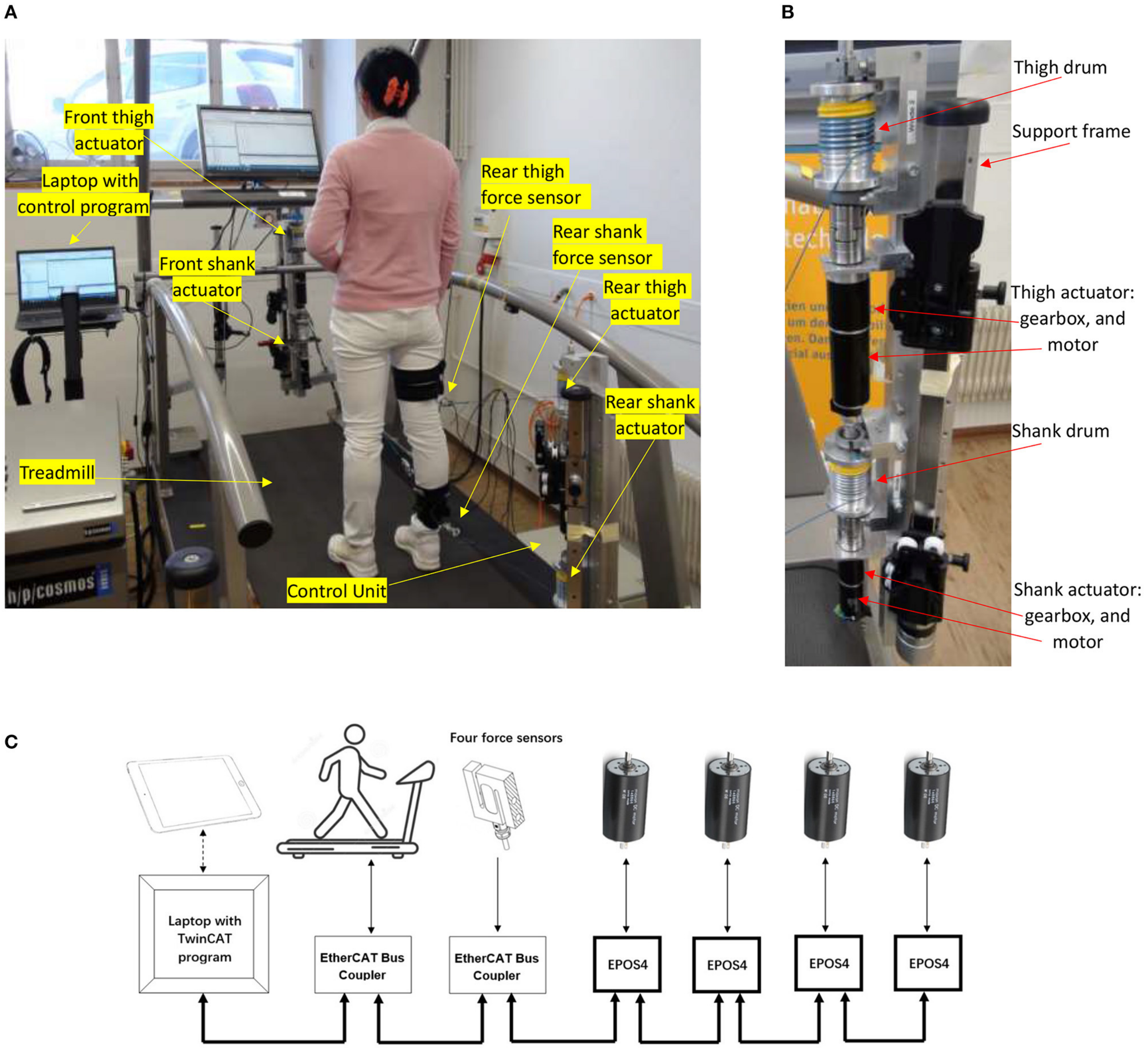


FIGURE 1 | The active cable-driven robotic system. **(A)** The physical system with a participant. The modified robowalk expander units are at the front-right and rear-right. **(B)** Modified expander unit with support frame and two actuators (rear-right). **(C)** The communication structure.

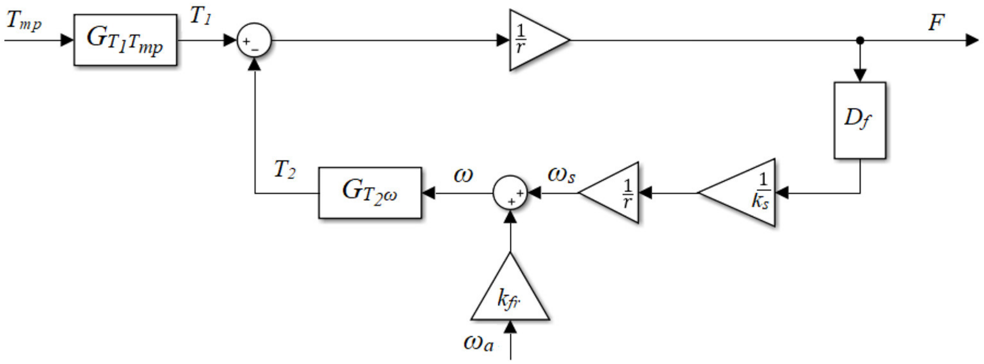


FIGURE 2 | Block diagram of the active cable-driven robotic system.

The cut-off frequency ω_{lp} was set to be 5×10^4 rad/s ($f_{lp} \approx 8{,}000$ Hz), which is much higher than the system sampling frequency of 1,000 Hz, therefore the filter will not influence simulation of the system dynamics within the frequencies that are of interest in this study.

The final angular velocity of the actuator ω is primarily influenced by two variables: the force-related velocity ω_s and the motion-produced velocity ω_a. The force F on the spring with a stiffness of k_s results in an angular velocity of the actuator, ω_s:

$$F \rightarrow \omega_s : G_{\omega_s F} = \frac{D_f}{r k_s}, \tag{4}$$

where D_f is the same derivative term in combination with a low pass filter, as described in Equation (3). The leg motion, i.e., the walking movement, is related to the angular velocity of the actuator. This velocity was approximated by the measured actuator velocity ω_a, which is the motor velocity divided by the gear ratio.

Due to the contact stress between the cable and the drum, the system has complex friction dynamics. The dynamic friction in general is related to the leg-motion induced velocity ω_a. Thus, a velocity scale k_{fr} was introduced (**Figure 2**). This scaled velocity passes through the actuator dynamics, i.e., Equation (2), representing the friction between the cable and the drum, apart from the leg-motion required torque. Thus, the final angular velocity ω is

$$\omega = \omega_s + k_{fr}\omega_a. \tag{5}$$

The force output F of the active cable-driven robotic system, i.e., the cable tension is

$$F = \frac{T_1 - T_2}{r}, \tag{6}$$

where r is the drum radius.

The active cable-driven robotic system has two inputs driving the final force output, which are the torque command of the actuator T_{mp} and the actuator velocity ω_a. This study considered the torque command T_{mp} as the main input, while the actuator velocity ω_a was regarded as a plant disturbance. The transfer function from the actuator control signal to the force output, which is regarded as the plant open-loop transfer function, is:

$$T_{mp} \rightarrow F : G_{FT_{mp}} = \frac{G_{T_1 T_{mp}}}{r + \frac{D_f G_{T_2 \omega}}{k_s r}}. \tag{7}$$

The transfer function from the disturbance, i.e., the angular velocity of the actuator, to the force output is:

$$\omega_a \rightarrow F : G^A_{F\omega_a} = -\frac{k_{fr} G_{T_2 \omega}}{r + \frac{D_f G_{T_2 \omega}}{k_s r}}. \tag{8}$$

System Parameter Determination

The parameters τ_a and k_a (Equation 1) were determined using the hardware controller settings, while J_a (Equation 2) and r (Equations 4, 6) were determined from the geometric configuration. The remaining parameters B_a (Equation 2), k_s (Equation 4), and k_{fr} (Equation 5) were determined using the experiment-assisted simulation approach described below.

In order to determine k_{fr}, the system has to move at different speeds. However, due to the high friction, it is difficult to pull the cable in the open-loop active cable-driven robotic system to make the actuators move. A tentative closed-loop force controller (**Figure 3**) was therefore developed for each actuator. Conventional closed-loop controllers such as PD algorithms were reported in studies (Marchal-Crespo et al., 2012; Alamdari and Krovi, 2015b). A pretest using a PD controller showed steady-state errors, therefore a PID controller was implemented for each actuator as the tentative controller C_{pid} in this phase (**Figure 3**):

$$C_{pid}(s) = k_p + \frac{k_i}{s} + k_d \frac{s}{\tau_d s + 1}, \tag{9}$$

where k_p, k_i, and k_d are the gains for the proportional, integral, and derivative terms. The derivative term was implemented with a low pass filter, where τ_d is the time constant of the low pass filter.

In the parameter determination phase, static and dynamic tests were implemented on a female participant with a height of 1.6 m and a mass of 53 kg. During the test setup, she stood on the treadmill with her right thigh and shank segments attached to the actuator-driven cables via cuffs, while her left leg was free (**Figure 1A**). Each cable was connected to an independent cuff to prevent interaction between the front and rear actuators. Square-wave and sinusoidal force references were implemented. By resisting the force from the cables throughout the test, the participant tried to stand still on the static treadmill in the static test, while in the dynamic test she tried to walk normally on the treadmill which ran at different speeds.

A Matlab/Simulink model simulating four actuators with the tentative PID controllers of the active cable-driven robotic system was developed (**Figure 3**). The experiment-used reference F_{ref} and the experiment-recorded actuator velocity ω_a were used as the inputs in the model. Using the calculated parameters τ_a, k_a, J_a, and r, and trying different possible values of B_a, k_s, and k_{fr}, the model simulated the dynamics of the PID-controlled active cable-driven robotic system. The simulated force output F_{sim} was then compared with the experimental force $F_{\mathrm{exp.}}$ It should be noted that the current simulation task was to determine the suitable B_a, k_s, and k_{fr}, which yielded the simulated force F_{sim} as close as possible to the experimental data F_{exp}. As the active cable-driven robotic system focused on walking training, the goodness of fit of the force output during the dynamic test was specifically investigated using a normalized root-mean-square error (NRMSE) on an evaluation interval from t_o to t_1:

$$\text{NRMSE} = (1 - \sqrt{\frac{\sum_{t=t_0}^{t_1} \left(F_{\mathrm{exp}}(t) - F_{sim}(t)\right)^2}{\sum_{t=t_0}^{t_1} \left(F_{\mathrm{exp}}(t) - \overline{F_{\mathrm{exp}}(t)}\right)^2}}) \times 100\%, \tag{10}$$

where $\overline{F_{\mathrm{exp}}}$ is the mean of F_{exp} between t_o and t_1. The values of B_a, k_s, and k_{fr} that yielded an NRMSE higher than 60%

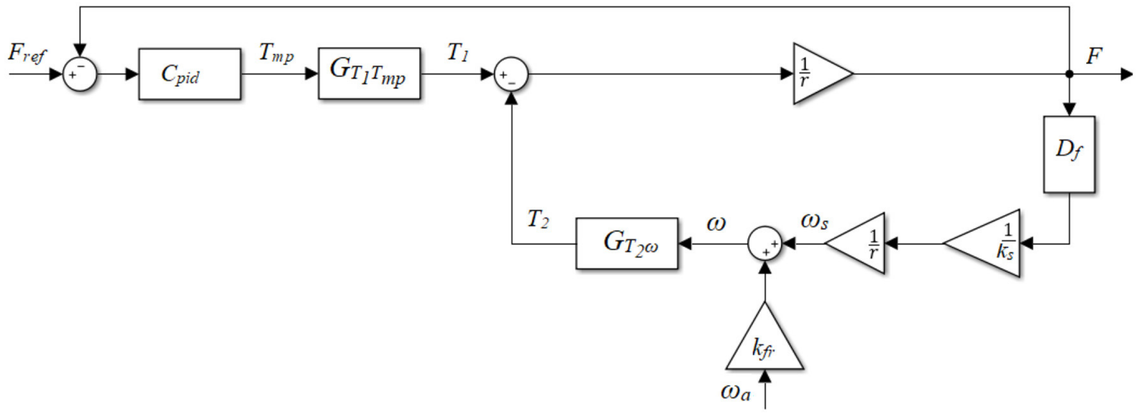


FIGURE 3 | Block diagram of the plant with the tentative PID controller.

were considered suitable parameters for the active cable-driven robotic system.

Force Control Strategy

The plant model was used to develop accurate force control strategies. The force control structure (**Figure 4**) included a force-feedback lead controller and a velocity-feedforward lead compensator for each actuator.

Force-Feedback Lead Controller

As the tests in the parameter determination phase showed an obvious delay (e.g., **Figure 5G**), a lead compensator C_F was developed initially (**Figure 4**, with $C_{\omega_a} = 0$):

$$C_F(s) = k_F \frac{\tau_F s + 1}{a_F \tau_F s + 1}, \tag{11}$$

where k_F is the gain, and $a_F < 1$. The controller parameters a_F and τ_F are linked with the center frequency f_F by

$$\tau_F = \frac{1}{f_F \sqrt{a_F}}. \tag{12}$$

The center frequency f_F is the frequency when the lead compensator C_F has the largest phase. The transfer function from the reference F_{ref} to the force output F is therefore:

$$F_{ref} \rightarrow F : G_{FF_{ref}} = \frac{C_F G_{FT_{mp}}}{1 + C_F G_{FT_{mp}}}. \tag{13}$$

Furthermore, the transfer function from F_{ref} to the control signal T_{mp} is:

$$F_{ref} \rightarrow T_{mp} : G_{T_{mp}F_{ref}} = \frac{C_F}{1 + C_F G_{FT_{mp}}}. \tag{14}$$

The transfer function from the disturbance term (angular velocity ω_a) to the force output F is:

$$\omega_a \rightarrow F : G^B_{F\omega_a} = \frac{G^A_{F\omega_a}}{1 + C_F G_{FT_{mp}}} = -\frac{k_{fr} G_{T_2\omega}}{G_{T_1T_{mp}} C_F + r + \frac{D_f G_{T_2\omega}}{k_s r}}. \tag{15}$$

To evaluate the stability of the feedback system, Nyquist analysis was performed using the loop gain G_{Lo}:

$$G_{L_o} = C_F G_{FT_{mp}}. \tag{16}$$

Apart from the gain and phase margins, the modulus margin was calculated, which is the minimal distance between the critical point $[-1, 0]$ and the Nyquist plot G_{Lo} (Aström and Murray, 2010).

Velocity-Feedforward Lead Compensator

The velocity compensator introduces an extra input T_{ω_a} to the torque command (**Figure 4**), with the aim of reducing the disturbance on the force output induced by the actuator velocity. In this case, the influence of T_{ω_a} on the force output is:

$$T_{\omega_a} \rightarrow F : G_{FT_{\omega_a}} = \frac{G_{FT_{mp}}}{1 + C_F G_{FT_{mp}}} = -\frac{G_{T_1T_{mp}}}{C_F G_{T_1T_{mp}} + r + \frac{D_f G_{T_2\omega}}{k_s r}}. \tag{17}$$

Experimental tests in section System Parameter Determination showed that the actual force output followed a similar shape of the angular velocity curve, but with some phase shifts (comparing **Figures 5D,F**, as well as **Figures 5J,L**). Therefore, the angular velocity was used as the input for the compensator channel, and a feedforward lead compensator C_{ω_a} was further implemented (**Figure 4**):

$$C_{\omega_a}(s) = k_{\omega_a} \frac{\tau_{\omega_a} s + 1}{a_{\omega_a} \tau_{\omega_a} s + 1}, \tag{18}$$

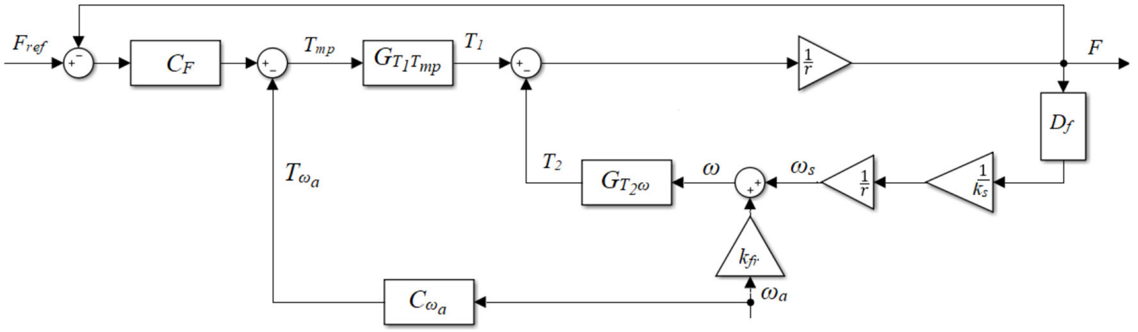


FIGURE 4 | Block diagram of the plant with the force-feedback lead controller C_F and the velocity-feedforward lead compensator C_{ω_a}.

where k_{ω_a} is the gain, and $a_{\omega_a} < 1$. The control parameters a_{ω_a}, τ_{ω_a} and the center frequency f_{ω_a} are interlinked:

$$\tau_{\omega_a} = \frac{1}{f_{\omega_a}\sqrt{a_{\omega_a}}}. \tag{19}$$

With this compensator C_{ω_a}, the transfer function of the velocity compensation channel is

$$\omega_a \to F : G^C_{F\omega_a} = G_{FT_{\omega_a}} C_{\omega_a}, \tag{20}$$

where $G_{FT_{\omega_a}}$ is given by Equation (17).

With the force-feedback controller and the velocity-feedforward compensator, the overall transfer functions from the angular velocity ω_a to the force output F and to the torque command T_{mp}, are, respectively:

$$\begin{aligned} \omega_a \to F : G^D_{F\omega_a} &= \frac{G^A_{F\omega_a} + C_{\omega_a} G_{FT_{mp}}}{1 + C_F G_{FT_{mp}}} \\ &= -\frac{G_{T_1T_{mp}} C_{\omega_a} + k_{fr} G_{T_2\omega}}{C_F G_{T_1T_{mp}} + r + \frac{D_f G_{T_2\omega}}{k_s r}}. \end{aligned} \tag{21}$$

$$\begin{aligned} \omega_a \to T_{mp} : G_{T_{mp}\omega_a} &= -\frac{C_F G^A_{F\omega_a} + C_{\omega_a}}{1 + C_F G_{FT_{mp}}} \\ &= \frac{k_{fr} G_{T_2\omega} - \frac{C_{\omega_a}\left(D_f G_{T_2\omega} + k_s r^2\right)}{k_s r C_F}}{G_{T_1T_{mp}} + \frac{D_f G_{T_2\omega} + k_s r^2}{k_s r C_F}}. \end{aligned} \tag{22}$$

Simulation and Evaluation of the Control Algorithms

A model was developed in Matlab/Simulink to design, simulate, and evaluate the force control algorithms for the four actuators (developed in section Force Control Strategy). Using the system parameters determined in section System Parameter Determination, two dynamic cases of the active cable-driven robotic system were simulated: (i) the nominal plant without consideration of the velocity disturbance; and (ii) the plant with consideration of the velocity disturbance. In simulation (i), only the closed-loop force-feedback controller C_F was simulated (**Figure 4**) with $C_{\omega_a} = 0$ and $\omega_a = 0$. The simulated force output F_{nom} served as the target for the force control. Simulation (ii) used the force-feedback lead controller C_F and the velocity-feedforward lead compensator C_{ω_a} (**Figure 4**), where the actuator velocity when the participant walked at 0.2 m/s was used as the velocity input ω_a. Different control parameters, i.e., k_F, a_F, and τ_F for the force-feedback controller C_F, and k_{ω_a}, a_{ω_a}, and τ_{ω_a} for the velocity-feedforward compensator C_{ω_a}, were simulated, with the aim of producing force outputs that well-tracked square-wave and sinusoidal reference forces. To evaluate the force tracking accuracy, the root-mean-square error (RMSE) between the simulated force output F_{sim} from simulation (ii) and the nominal force output F_{nom} from simulation (i) was calculated on an evaluation interval from t_o to t_1 as follows:

$$\mathrm{RMSE}_{F\mathrm{sim}} = \sqrt{\frac{1}{N}\sum_{t=t_0}^{t_1} \left(F_{\mathrm{sim}}(t) - F_{\mathrm{nom}}(t)\right)^2}, \tag{23}$$

where N is the number of data points in the interval t_o and t_1. When the simulated force output had an $\mathrm{RMSE}_{F\mathrm{sim}} < 5\,\mathrm{N}$, which was considered satisfactory, the control algorithms were implemented in the four actuators of the active cable-driven robotic system for evaluation.

The evaluation tests were performed on the same participant who performed the tests in section System Parameter Determination and included two sections: "static" tests where the participant stood still, and "dynamic" tests where the participant walked at 0.2 m/s. Both tests used the square-wave and sinusoidal force references and implemented two control algorithms: (1) using the force-feedback controllers C_F only (**Figure 4**, with $C_{\omega_a} = 0$), and (2) using the force-feedback controllers C_F and the velocity-feedforward compensators C_{ω_a} (**Figure 4**). Replacing F_{sim} with F_{exp} in Equation (23) yielded the difference between the experimental and nominal force, $\mathrm{RMSE}_{F\mathrm{exp}}$. To enable comparison between different force strategies, the experimental force-tracking error is presented in percentage terms after $\mathrm{RMSE}_{F\mathrm{exp}}$ was divided by the maximal reference force.

After the force control algorithms were evaluated to be feasible in four actuators, the participant finally walked at 0.2 m/s with a constant reference force within the four cables in the active cable-driven robotic system.

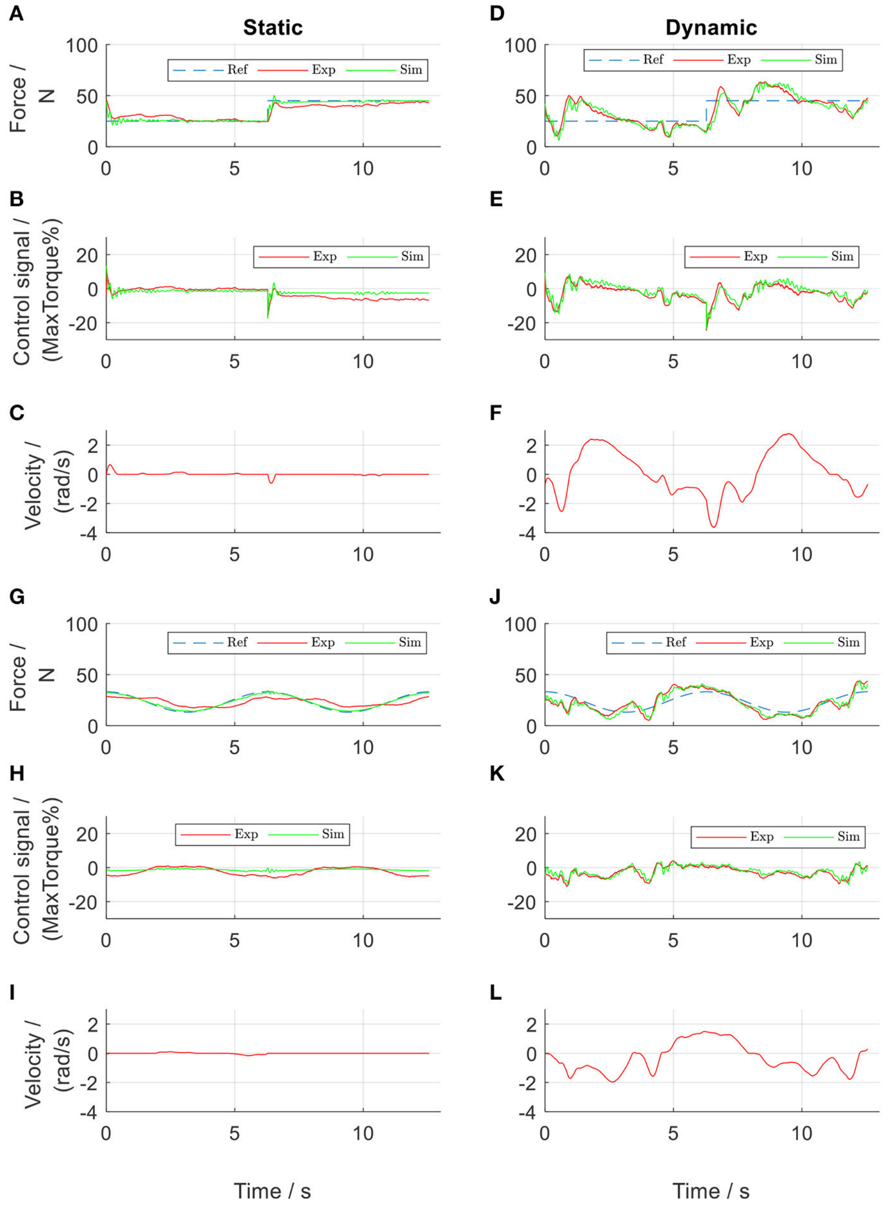


FIGURE 5 | PID-regulated force control of the front thigh actuator. Subplots **(A–F)** show results from square-wave reference test, while subplots **(G–L)** show results from sinusoidal reference test. Ref, references; Exp, experimental values; Sim, simulation results.

RESULTS

The results from the parameter determination tests, followed by the controller design and evaluation tests are presented. During the tests in the active cable-driven robotic system, the four actuators were controlled simultaneously. To reduce the figure numbers, **Figures 5–11** only show the representative results of the front thigh actuator, while the results of all actuators are presented in **Figures 12**, **13**. During evaluation of the force control algorithms (section Evaluation of the Force Control Algorithms), the simulated force output F_{sim} and the control signal $T_{mp\mathrm{sim}}$ are presented in figures using green lines, serving as a comparison for the experimental results, which are presented as red lines. The force output from the nominal plant, F_{nom}, is presented in black lines, serving as the target during the calculation of $\mathrm{RMSE}_{F\mathrm{exp}}$. The positive angular velocity of the actuator in **Figures 5C,F,I,L**, **10C,F,I,L**, **11C,F,I,L** means the right leg moved backwards.

Parameter Determination Tests

By trial and error, the parameters of the tentative PID controllers for both thigh actuators were tuned to be: $k_p = 0.4$, $k_i = 0.2$, $k_d = 0.04$, and $\tau_d = 0.1$. Those for both shank actuators were: $k_p = 0.6$, $k_i = 0.3$, $k_d = 0.06$, and $\tau_d = 0.1$. These PID controllers allowed implementation of a series of tests, which produced experimental data for determination of the system

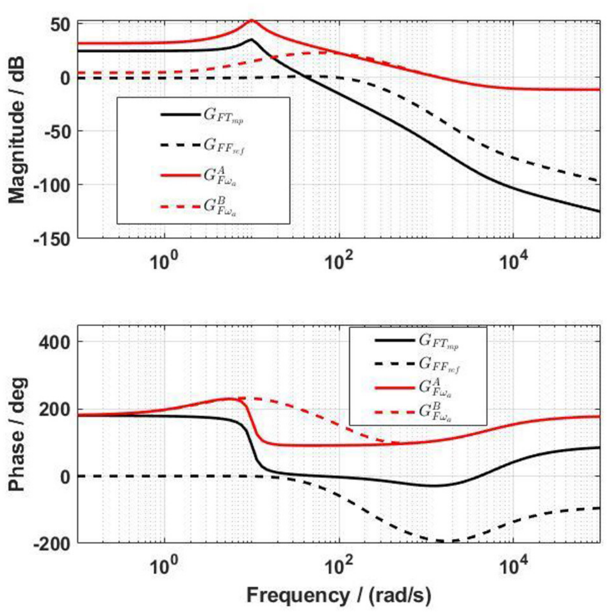


FIGURE 6 | Bode diagram of the transfer functions from the actuator control signal to the force output before ($G_{FT_{mp}}$) and after ($G_{FF_{ref}}$) including the force-feedback lead controller, and also the transfer functions from the angular velocity of the actuator to the force output before ($G^A_{F\omega_a}$) and after ($G^B_{F\omega_a}$) including the force-feedback lead controller.

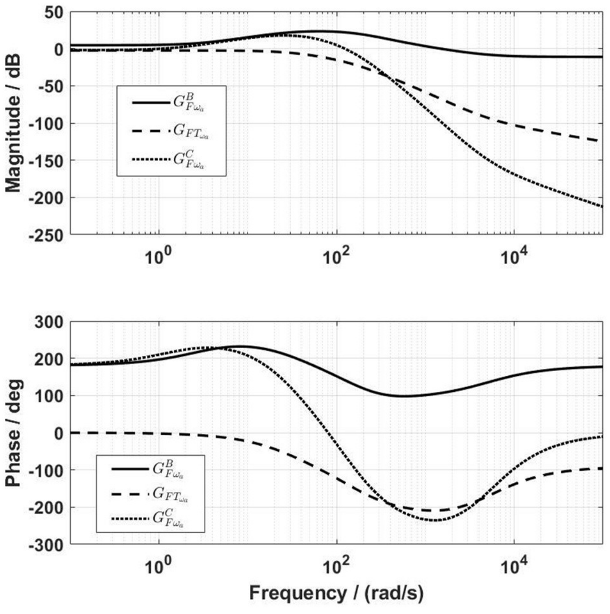


FIGURE 7 | Frequency characteristics of the compensation channel before ($G_{FT\omega_a}$) and after ($G^C_{F\omega_a}$) the inclusion of the velocity-feedforward lead compensator. The amplitude of $G^C_{F\omega_a}$ is close to $G^B_{F\omega_a}$ for frequencies up to 40 rad/s.

parameters. After repetitive comparisons between the simulated force F_{sim} and the experimental force F_{exp}, the system parameters were determined (**Table 1**). The PID-regulated results when the participant walked at 0.1 m/s are presented in **Figure 5**. It was observed that the model yielded more similar results in the dynamic tests than in the static tests. This might be due to static friction, which was neglected in the Simulink model. Nevertheless, the simulated force outputs (green lines) present similar curves with comparable amplitudes to those of the experimental results (red lines). The NRMSE between F_{sim} and F_{exp} during the dynamic tests with the sinusoidal reference in

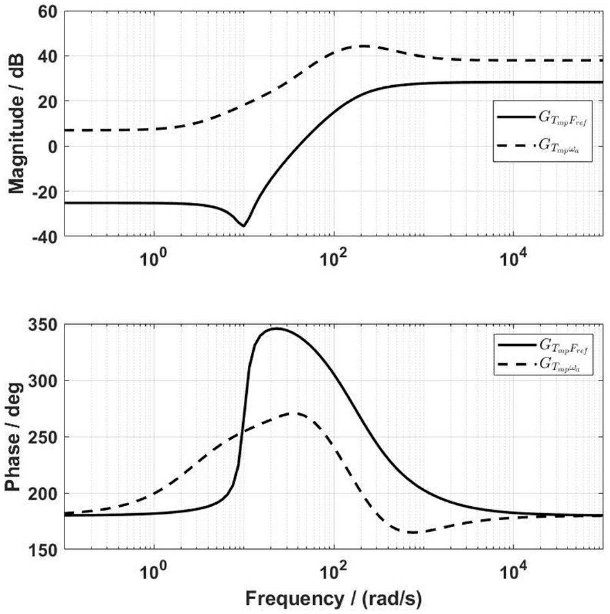


FIGURE 8 | Frequency characteristics of the transfer functions from the reference F_{ref} and from the angular velocity ω_a to the control signal T_{mp}.

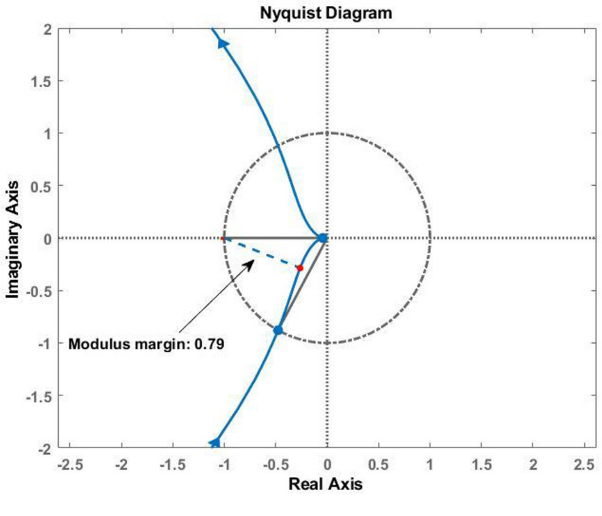


FIGURE 9 | Stability analysis of the closed-loop system using loop gain G_{Lo} (Equation 16). The gain margin is 27.4 dB (at 824 rad/s), and the phase margin is 61.5° (at 103 rad/s). The blue dashed line shows the modulus margin of 0.79.

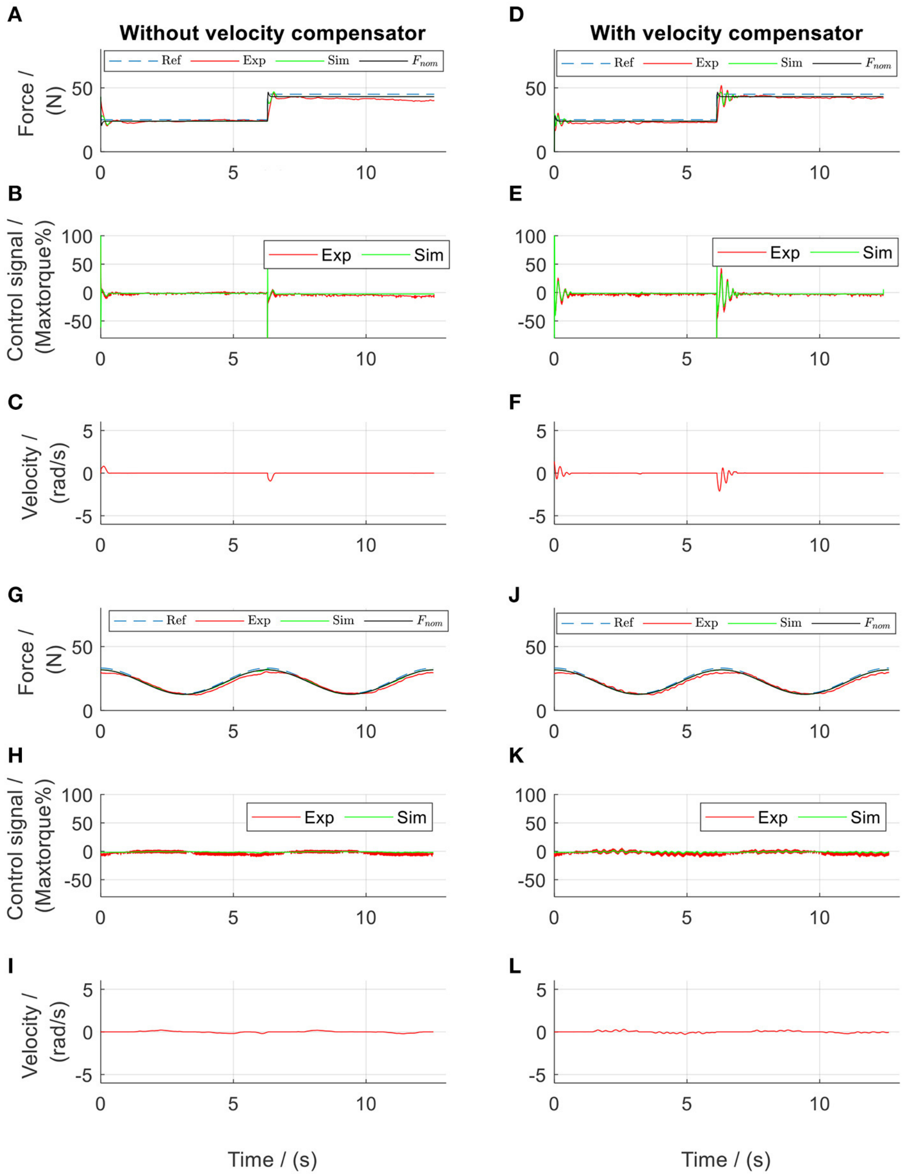


FIGURE 10 | Lead algorithms regulated force control of the front thigh actuator in the static tests. Subplots **(A–F)** show the results from square-wave reference test, while subplots **(G–L)** show the results from sinusoidal reference test. Ref, references; Exp, experimental values; Sim, simulation results; F_{nom}, the simulated target.

the front thigh actuator was 75 % (**Figure 5J** and **Table 2**). The four actuators showed a mean NRMSE ≥ 70 % (**Table 2**), which demonstrated that the determined system parameters in **Table 1** were satisfactory.

From the point of view of force control, the tentative PID controllers produced good force tracking during the static tests (**Figures 5A,G**). The experimental force showed a tracking error of 9.3% in **Figure 5A** and 15.7% in **Figure 5G**. This error was considered to come from the static friction. It should be noted that in the sinusoidal force control (**Figure 5G**), the force output showed a phase delay of 70.6° compared to the reference. Compared to the static tests, the PID controllers in the dynamic tests failed to produce satisfactory force tracking, in that the force output was either larger or smaller than the reference (an error of 24.0% in **Figure 5D** and 19.9% in **Figure 5J**). Comparison between **Figures 5D,F** and also comparison between **Figures 5J,L** both revealed that the force output was influenced by the walking velocity, in that the force curve took a similar shape of the velocity curve.

The actuator-driven cables were observed to be loose briefly at toe off and heel strike. These problems were more obvious when the walking speed increased. The participant described it as a difficult walking because of the continuously varying force on the leg.

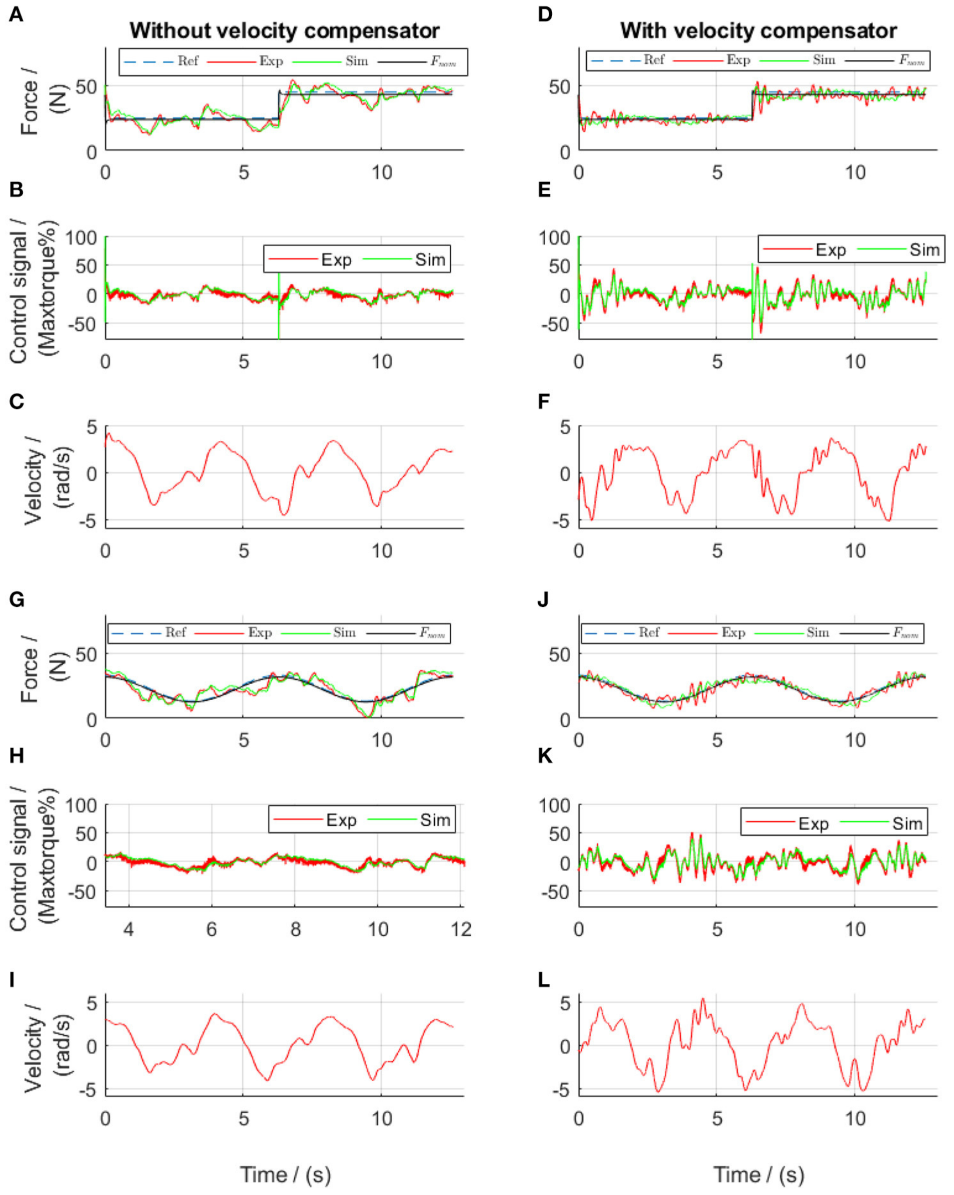


FIGURE 11 | Lead algorithms regulated force control of the front thigh actuator in the dynamic tests. Subplots **(A–F)** show the results from square-wave reference test, while subplots **(G–L)** show the results from sinusoidal reference test. Ref, references; Exp, experimental values; Sim, simulation results; F_{nom}, the simulated target.

Frequency-Domain Analysis for the Controller Development

Using the determined system parameters (**Table 1**), the transfer functions of the force output from the actuator torque command, $G_{FT_{mp}}$ in Equation (7), and from the actuator velocity, $G^{A}_{F\omega_a}$ in Equation (8), were obtained. The Bode diagram of $G_{FT_{mp}}$ shows a phase shift of 150° at frequencies between 10 and 20 rad/s (black solid lines in **Figure 6**). The Bode diagram of $G^{A}_{F\omega_a}$ shows a large influence (amplitude of 54 dB) of the velocity on the force output at a frequency of 10 rad/s (red solid lines in **Figure 6**).

Simulation of the force-feedback lead compensator C_F with different parameters showed that a low f_F brought an overshoot to the closed-loop system $G_{FF_{ref}}$, and a high f_F reduced the bandwidth of $G_{FF_{ref}}$. The parameter a_F influenced $G_{FF_{ref}}$ differently, in that the bigger a_F was, the lower the bandwidth was, and the higher the overshoot was. The closed-loop system $G_{FT_{mp}}$ required a high bandwidth to secure a fast response. However, a high bandwidth requires a high sample rate in controlling the physical system. After many simulations, suitable values for the force-feedback lead controller for the front thigh actuator were determined as: $a_F = 0.05$, $f_F = 100$ rad/s, and $k_F = -1.3$. The bandwidth of $G_{FF_{ref}}$ (black dashed lines in **Figure 6**) was 172.4 rad/s. The same values of a_F and f_F were used for the rest actuators, but k_F for the rear thigh,

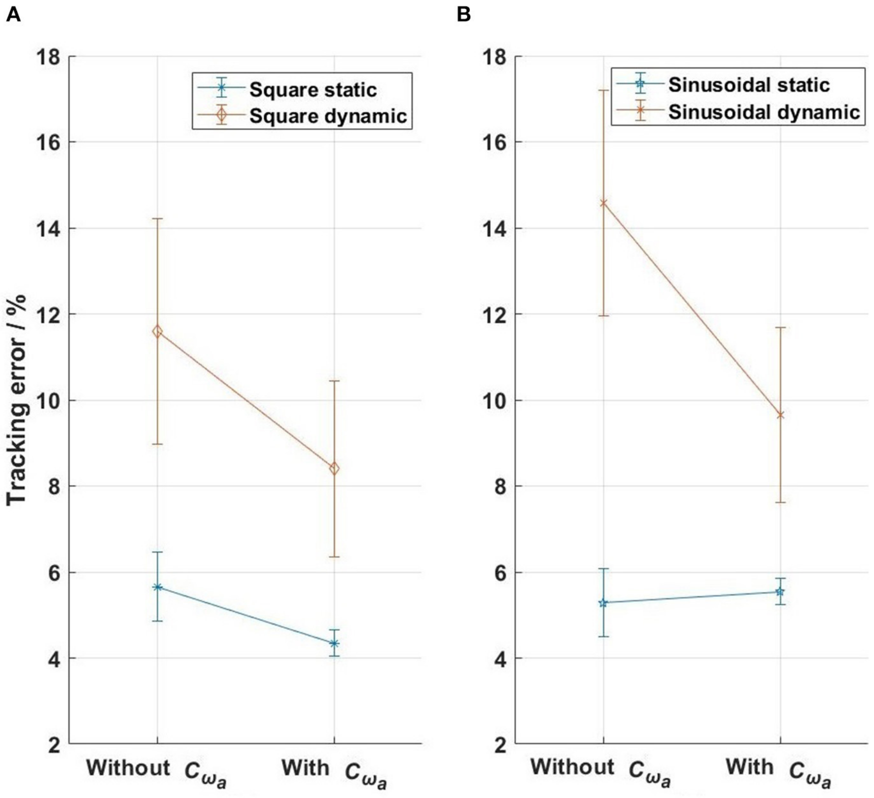


FIGURE 12 | The tracking errors of the four actuators in the evaluation tests. **(A)** Square-wave reference test, **(B)** sinusoidal reference test. The bars with points in the middle show the standard deviations and the means of the errors from the four actuators.

front shank and rear shank actuators were −2.5, −2.5, and −3.5, respectively.

With the lead controller for the force-feedback control, the influence of the velocity on the force output was reduced (red dashed lines in **Figure 6**), but was not yet negligible. The design target of the velocity-feedforward compensator was that its final influence on the force output, $G^{C}_{F\omega_a}$, should compensate as much as possible the influence of the velocity on the force output, $G^{B}_{F\omega_a}$. By trial and error, the parameters of the velocity-feedforward lead compensator for the front thigh actuator were determined as $a_{\omega_a} = 0.02$, $f_{\omega_a} = 10$ rad/s, and $k_{\omega_a} = -1.4$. With this compensator, the magnitude of $G^{C}_{F\omega_a}$ is close to $G^{B}_{F\omega_a}$, especially at frequencies between 1 and 40 rad/s (**Figure 7**). The Fourier transform of the experimental angular velocity during walking at 0.2 m/s showed a maximal frequency of 20 rad/s. Therefore, the current velocity-feedforward lead compensator should be feasible for the front thigh actuator. With the combined force algorithms, the velocity (dashed lines) compared to the reference (solid lines) force had a large influence on the control signal in the overall system (**Figure 8**).

Using similar methods to those described above, the parameters of the velocity-feedforward lead compensators for the rear thigh actuator were determined as: $a_{\omega_a} = 0.02$, $f_{\omega_a} = 10$ rad/s, and $k_{\omega_a} = -1.2$. The values of $a_{\omega_a} = 0.01$, $f_{\omega_a} = 30$ rad/s were used for both shank actuators, but k_{ω_a} for the front and rear shank actuators were, respectively, −2.2 and −1.8.

Frequency analysis of the loop gain G_{Lo} of the front thigh actuator yielded a gain margin of 27.4 dB (at 824 rad/s) and a phase margin of 61.5° (at 103 rad/s). The Nyquist diagram also shows the modulus margin, which was computed to be 0.79 (**Figure 9**). Therefore, the force control system is stable with adequate stability margins.

Evaluation of the Force Control Algorithms

Different force references were implemented and tests performed using the force control algorithms developed in section Frequency-Domain Analysis for the Controller Development. During the static tests (**Figure 10**), the experimental force showed an error of 6 and 4.79% during the tests with the square-wave and sinusoidal references, respectively. It should be noted that the phase delay in the sinusoidal force control, which was 70.6° in the PID-regulated test (**Figure 5G**) was reduced to be 28.0° (**Figure 10G**) using the force-feedback lead controller. Although the participant stood still, the actuator still moved slightly. The mean velocity during the rising edge of the square pulse was 1.5 rad/s in **Figures 10C,F**, and the velocity during sinusoidal test was 0.01 rad/s in **Figures 10I,L**. Using the velocity-feedforward

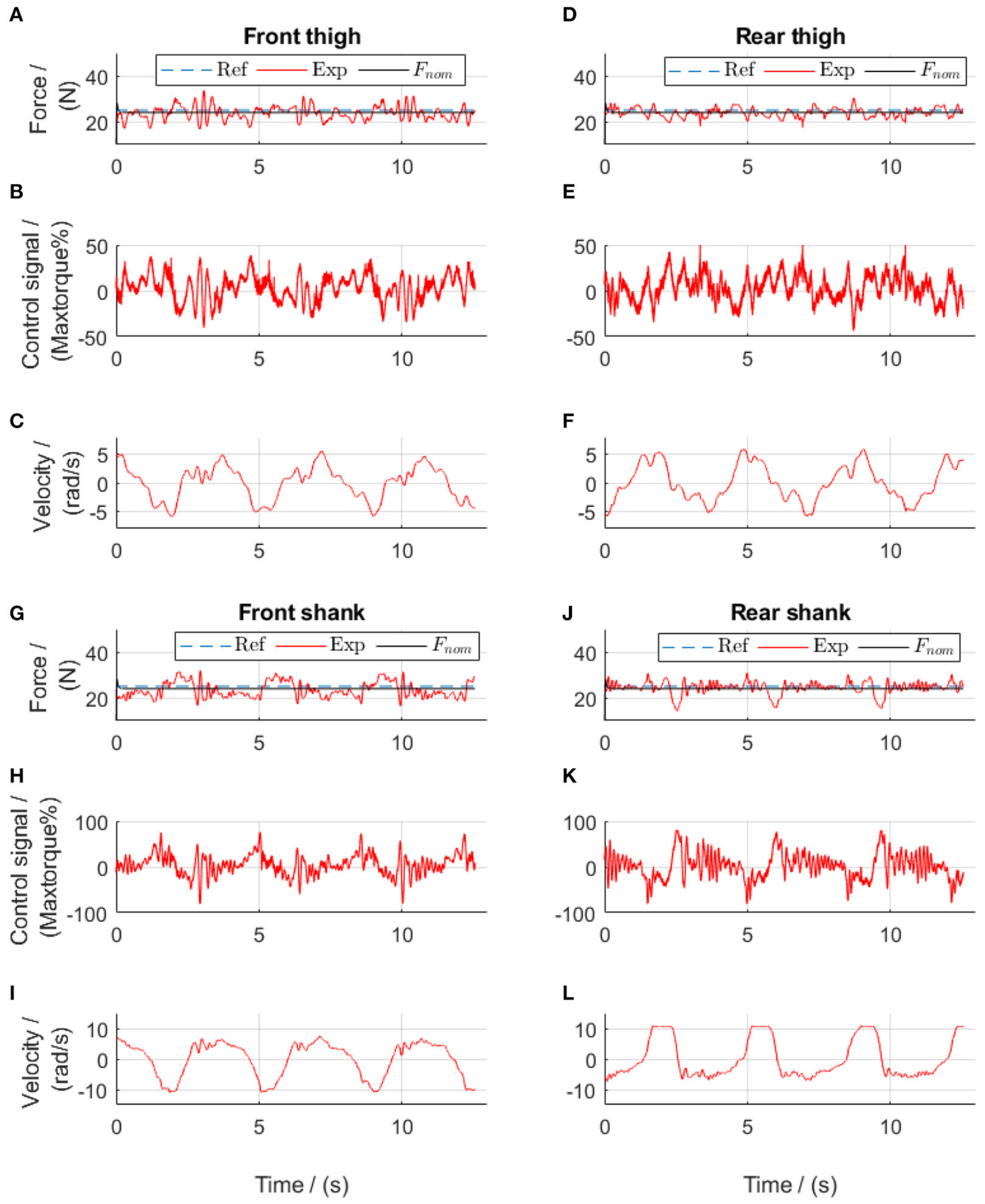


FIGURE 13 | Constant-reference force control during walking using force-feedback lead controllers and velocity-feedforward lead compensators. Subplots **(A–C)** show results from the front thigh actuator, **(D–F)** from the rear thigh actuator, **(G–I)** from the front shank actuator while **(J–L)** from the rear shank actuator. Ref, references; Exp, experimental values; F_{nom}, the simulated target. The positive angular velocity of the front actuators in **(C,L)** means that the right leg moved backwards.

TABLE 1 | System parameters.

Geometry-determined parameters				Test-determined parameters			
	Thigh actuators	**Shank actuators**	**Units**		**Thigh actuators**	**Shank actuators**	**Units**
τ_a	0.0013	0.0012	s	B_a	0.0455	0.03	Nms/rad
k_a	−0.4	−0.2	Nm	k_s	3000	3000	N/m
J_a	0.0156	0.0132	kgm^2	k_{fr}	23	15	
r	0.023	0.023	m				

TABLE 2 | NRMSE [goodness of fit, %, calculated using Equation (10)] between the simulated force F_{sim} and the experimental force F_{exp} during the dynamic tests.

	Front thigh actuator	Rear thigh actuator	Front shank actuator	Rear shank actuator	Mean ± STD
Square	68	66	72	76	71 ± 4.4
Sine	75	68	74	62	70 ± 6.0

lead compensation, the force-tracking error was reduced by 0.6% (comparing **Figures 10A,D**) and by 0.01% (comparing **Figures 10G,J**) during the tests with the square-wave and sinusoidal references, respectively.

During the dynamic evaluation tests, the force-feedback lead controllers alone yielded improved force tracking compared to the tentative PID controllers, with a mean tracking error of 12.5% in square-wave and 15.5% in sinusoidal references (**Figures 11A,G**). However, similar to the force output using the PID controllers, the force output using only the force-feedback lead controllers still showed the influence of the walking movement (comparisons between **Figures 11A,C**, and also between **Figures 11G,I**). After further implementation of the velocity-feedforward lead compensators, the mean force control error reduced to 8.1% in square-wave and 9.2% in sinusoidal references (**Figures 11D,J**) during the dynamic tests. These results demonstrated that the control algorithms were satisfactory. The improved force tracking was obtained by sending a more dynamic control signal to the system. This dynamic control signal thereby brought a more dynamic actuator velocity. Nevertheless, during the dynamic tests, the participant reported much more comfortable walking when the velocity-feedforward lead compensator was further implemented compared to the tests without.

Similar results were obtained in all four actuators (**Figure 12**). The mean force-tracking errors of all four actuator-driven cables in the static tests were 5.7 and 5.3% for the square-wave and sinusoidal references, respectively, which showed that the force-feedback lead controllers produced satisfactory force tracking (mean error of 5.5%) in the static tests. But in the dynamic tests, the mean tracking errors were 11.6 and 14.8% for the square-wave and sinusoidal references (mean error of 13.2%), respectively. Further implementation of the velocity-feedforward compensators reduced the force control error in the dynamic tests to 8.4 and 9.6% for the square-wave and sinusoidal references, respectively. This showed that the force-feedback controllers with velocity-feedforward compensators produced satisfactory force tracking (mean error of 9.0%) in the dynamic tests.

In the final walking test with a constant reference in each cable (see **Supplementary Material** "Walk with the active cable-driven robotic system. mp4"), both thigh actuators yielded good force tracking (**Figure 13**). The shank actuators had velocities almost twice those of the thigh actuators, thereby having a larger influence of velocity on the force output. The slightly noisy control signals came mainly from the velocity input (**Figure 8**). Nevertheless, all actuators yielded satisfactory force tracking, with tracking errors from the actuators for the front thigh, rear thigh, front shank and rear shank of 12.1, 8.2, 12.4, and 8.5%, respectively, resulting in a mean error of 10.3%. It should be noted that the reference force was only 25 N, therefore the absolute mean force tracking error was 2.6 N.

DISCUSSION

The aim of this work was to develop an active cable-driven robotic system, and to evaluate force control strategies for walking rehabilitation using frequency-domain analysis. The experiment-assisted simulation method determined the system parameters, which provided the basis for the development of the force control strategies. The control algorithms were firstly designed using frequency-domain analysis, then simulated in Simulink dynamic models, and finally evaluated in the active cable-driven robotic system. The force-feedback lead controllers provided good force tracking in the static tests, but were not satisfactory in the dynamic tests. Using velocity-feedforward lead compensators in addition to the force-feedback lead controllers, the active cable-driven robotic system produced satisfactory performance in tracking various force profiles during walking. The novelties of the current work include using the experimental-assisted simulation method to determine the system parameters, and deducing, simulating, and evaluating lead control algorithms for the force regulation and motion-related disturbance compensation in the cable-driven system. To the authors' knowledge (based on the extensive literature review), this control strategy hasn't been published in the literature. After being validated in four cable-driven actuators in the current work, this control strategy can be applied generally in active force-controlled cable-driven robotic systems. The robotic system developed in this work has the potential to produce user-defined assistance or resistance in rehabilitation and fitness training.

The experiment-assisted simulation method determined the system parameters, which helped to understand the system dynamics, thereby developing the proper control algorithms. The dynamics of the cable-driven robotic system is complex, with the transfer function from the actuator control signal to the force output (Equation 7) as a fourth-order system. Determination of the system parameters was the prerequisite for understanding the system dynamics. Study (Jung and Bae, 2016) used system identification and linear quadratic methods to determine the plant transfer functions and the controller. Other studies (Zou et al., 2018a, 2019; Wang et al., 2020) listed the system parameters without describing how they were obtained. Using the tentative PID controllers, the current work performed several tests on the active cable-driven robotic system, where the actual force outputs were recorded. Then repetitive model simulations found suitable system parameters, which yielded similar force outputs to the experiments (the mean NRMSE $\geq$ 70%). This experiment-assisted simulation method searched for suitable parameters for the active cable-driven robotic system during walking, which provided the basis for the following development of the control strategies. However, there might be other suitable system parameters, and a systematic determination of the optimal parameters would require further investigation.

Based on the system dynamics, force-feedback lead controllers were developed, which yielded satisfactory static force tracking. Traditional PD controllers are often reported in the literature (Marchal-Crespo et al., 2012; Alamdari and Krovi, 2015b). The current study implemented PID controllers during the system parameter determination phase. But the phase delay observed in the sinusoidal-reference force tests motivated the development of force-feedback lead controllers. Comparing the lead controllers with the PID controllers in the current study revealed that the PID controllers produced better tracking of the square wave force, due to the inclusion of the integral term (comparing **Figure 5A** with **Figure 10A**), while the lead controllers reduced the phase delay during sinusoidal force tracking (comparing **Figure 10G** with **Figure 5G**). It is believed that by tuning the parameters of both types of force-feedback controllers, the static force tracking could be improved. The force-feedback lead controllers showed a steady-state error of 2.2%. An integrator could be further included, as demonstrated in studies (Zou et al., 2018a, 2019), to remove this error.

Frequency-domain analysis guided the development of the velocity-feedforward lead compensators, which improved the force-tracking accuracy during walking. Dynamic force control of cable-driven systems is challenging because the velocity-related disturbance cannot be removed, but can only be partially compensated (Zou et al., 2018a). Based on the frequency-domain analysis of the angular velocity of the actuator and the force output (Equations 15 and 20), the parameters of the lead compensators were straightforwardly determined by simulation. Dynamic tests on the active cable-driven robotic system showed that the disturbance was similarly suppressed as demonstrated in the simulation. The velocity-feedforward compensators in the current study were simpler than the compensator developed using the structure invariance principle (Zou et al., 2018a,b). The current velocity-feedforward compensators could be further improved by using double lead elements. Nevertheless, the mean error in all actuators was 10.3% in the force control with a constant-reference within the four cables, which can be considered negligible. The control algorithms have the potential to produce user-defined assistance or resistance in rehabilitation and fitness training (Marchal-Crespo et al., 2012; Von Zitzewitz et al., 2013) by modifying the force references of each actuator.

The limitations of this study included the mechanical and control development of the active cable-driven robotic system. The cables did not always follow the threads on the drum during the dynamic walking tests. A deflection unit as presented in Marchal-Crespo et al. (2012) is required to secure regular winding of each cable on each drum. In order to achieve accurate measurement of the pulling force on the leg, the force sensors were fixed close to the leg. However, this resulted in a situation where the force sensors moved with the leg during the dynamic tests. This movement, especially the shock at heel strike, was believed to influence the force measurements. Due to mechanical constraints, the current cable-driven robotic system could produce a maximal speed of 0.6 m/s. Therefore, the force control at high speed conditions was not investigated. Four actuator-drum assemblies with a larger drum size and proper fixation of the force sensors will be developed for accurate control of high-speed walking training. In order to compensate potentially large motion-dependent disturbances at high speed, double lead compensators could be implemented, especially for the shank actuators. The force control algorithms, although tested to be feasible in the current study, were not optimal. The force-feedback lead controllers could include an integrator to remove the steady-state errors. New control algorithms such as PID controllers with adaptive parameters, and disturbance observers as used in Jung and Bae (2016) and Wang et al. (2019) will be investigated. Many different control strategies exist in the literature for control of cable-driven systems. The control strategies presented in Von Zitzewitz et al. (2013) are also candidates for further investigation in our system in future.

The current paper evaluated lead control algorithms for force regulation and velocity compensation, with different target force profiles (sine and square wave profiles) on each cable tested. From the view of force controller evaluation, the tests were successful with straightforward results. The non-parallel cables during walking did not cause problems for the force regulation. But more work is required from the view of promotion of walking rehabilitation. Kinematic and kinetic analysis will be performed to define a suitable target force profile for each cable that jointly produce the required kinetics on the leg during normal gait (Fang et al., 2014), as a strategy to effectively guide patients in walking relearning. As another strategy to promote rehabilitation, impedance control will be implemented so as to guide patients to generate appropriate gait and to exert the appropriate force during walking training. The current work provides the important basis for the further application of the cable-driven robotic systems in walking rehabilitation.

CONCLUSIONS

An active cable-driven, force-controlled robotic system for walking rehabilitation was developed. Using frequency-domain analysis, force-feedback controllers, and velocity-feedforward compensators were developed. The technical evaluation showed that the combined control algorithms produced satisfactory force tracking during walking in the active cable-driven robotic system. This study demonstrated that the force control algorithms were technically feasible. In future, four actuator-drum assemblies with proper fixation of the force sensors will be developed, and the force control algorithms will be further improved by adding integrator elements, implementing double-lead velocity compensators, and tuning the control parameters. As strategies to effectively guide patients in walking relearning, kinematic and kinetic analysis will be performed, and impedance control will be further implemented. The active cable-driven, force-controlled robotic system has the potential to produce user-defined assistance or resistance in rehabilitation and fitness training.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

JF, LM-C, and KH developed the system concept. MH assembled the active cable-driven robotic system. JF developed the control algorithms. JF and MH implemented the control programs. JF conducted the experiments, analyzed the data, and drafted the manuscript. KH, LM-C, and MH revised it critically for important intellectual content. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbot.2021.651177/full#supplementary-material

REFERENCES

Alamdari, A., and Krovi, V. (2015a). "Robotic Physical Exercise and System (ROPES): a cable-driven robotic rehabilitation system for lower-extremity motor therapy," in *ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Boston, MA), 1–10. doi: 10.1115/DETC2015-46393

Alamdari, A., and Krovi, V. (2015b). Design and analysis of a cable-driven articulated rehabilitation system for gait training. *J. Mech. Robot.* 8:051018. doi: 10.1115/1.4032274

Aprile, I., Cruciani, A., Germanotta, M., Gower, V., Pecchioli, C., Cattaneo, D., et al. (2019). Upper limb robotics in rehabilitation: an approach to select the devices, based on rehabilitation aims, and their evaluation in a feasibility study. *Appl. Sci.* 9:3920. doi: 10.3390/app9183920

Aström, K. J., and Murray, R. M. (2010). *Feedback Systems: An Introduction for Scientists and Engineers*: Princeton, NJ: Princeton University Press. doi: 10.2307/j.ctvcm4gdk

Beckmann-Hemmers, D. (2012). *Einsatz des h/p/Cosmos Robowalk Expander Systems® bei Patienten mit Hemiparese - Eine Kritische Analyse*. Bachelor of Arts, Medizinalfachberufe, Europäische Hochschule Nordhessen, Bochum.

Fang, J., Galen, S., Vuckovic, A., Conway, B. A., and Hunt, K. J. (2014). Kinetic analysis of supine stepping for early rehabilitation of walking. *Proc. Inst. Mech. Eng. H* 228, 456–464. doi: 10.1177/0954411913518031

Hesse, S., Schmidt, H., Werner, C., and Bardeleben, A. (2003). Upper and lower extremity robotic devices for rehabilitation and for studying motor control. *Curr. Opin. Neurol.* 16, 705–710. doi: 10.1097/00019052-200312000-00010

Hidler, J., Nichols, D., Pelliccio, M., and Brady, K. (2005). Advances in the understanding and treatment of stroke impairment using robotic devices. *Top. Stroke Rehabil.* 12, 22–35. doi: 10.1310/RYT5-62N4-CTVX-8JTE

Jin, X., Cui, X., and Agrawal, S. K. (2015). "Design of a cable-driven active leg exoskeleton (C-ALEX) and gait training experiments with human subjects," in *IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA), 5578–5583. doi: 10.1109/ICRA.2015.7139979

Jöllenbeck, T., and Pietschmann, J. (2018). "Effektivität eines taktilen Gangtrainings zur Normalisierung des Gangbildes während der orthopädischen Rehabilitation nach Hüft-TEP," in *Rehabilitation Bewegt!: 27. Rehabilitationswissenschaftliches Kolloquium: Deutscher Kongress Für Rehabilitationsforschung*, eds R. Buschmann-Steinhage, H.-G. Haaf, and U. Koch (München).

Jung, Y., and Bae, J. (2016). "An asymmetric cable-driven mechanism for force control of exoskeleton systems," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Daejeon), 273–278. doi: 10.1109/IROS.2016.7759066

Maier, M., Ballester, B. R., and Verschure, P. F. M. J. (2019). Principles of neurorehabilitation after stroke based on motor learning and brain plasticity mechanisms. *Front. Syst. Neurosci.* 13:74. doi: 10.3389/fnsys.2019.00074

Marchal-Crespo, L., Rauter, G., Wyss, D., Zitzewitz, J. V., and Riener, R. (2012). Synthesis and control of an assistive robotic tennis trainer. in *4th IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)* (Rome), 355–360. doi: 10.1109/BioRob.2012.6290262

Rosati, G., Masiero, S., and Rossi, A. (2017). On the use of cable-driven robots in early inpatient stroke rehabilitation. in *Advances in Italian Mechanism Science*, Vol. 47, eds G. Boschetti and A. Gasparetto (Cham: Springer), 551–558. doi: 10.1007/978-3-319-48375-7_59

Schulze, S., Lauenroth, A., Delank, K., and Schwesig, R. (2019). Einfluss von Expanderzügen auf dem laufband auf spatio-temporale und dynamische gangparameter bei gesunden probanden. *Physioscience* 15, 124–130. doi: 10.1055/a-0965-5506

Von Zitzewitz, J., Morger, A., Rauter, G., Marchal-Crespo, L., Crivelli, F., Wyss, D., et al. (2013). A reconfigurable, tendon-based haptic interface for research into human-environment interactions. *Robotica* 31, 441–453. doi: 10.1017/S026357471200046X

Wang, Y., Wang, K., Zhang, Z., and Mo, Z. (2020). Control strategy and experimental research of a cable-driven lower limb rehabilitation robot. *Proc. Inst. Mech. Eng. C J. Mec. Eng. Sci.* 1–14. doi: 10.1177/0954406220952510

Wang, Z., Zi, B., Wang, D., Qian, J., You, W., and Yu, L. (2019). External force self-sensing based on cable-tension disturbance observer for surgical robot end-effector. *IEEE Sens. J.* 19, 5274–5284. doi: 10.1109/JSEN.2019.2903776

Wu, M., Hornby, T. G., Landry, J. M., Roth, H., and Schmit, B. D. (2011). A cable-driven locomotor training system for restoration of gait in human SCI. *Gait Posture* 33, 256–260. doi: 10.1016/j.gaitpost.2010.11.016

Zhang, L., Li, L., Zou, Y., Wang, K., Jiang, X., and Ju, H. (2017). Force control strategy and bench press experimental research of a cable driven astronaut rehabilitative training robot. *IEEE Access.* 5, 9981–9989. doi: 10.1109/ACCESS.2017.2702188

Zou, Y., Ma, H., Han, Z., Song, Y., and Liu, K. (2018a). Force control of wire driving lower limb rehabilitation robot. *Technol. Health Care* 26, S399–S408. doi: 10.3233/THC-174714

Zou, Y., Wang, N., Wang, X., Ma, H., and Liu, K. (2019). Design and experimental research of movable cable-driven lower limb rehabilitation robot. *IEEE Access.* 7, 2315–2326. doi: 10.1109/ACCESS.2018.2887233

Zou, Y., Zhang, L., Li, L., Ma, H., and Liu, K. (2018b). Running experimental research of a wire driven astronaut rehabilitative training robot. *IEEE Access.* 6, 11464–11471. doi: 10.1109/ACCESS.2018.2809735

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Neurorobotics

BRIEF RESEARCH REPORT
published: 07 June 2021
doi: 10.3389/fnbot.2021.688344

Spiking Neural Network for Fourier Transform and Object Detection for Automotive Radar

*Javier López-Randulfe**†, *Tobias Duswald*†, *Zhenshan Bing and Alois Knoll*

Department of Informatics, Technical University of Munich, Munich, Germany

The development of advanced autonomous driving applications is hindered by the complex temporal structure of sensory data, as well as by the limited computational and energy resources of their on-board systems. Currently, neuromorphic engineering is a rapidly growing field that aims to design information processing systems similar to the human brain by leveraging novel algorithms based on spiking neural networks (SNNs). These systems are well-suited to recognize temporal patterns in data while maintaining a low energy consumption and offering highly parallel architectures for fast computation. However, the lack of effective algorithms for SNNs impedes their wide usage in mobile robot applications. This paper addresses the problem of radar signal processing by introducing a novel SNN that substitutes the discrete Fourier transform and constant false-alarm rate algorithm for raw radar data, where the weights and architecture of the SNN are derived from the original algorithms. We demonstrate that our proposed SNN can achieve competitive results compared to that of the original algorithms in simulated driving scenarios while retaining its spike-based nature.

Keywords: spiking neural network, FMCW radar, Fourier tranform, constant false-alarm rate, autonomous driving

OPEN ACCESS

Edited by:
Shuai Li,
Swansea University, United Kingdom

Reviewed by:
Dechao Chen,
Hangzhou Dianzi University, China
Ameer Tamoor Khan,
Hong Kong Polytechnic University, China

****Correspondence:***
Javier López-Randulfe
lopez.randulfe@tum.de

† These authors have contributed equally to this work

Received: *30 March 2021*
Accepted: *06 May 2021*
Published: *07 June 2021*

Citation:
López-Randulfe J, Duswald T, Bing Z and Knoll A (2021) Spiking Neural Network for Fourier Transform and Object Detection for Automotive Radar. Front. Neurorobot. 15:688344. doi: 10.3389/fnbot.2021.688344

1. INTRODUCTION

Autonomous driving is a billion-dollar business with high demand for efficient computing systems. This introduces a limitation for highly automated vehicles, where the systems that process sensor data can drain more than 10% of the power stored for driving (Lin et al., 2018). Radar sensors are a fundamental component of most autonomous vehicles. Their low price and robustness against bad weather and lighting conditions make them great companions for lidar and vision cameras (Hasch et al., 2012; Winner et al., 2014; Patole et al., 2017). Whereas, the algorithms in this field are typically implemented on graphical processing units (GPUs), field-programmable gate arrays (FPGAs), or application-specific integrated circuits (ASICs) (Lin et al., 2018), neuromorphic hardware (NHW) offers an efficient alternative environment (Furber et al., 2014; Davies et al., 2018; Sangwan and Hersam, 2020). It provides a low-energy-footprint platform for a new generation of neural networks called spiking neural networks (SNNs), which reduce the high energy consumption of the popular artificial neural networks (ANNs) (Maass and Schmitt, 1999; Bouvier et al., 2019; Strubell et al., 2019).

SNN-based applications for sensor signal processing comprise a novel area within autonomous driving and have been applied to traditional vision cameras for tracking applications (Cao et al., 2015; Piekniewski et al., 2016); raw temporal pulses of lidar sensors for object detection (Wang et al., 2021); dynamic vision sensors for tasks including lane keeping (Bing et al., 2020), feature extraction and motion perception (Paredes-Vallés et al., 2019), and collision avoidance (Salvatore et al., 2020); as well as remote sensing images for object detection (Liu et al., 2018). Likewise, bio-inspired neural networks have also been applied to control problems in mobile applications, including rotation control for unmanned aerial vehicles (Stagsted et al., 2020); and tracking control (Khan et al., 2021) and obstacle avoidance (Khan et al., 2020) for smart-home manipulators.

Current trends in signal processing include optimization of the energy consumption, memory resources, and computational speed of the Fourier transform (FT) (Gilbert et al., 2014). In the case of radar processing, the FT converts the raw data from the sensor into a range-Doppler map, followed by an object-detection algorithm that separates the target values from noise (Patole et al., 2017). Later stages extract high-level information, such as a list of labeled objects or target trajectories. A few recent works have explored the application of SNNs to decomposing a time signal into a frequency spectrum, e.g., by applying sequential spiking band-pass filters to audio signals (Jiménez-Fernández et al., 2016) or using neurons that spike at specific input frequencies (Auge and Mueller, 2020). The former offers an efficient bio-inspired solution, but its applications are limited to extracting a small set of frequency components. The latter provides a simple and elegant solution, but fails to provide an accurate measurement of the frequency-components angles, which are crucial for computing a 2D FT.

In this letter, we present a novel SNN that is able to effectively replace the discrete FT (DFT) and constant false-alarm rate (CFAR) algorithms (Rohling, 1983). Its weights are fixed, as they are mathematically derived from the equations defining the two algorithms, leading to equivalent results in simulated driving scenarios. Therefore, the network does not require learning to adapt its weights. Modern radar applications use the fast FT (FFT), which consists of a recursive decomposition of the problem into smaller DFTs over subsets of the original data (Frigo and Johnson, 2005), i.e., the DFT is a generalization of the FFT. The proposed network is a theoretical approximation of the DFT, and further work can explore network topologies with smaller layers that mimic the desired FFT structure. The quantitative results show high similarity between the outputs of the proposed SNN and the original algorithms for both one and two dimensions. In combination with NHW, this work could provide an efficient alternative for processing radar data while maintaining analogous performance. Designing SNNs that can process sensor data is a crucial step for obtaining full neuromorphic sensor processing pipelines and, together with NHW, bring a new generation of signal-processing solutions with higher energy efficiencies and shorter latency response-times. Furthermore, finding SNN equivalents for all radar-processing stages is paramount, as hybrid pipelines introduce additional complexity through communication and spike conversion when data flows between the NHW and traditional hardware.

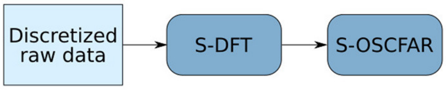


FIGURE 1 | Diagram of the implemented pipeline. The first block implements a spiking DFT on raw discretized FMCW radar data, and the second block implements the spiking OS-CFAR.

2. SPIKING NEURAL NETWORK

In this section, we explain an SNN that substitutes the 2D DFT (S-DFT) and object detection (S-OSCFAR) on raw frequency-modulated continuous-wave (FMCW) radar data. To do so, the network is split into two smaller networks that process each of the steps (see **Figure 1**).

2.1. Fourier Transform

The 2D DFT is a linear transformation that can be represented by two successive matrix multiplications. Hence, each of the dimensions of the 2D DFT can be represented by a one-layer neural network with a linear activation function.

We have implemented the S-DFT with a two-layer network, where the first layer provides information about the target ranges and the second extracts their velocities. The input data has dimension $N \times M$, where N is the total number of samples per chirp, and M is the number of chirps in a radar frame. Owing to the complex nature of the DFT, both the real and imaginary values must be computed. Therefore, each layer contains $2N \times M$ neurons (**Figure 2**).

The weights of the network have been calculated based on the trigonometric equation of the DFT:

$$Y_k = \sum_{l=0}^{L-1} X_l \left[cos\left(\frac{2\pi}{L}kl\right) - i \cdot sin\left(\frac{2\pi}{L}kl\right)\right] . \qquad (1)$$

Both the input vector X and output vector Y are formed by L values; hence, k and l take values between 0 and $L-1$. L takes the value of N for the first layer and M for the second. We rewrite Equation (1) as

$$Y = W_{Re}X + i \cdot W_{Im}X , \qquad (2)$$

where Y is the result of the transform, X is the input vector, and W_{Re} and W_{Im} are the real and imaginary coefficients, respectively. From Equation (1), the individual weight coefficients can be calculated using the following equations:

$$w_{Re_{kl}} = \quad cos\left(\frac{2\pi}{L}kl\right) , \qquad (3)$$

$$w_{Im_{kl}} = -sin\left(\frac{2\pi}{L}kl\right) . \qquad (4)$$

The first layer is replicated in parallel M times using the same weights. The output consists of the frequency spectrum of the input chirps, formed each by $2N$ values.

The second layer applies the same algorithm to the output of the first layer, $Y^{(1)}$. The main difference is that the input in this case is formed by complex values, so the real and imaginary

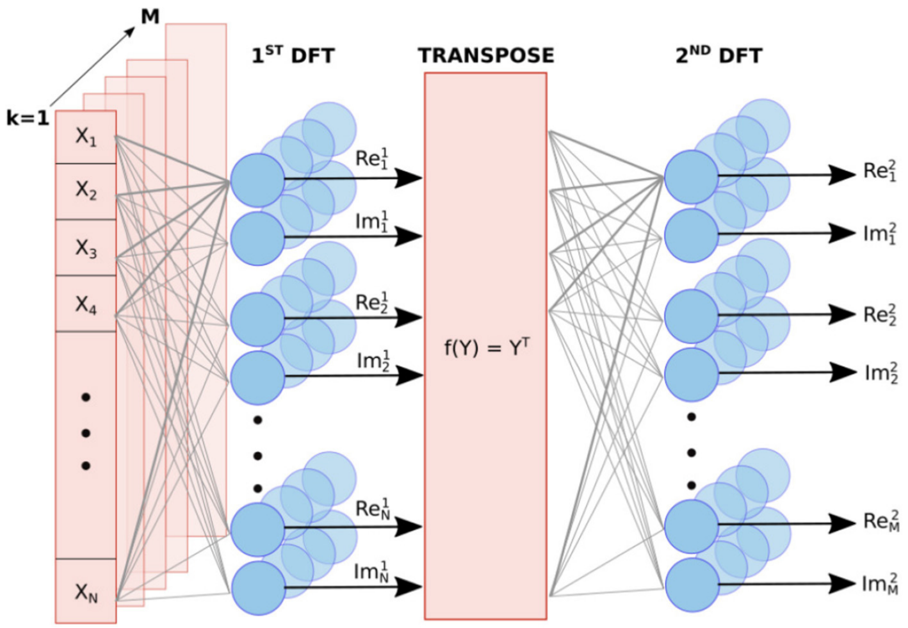


FIGURE 2 | Diagram of the S-DFT. The first layer implements one DFT per chirp to obtain the range-dependent frequency bins, and the second layer performs a DFT along the different chirps to obtain the Doppler frequency components. A transpose operation is implemented between these steps, as each layer is applied along a different dimension.

parts obtained from Equation (2) for the first layer are fed into the equation for the second layer. Therefore, Equation (2) can be generalized for the second layer by using the following algebraic system:

$$\begin{bmatrix} Re\left(Y^{(N)}\right) \\ Im\left(Y^{(N)}\right) \end{bmatrix} = \begin{bmatrix} W_{Re} & W_{Im} \\ -W_{Im} & W_{Re} \end{bmatrix} \begin{bmatrix} Re\left(Y^{(N-1)}\right)^T \\ Im\left(Y^{(N-1)}\right)^T \end{bmatrix}. \tag{5}$$

For this layer, W_{Re} and W_{Im} are $M \times M$ matrices.

With appropriate weights, a two-layer feed-forward network with linear activation can represent Equations (2) and (5). A rate-based SNN can approximate a neural network based on rectified linear unit (ReLU) functions (Rueckauer et al., 2017). For one layer, such a network transforms its input $a \in \mathbb{R}^n$ as

$$z_j = \text{ReLU}\left(\sum_i W_{ji} a_i\right), \tag{6}$$

where $W \in \mathbb{R}^{m \times n}$ is the weight matrix and $z \in \mathbb{R}^m$ is the result. In the rate-based approximation, the input is represented by n neurons whose spike frequencies are proportional to the components of a. Assuming that we simulate the SNN for k timesteps, we denote the spike train of the input neuron i as a binary vector $\Theta^{(i)} \in \{0, 1\}^k$, where 1 in indicates a spike and $\Theta^{(i)}$ is computed from a_i. Rueckauer et al. (2017) showed that z can be approximated by the spike frequency of m output neurons. The membrane potential V_j fully describes the state of the output neuron j. At a certain timestep $1 < t < k$, it is computed according to

$$V_j^t = V_j^{t-1} + V_{thr} \sum_i W_{ji} \Theta_t^{(i)}, \tag{7}$$

where V_{thr} is a model parameter. We compute the output spike train with a threshold behavior. Whenever $V_j^t > V_{thr}$, the output neuron j spikes and V_j^t is reduced by V_{thr}. The spike train defines the spike frequency, and we obtain an approximation to y_j. For deeper networks, the neuron model governs all hidden layers as well as the output layer (Rueckauer et al., 2017).

Latest conversion approaches are based on ReLU networks. However, a single ReLU function does not span over the entire real spectrum, so the S-DFT implements the activation by combining two ReLU functions with opposite signs in their weights. To still use the existing conversion theory, we rewrite a regular matrix multiplication as a sum of two ReLU layers

$$\sum_i W_{ji} a_i = \text{ReLU}\left(\sum_i W_{ji} a_i\right) - \text{ReLU}\left(\sum_i \left(-W_{ji}\right) a_i\right), \tag{8}$$

which leads to two convertible sub-networks with weights of opposite signs. This logic is applied to Equations (1) and (5) in order to obtain a convertible network and ultimately an SNN for the DFT.

2.2. OS-CFAR

CFAR algorithms differentiate signals from noise. In the context of radar processing, the ordered-statistics (OS) CFAR is a very

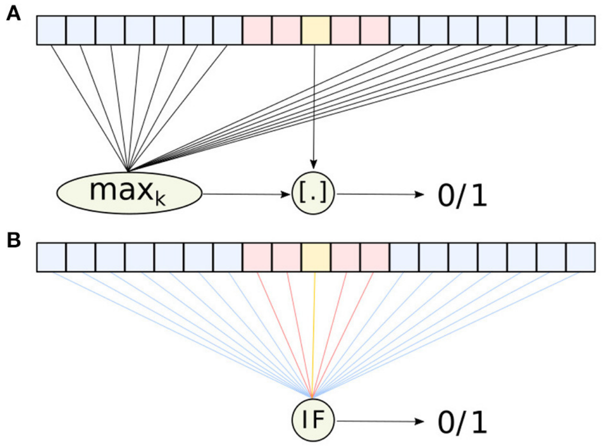


FIGURE 3 | Diagrams representing the functionality of **(A)** the traditional OS-CFAR algorithm, and **(B)** the spiking OS-CFAR algorithm described in section 2.2. The diagrams show the N neighbor cells, G guarding cells, and the value under consideration x_c in blue, red, and yellow, respectively. The different connection colors indicate the weights: −1 (blue), 0 (red), and $1 + (k - 1)$ (yellow).

prominent candidate to detect targets in range-Doppler maps (Rohling, 1983).

It operates on a 1D array x containing $N + G + 1$ values, where N is the number of neighbor cells, G is the number of guarding cells, and the remaining value is the one to be classified (see **Figure 3**). Let us denote the latter as x_c and the set of all neighbor cells as $\mathcal{N}$. The OS-CFAR algorithm compares the *kth* largest value of $\mathcal{N}$ against αx_c, where $0 < \alpha < 1$ is a scale factor. If αx_c is larger, it is classified as a signal; otherwise, it is considered noise. This can be expressed as

$$\text{OS-CFAR}(x) = [\alpha x_c > max_k\{y|y \in \mathcal{N}\}]\,, \tag{9}$$

where max_k selects the *kth* largest element and [.] denotes the Iverson bracket[1].

The OS-CFAR algorithm can be replaced by an SNN consisting of a single integrate-and-fire (IF) neuron, see Equation (7). We refer the reader to chapter 1.3 of Gerstner et al. (2014) for more details about the neuron model. If the IF neuron receives a spike from a pre-synaptic neuron connected with weight w, the membrane potential v instantaneously increases to the new value, $v + w$. The information of the real-valued vector $x \in \mathbb{R}^q$ is encoded in the precise spike time of a set of q pre-synaptic neurons. For a given value x_i, the associated pre-synaptic neuron i spikes at

$$t(x_i) = -\left(t_{max} - t_{min}\right)\frac{x_i - x_{min}}{x_{max} - x_{min}} + t_{max}\,, \tag{10}$$

where $x_{min} < x_i < x_{max}\ \forall i$, and t_{min} and t_{max} define the earliest and latest spike time, respectively. Equation (10) is a linear transformation that maps larger values to earlier spike times.

The defined encoding allows us to rewrite the CFAR problem. Instead of finding the *kth* largest value and comparing it against αx_c, we evaluate whether fewer than k spikes arrive before $t(\alpha x_c)$. Therefore, we need to choose the parameters of the IF neuron appropriately. Among many different choices, we propose defining the initial membrane potential, firing threshold, neighbor weights, guarding weights, and weight for the value under consideration as $v_0 = 0$, $v_{th} = 1$, $w_{\mathcal{N}} = -1$, $w_{\mathcal{G}} = 0$, and $w_c = 1 + (k - 1)$, respectively. With these parameters, the neuron spikes if and only if fewer than k spikes arrived before $t(\alpha x_c)$. The neuron spikes in the same manner in which the OS-CFAR computes 0 or 1; thus, in that sense, both methods are mathematically equivalent. To guarantee that the classic algorithm and its spiking counterpart detect the same peaks, the time step Δt must comply with $\Delta t < |t_{k-1} - t(\alpha x_c)|$; i.e., the $(k-1)th$ spike must arrive at least one time step before the reference spike.

The method can be generalized to the 2D case by defining the neighbor and guarding cells on a 2D array, as done for the traditional OS-CFAR algorithm. The interested reader may find more information in Kronauge and Rohling (2013). Although the weights and spike times remain unchanged, the algorithmic parameters k and α need to be adapted to the new problem to obtain meaningful results (Rohling, 1983).

3. EXPERIMENT RESULTS

The data used in the experiments has been gathered using an automotive radar simulator. This tool simulates a 77 GHz radar with a user-specified number of targets in the sensed scene. The generated data also includes several sources of noise typically present in real radar data, e.g., analog front-end noise, ADC noise and saturation, phase noise, and thermal noise. The generated data frames are formed by 128 chirps and 1,024 samples per chirp. The bandwidth and duration of each chirp are 275 MHz, and 54 μs, respectively.

The simulated sensor is a long-distance radar (range up to 300 m), and the scene is populated with three targets at ranges $R = \{5, 9, 100\}\ m$. The velocities of the targets relative to the radar are $V = \{0, 2, 14\}\ m/s$, and the radar cross section (RCS) $\sigma = \{0, 5, 40\}\ dBsm$. The first two objects represent two pedestrians, whereas the last target represents a vehicle. These values are based on the typical velocities of such targets, and the RCS values are obtained from previous studies (Kamel et al., 2017; Deep et al., 2020). The simulation also includes 20-mm sensor packaging and the 1-cm car bumper, with RCS $\sigma_P = -40\ dBsm$, and $\sigma_B = -23\ dBsm$, respectively.

The proposed algorithm was implemented in Python, and the code is open-source[2].

3.1. One-Dimensional DFT and CFAR

In the first experiment, we applied the S-DFT and S-OSCFAR to a single chirp of the simulated scene. The 1D S-DFT was simulated with 1,000 time steps of size $\Delta t = 0.01$ ms, i.e., for a total time of 10 ms. The input values are converted to spike trains with

---

[1] The Iverson bracket $[C]$ yields 1 if the condition C is true and 0 if false.

[2] https://github.com/KI-ASIC-TUM/spiking-dft-cfar

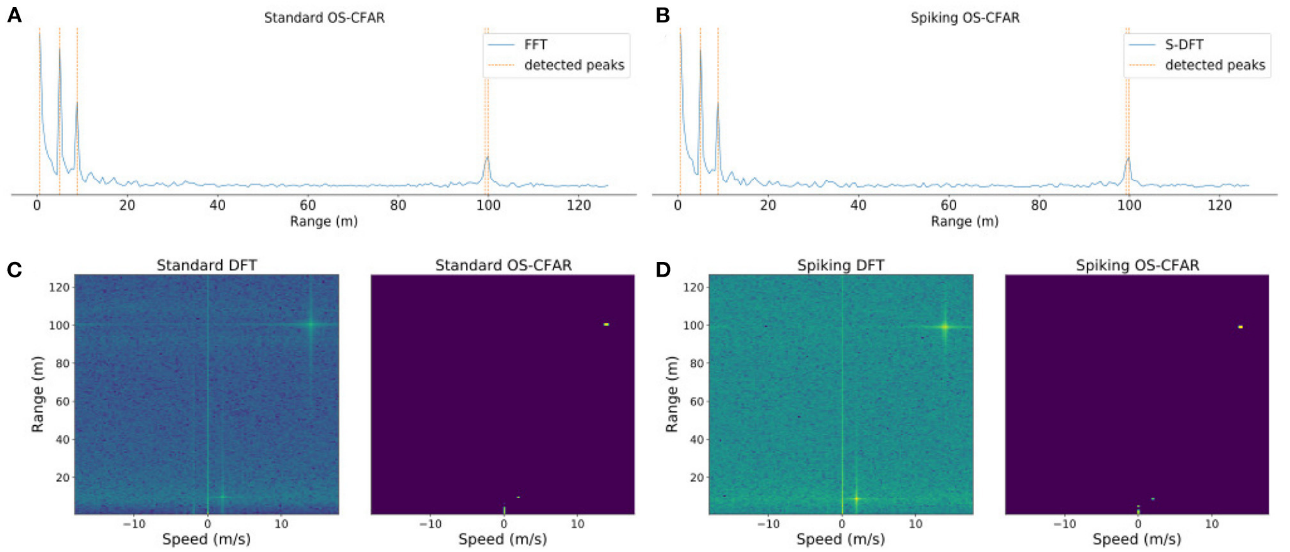


FIGURE 4 | Results of the experiments on a simulation with three targets and a car bumper next to the sensor for **(A)** the 1D standard DFT and OS-CFAR algorithms, **(B)** the 1D S-DFT and S-OSCFAR, **(C)** the 2D standard DFT and OS-CFAR algorithms, and **(D)** the 2D S-DFT and S-OSCFAR. The sensor range is up to 300 m, but for visualization purposes the range dimension has been cropped until 125 m.

regular spike time intervals $\Delta T \in [0.2, 10.0]$ ms. To encode the S-OSCFAR values, we set the parameters t_{min} and t_{max} to 0 and 50 ms, respectively. This sub-network was also simulated with a temporal resolution $\Delta t = 0.01$ ms. Both the classical and spiking CFAR algorithms used $G = 12$ guarding cells, $N = 30$ neighbor cells, a scale factor $\alpha = 0.2$, and compared against the $kth = 6th$ largest value.

Figure 4A shows the results of the original algorithms, and **Figure 4B** shows the results of the proposed algorithm for a single dimension, i.e., range. The performance of both the standard and spiking versions are comparable, and they were able to detect the three objects present in the scene.

For evaluating the performance of the S-DFT, we measured the root mean square error (RMSE) between the DFT and S-DFT, which yielded RMSE $=$ 0.0056. This measure was taken after normalizing the output values between 0 and 1. This small error is generated during the encoding process, as the S-DFT is a mathematical approximation of the original DFT.

3.2. Two-Dimensional DFT and CFAR

In the second experiment, we applied the S-DFT and S-OSCFAR to the complete radar frame of the simulated scene. The 2D S-DFT simulation parameters were identical to the 1D case, except for the total simulation time, which was increased to 50 ms. Furthermore, the encoding parameters for the S-OSCFAR algorithm and the temporal resolution were identical to the 1D experiment. The CFAR parameters were set to $G = 48$ guarding cells, $N = 176$ neighbor cells, a scale factor $\alpha = 0.2$, and compared against the $kth = 9th$ largest value.

Figure 4C shows the results for the standard DFT and OS-CFAR, and **Figure 4D** shows the results for the spiking versions. As in the 1D experiment, the traditional and the spiking approaches were able to detect the three targets.

The error between the traditional algorithm and the S-DFT resulted in an RMSE of 0.0060, which is slightly larger than in the 1D case. As the number of layers in the network increases, encoding errors accumulate and the results increasingly differ from the desired goal (Rueckauer et al., 2017). However, this error is small, and the S-DFT offers very similar output to that of the standard DFT. Moreover, the small deviations did not affect the overall detection performance.

4. CONCLUSION

In this work, we presented an SNN that approximates the first stages of a typical automotive radar pipeline i.e., the Fourier transform and object detection. In contrast to the majority of ANN and SNN applications, our network is mathematically derived from the original algorithms and is not based on learning, as the DFT and CFAR are efficient and fast methods that can be translated into a neural network. Thus, we avoid the difficulties of network verification arising from high-dimensional optimization algorithms.

The implementation of every step of the radar signal processing with SNNs is fundamental in order to unfold the full potential of SNNs and NHW. For instance, the input and output of the proposed method are binary temporal events. In the present work, this data structure is inferred with encoding and decoding techniques. However, they introduce additional overhead and consequently penalize the energy footprint of the method. Instead of generating spikes after the analog-to-digital converter, future work should focus on designing *ad-hoc* electronics for generating spike-trains directly from the voltage signal, taking inspiration from the most recent advances in neuromorphic sensors. Future work should also focus on encoding the information using temporal coding, due to the efficiency gains that it provides, as well as extending the SNN to higher-level processing stages e.g., classification, tracking, or semantic segmentation. Furthermore, it is necessary to implement the developed algorithms in NHW and conduct benchmarking experiments that compare them with traditional methods. An exhaustive assessment that evaluate performance parameters (e.g., latency, energy, and memory usage) is paramount in order to determine its real potential for automotive applications.

The application of SNNs has typically been limited to computer vision datasets or *ad-hoc* neuromorphic sensors, and, to our knowledge, this is the first implementation of an SNN to the processing pipeline of automotive radar. We anticipate a rise in the application of SNNs for radar processing in upcoming years, due to increasing interest in efficient processing chains for autonomous driving applications and the major role that radar sensors play in these systems.

DATA AVAILABILITY STATEMENT

The data for conducting the experiments is available here: https://github.com/KI-ASIC-TUM/spiking-dft-cfar.

AUTHOR CONTRIBUTIONS

JL-R designed and implemented the S-DFT. TD designed and implemented the S-OSCFAR. The manuscript was written by JL-R and TD, with support from ZB. AK supervised the project and provided the funding. All authors contributed to the article and approved the submitted version.

FUNDING

This research has been funded by the Federal Ministry of Education and Research of Germany in the framework of the KI-ASIC project (16ES0995).

ACKNOWLEDGMENTS

We would like to thank Infineon Technologies AG for providing the automotive radar simulator used for the experiments. We would also like to thank Nico Reeb, Bernhard Vogginger, Leon Bonde Larsen, Katrine Nielsen, and all the reviewers for their useful comments and suggestions.

REFERENCES

Auge, D. and Mueller, E. (2020). *Resonate-and-Fire Neurons as Frequency Selective Input Encoders for Spiking Neural Networks*. Technical Report TUM-I2083. Available online at: https://mediatum.ub.tum.de/1544346 (accessed March 25, 2021).

Bing, Z., Meschede, C., Chen, G., Knoll, A., and Huang, K. (2020). Indirect and direct training of spiking neural networks for end-to-end control of a lane-keeping vehicle. *Neural Netw*. 121, 21–36. doi: 10.1016/j.neunet.2019.05.019

Bouvier, M., Valentian, A., Mesquida, T., Rummens, F., Reyboz, M., Vianello, E., et al. (2019). Spiking neural networks hardware implementations and challenges: a survey. *ACM J. Emerg. Technol. Comput. Syst*. 15, 1–35. doi: 10.1145/3304103

Cao, Z., Cheng, L., Zhou, C., Gu, N., Wang, X., and Tan, M. (2015). Spiking neural network-based target tracking control for autonomous mobile robots. *Neural Comput. Appl*. 26, 1839–1847. doi: 10.1007/s00521-015-1848-5

Davies, M., Srinivasa, N., Lin, T.-H., Chinya, G., Cao, Y., Choday, S. H., et al. (2018). Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro* 38, 82–99. doi: 10.1109/MM.2018.112130359

Deep, Y., Held, P., Ram, S. S., Steinhauser, D., Gupta, A., Gruson, F., et al. (2020). Radar cross-sections of pedestrians at automotive radar frequencies using ray tracing and point scatterer modelling. *IET Radar Sonar Navig*. 14, 833–844. doi: 10.1049/iet-rsn.2019.0471

Frigo, M., and Johnson, S. G. (2005). The design and implementation of fftw3. *Proc. IEEE* 93, 216–231. doi: 10.1109/JPROC.2004.840301

Furber, S. B., Galluppi, F., Temple, S., and Plana, L. A. (2014). The spinnaker project. *Proc. IEEE* 102, 652–665. doi: 10.1109/JPROC.2014.2304638

Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). *Neuronal Dynamics - From Single Neurons to Networks and Models of Cognition*. Cambridge: Cambridge University Press.

Gilbert, A. C., Indyk, P., Iwen, M., and Schmidt, L. (2014). Recent developments in the sparse fourier transform: a compressed fourier transform for big data. *IEEE Signal Process. Mag*. 31, 91–100. doi: 10.1109/MSP.2014.2329131

Hasch, J., Topak, E., Schnabel, R., Zwick, T., Weigel, R., and Waldschmidt, C. (2012). Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band. *IEEE Trans. Microwave Theor. Tech*. 60, 845–860. doi: 10.1109/TMTT.2011.2178427

Jiménez-Fernández, A., Cerezuela-Escudero, E., Miró-Amarante, L., Dominguez-Morales, M. J., de Asis Gómez-Rodriguez, F., Linares-Barranco, A., et al. (2016). A binaural neuromorphic auditory sensor for FPGA: a spike signal processing approach. *IEEE Trans. Neural Netw. Learn. Syst*. 28, 804–818. doi: 10.1109/TNNLS.2016.2583223

Kamel, E. B., Peden, A., and Pajusco, P. (2017). "RCS modeling and measurements for automotive radar applications in the W band," in *2017 11th European Conference on Antennas and Propagation (EUCAP)* (IEEE), 2445–2449. doi: 10.23919/EuCAP.2017.7928266

Khan, A. T., Li, S., and Cao, X. (2021). Control framework for cooperative robots in smart home using bio-inspired neural network. *Measurement* 167:108253. doi: 10.1016/j.measurement.2020.108253

Khan, A. T., Li, S., and Li, Z. (2020). Obstacle avoidance and model-free tracking control for home automation using bio-inspired approach. *Adv. Control Appl. Eng. Indus. Syst*. e63. doi: 10.1002/adc2.63

Kronauge, M., and Rohling, H. (2013). Fast two-dimensional CFAR procedure. *IEEE Trans. Aerospace Electron. Syst*. 49, 1817–1823. doi: 10.1109/TAES.2013.6558022

Lin, S.-C., Zhang, Y., Hsu, C.-H., Skach, M., Haque, M. E., Tang, L., et al. (2018). "The architectural implications of autonomous driving: constraints and acceleration," in *Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems*, 751–766. doi: 10.1145/3173162.3173191

Liu, Y., Cai, K., Zhang, M.-H., and Zheng, F.-B. (2018). "Target detection in remote sensing image based on saliency computation of spiking neural network," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (Valencia: IEEE), 2865–2868. doi: 10.1109/IGARSS.2018.8518345

Maass, W., and Schmitt, M. (1999). On the complexity of learning for spiking neurons with temporal coding. *Inform. Comput*. 153, 26–46. doi: 10.1006/inco.1999.2806

Paredes-Vallés, F., Scheper, K. Y. W., and De Croon, G. C. H. E. (2019). Unsupervised learning of a hierarchical spiking neural network for optical flow estimation: from events to global motion perception. *IEEE Trans. Pattern Anal. Mach. Intell*. 42, 2051–2064. doi: 10.1109/TPAMI.2019.2903179

Patole, S. M., Torlak, M., Wang, D., and Ali, M. (2017). Automotive radars: a review of signal processing techniques. *IEEE Signal Processing Mag*. 34, 22–35. doi: 10.1109/MSP.2016.2628914

Piekniewski, F., Laurent, P., Petre, C., Richert, M., Fisher, D., and Hylton, T. (2016). Unsupervised learning from continuous video in a scalable predictive recurrent network. *arXiv preprint arXiv:1607.06854*.

Rohling, H. (1983). Radar CFAR thresholding in clutter and multiple target situations. *IEEE Trans. Aerospace Electron. Syst*. 19, 608–621. doi: 10.1109/TAES.1983.309350

Rueckauer, B., Lungu, I.-A., Hu, Y., Pfeiffer, M., and Liu, S.-C. (2017). Conversion of continuous-valued deep networks to efficient event-driven networks for image classification. *Front. Neurosci*. 11:682. doi: 10.3389/fnins.2017.00682

Salvatore, N., Mian, S., Abidi, C., and George, A. D. (2020). "A neuro-inspired approach to intelligent collision avoidance and navigation," in *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)* (IEEE), 1–9. doi: 10.1109/DASC50938.2020.9256492

Sangwan, V. K., and Hersam, M. C. (2020). Neuromorphic nanoelectronic materials. *Nat. Nanotechnol*. 15, 517–528. doi: 10.1038/s41565-020-0647-z

Stagsted, R. K., Vitale, A., Renner, A., Larsen, L. B., Christensen, A. L., and Sandamirskaya, Y. (2020). "Event-based pid controller fully realized in neuromorphic hardware: a one dof study," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Las Vegas, NV: IEEE), 10939–10944. doi: 10.1109/IROS45743.2020.9340861

Strubell, E., Ganesh, A., and McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *arXiv preprint arXiv:1906.02243*. doi: 10.18653/v1/P19-1355

Wang, W., Zhou, S., Li, J., Li, X., Yuan, J., and Jin, Z. (2021). "Temporal pulses driven spiking neural network for time and power efficient object recognition in autonomous driving," in *2020 25th International Conference on Pattern Recognition (ICPR)* (Milano: IEEE), 6359–6366.

Winner, H., Hakuli, S., Lotz, F., and Singer, C. (2014). *Handbook of Driver Assistance Systems*. Springer International Publishing. doi: 10.1007/978-3-319-09840-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurorobotics

REVIEW
published: 09 June 2021
doi: 10.3389/fnbot.2021.658280

Robotics Dexterous Grasping: The Methods Based on Point Cloud and Deep Learning

Haonan Duan [1,2,3], Peng Wang [1,3,4], Yayu Huang [1,3], Guangyun Xu [1,3], Wei Wei [1,3] and Xiaofei Shen [1,3]*

[1] The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, [2] Department of Information Science, School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, United States, [3] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China, [4] Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China

Dexterous manipulation, especially dexterous grasping, is a primitive and crucial ability of robots that allows the implementation of performing human-like behaviors. Deploying the ability on robots enables them to assist and substitute human to accomplish more complex tasks in daily life and industrial production. A comprehensive review of the methods based on point cloud and deep learning for robotics dexterous grasping from three perspectives is given in this paper. As a new category schemes of the mainstream methods, the proposed generation-evaluation framework is the core concept of the classification. The other two classifications based on learning modes and applications are also briefly described afterwards. This review aims to afford a guideline for robotics dexterous grasping researchers and developers.

Keywords: robotics, dexterous grasping, point cloud, deep learning, review

Edited by:
Yan Wu,
*Institute for Infocomm Research (A*STAR), Singapore*

Reviewed by:
Yingbai Hu,
Technical University of Munich, Germany
Qujiang Lei,
Chinese Academy of Sciences, China

****Correspondence:***
Peng Wang
peng_wang@ia.ac.cn

Received: *25 January 2021*
Accepted: *14 May 2021*
Published: *09 June 2021*

Citation:
Duan H, Wang P, Huang Y, Xu G, Wei W and Shen X (2021) Robotics Dexterous Grasping: The Methods Based on Point Cloud and Deep Learning. Front. Neurorobot. 15:658280. doi: 10.3389/fnbot.2021.658280

INTRODUCTION

In the last decades, there has been an enormous proliferation in robotic community, both at in terms of research and attracting boundless varieties of imagination of general public, due to its diverse possibilities. The vast majority of robots in operation today consist of 6 degree of freedom (6-DOF) which are either rotary (articulated) or sliding (prismatic), with a simple end effector for interacting with the workpieces (Murray et al., 1994). Robot manipulation means it can use and control different objects according to certain specifications and essentials through the end effector to achieve the effect of making the best use of playing the role of object itself (Okamura et al., 2000; Saut et al., 2007). Grasping, as one of the most primitive manipulations, almost all high-level operations and complex tasks that people expect robots to complete are inseparable from the assistance of it. With the in-depth development of robotics, researchers begin to facilitate the transition from simple or even crude grasping of robots with less discrimination of objects to object-oriented dexterous grasping. Unlike simple grasping, dexterous grasping is able to determine which posture to be employed to grasp where of the object to ensure a higher grasping success rate (Ciocarlie et al., 2007; Prattichizzo et al., 2008; Ciocarlie and Allen, 2009).

The research on dexterous manipulation and grasping can be traced back to the 1980s. In the era when deep learning methods were not yet established, researchers came up with ideas for dexterous manipulations from the perspective of physics and geometry. Through the kinematic modeling of robots, plenty of research results that attracted widespread attention at the time were born (Moreno et al., 2011; Fischinger et al., 2015; Chen et al., 2016; Zhou Z. et al., 2018; Zito et al., 2019;

Monica and Aleotti, 2020). However, robot grasping algorithms based on physics and geometry presuppose many assumptions, making these methods hard to generalize. With the substantial increase in computing power from hardware, the artificial intelligence surge represented by deep learning methods has quickly penetrated into various research fields. Free from the limitation of manually extracting features, the grasping algorithms based on deep learning have achieved insurmountable effects in all aspects by traditional approaches, taking the robot's intelligence to a higher level. Specifically, with RGB images or depth images as input, robotic grasping based on convolutional neural network (CNN) which is a dominant deep learning framework in the field of computer vision, has obtained high grasping success rates in many tasks (Lenz et al., 2015; Varley et al., 2015; Johns et al., 2016; Finn and Levine, 2017; James et al., 2017; Kumra and Kanan, 2017; Zhang et al., 2017; Dyrstad et al., 2018; Levine et al., 2018; Schmidt et al., 2018; Schwarz et al., 2018). As shown in **Figure 1**, nowadays, based on visual information, robot dexterous grasp learning can be roughly divided into two categories based on whether the learning process is based on trial and error. Dexterous grasping learning that is not based on trial and error will determine the best grasp posture based on the visual information of the scene, and then execute it. On the contrary, the dexterous grasping learning based on trial and error is to accumulate the experience of grasping from failure through the interaction between the robot and the environment to improve grasping dexterity.

As a commonly used 3D visual data form, point cloud can be generated by 3D laser scanner (LIDAR), depth sensors or RGB-D images (Liu W. et al., 2019; Guo Y. et al., 2020; Lu and Shi, 2020). Compared with RGB or depth images, 3D point clouds can store more spatial information as their higher dimensions. With the improvement of point cloud processing methods (Fischler and Bolles, 1981; Rusu et al., 2010; Rusu and Cousins, 2011; Aldoma et al., 2012; Chen et al., 2016) and the introduction of CNN based on point cloud as input (Wu et al., 2015; Qi et al., 2017a,b), point clouds have become increasingly common for those tasks based on visual perception. Meanwhile, as more and more contributions on datasets of grasping based on point cloud (Goldfeder et al., 2009; Calli et al., 2015a,b, 2017; Kappler et al., 2015; Mahler et al., 2016, 2017; Depierre et al., 2018; Bauza et al., 2019; Bottarel et al., 2020; Fang H.-S. et al., 2020), robotic dexterous grasping based on point cloud and deep learning set off a tremendous wave of research in the field of robotics.

Based on the current work of robot dexterous grasping combining with point cloud and deep learning, this paper summarizes relevant important work from 2015 to present. As the earliest state-of-the-art work in related research, grasp pose detection (GPD) (Pas and Platt, 2015; Gualtieri et al., 2016; Pas et al., 2017) samples various grasp poses (candidate generation) in the point cloud and employ a deep learning method to assess these grasps (candidate evaluation) to obtain the optimal grasp pose. Inspired by this efficient pipeline, most subsequent works followed this framework to ameliorate generation or evaluation stages. From this perspective, this survey proposes a more generalized framework and summarizes relevant work as one or both of these two stages, that is, in which step the work contributes more. The articles reviewed in this paper are all based on deep learning framework. If deep learning strategy is not utilized in the generation stage, it will appear in evaluation stage, and vice versa. Remaining uncategorized models are provided in a separate section afterwards. The taxonomies from different perspectives of learning modes and applications are also briefly described. This paper aims to provide valuable insights and inspirations through the proposed taxonomy.

The remainder of the paper is organized as follows. Section Proposed Taxonomy presents related surveys and proposed category scheme. Section Grasping Candidate Generation and Grasp Candidate Evaluation are about the methodologies of grasp candidate generation and evaluation respectively. The

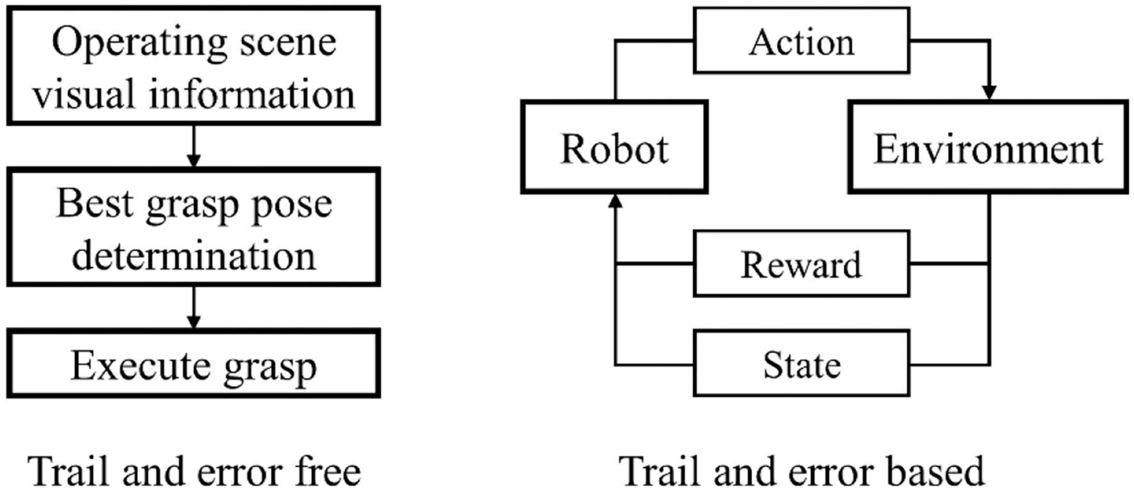


FIGURE 1 | Recent dexterous grasp pipeline.

TABLE 1 | The related surveys and corresponding topics.

References	Review topic	Journals
Du et al., 2020	Vision methods facilitate grasp estimation	Artificial intelligence review
Ruiz-del-Solar et al., 2018	Deep learning methods for robot vision	arXiv
Luo et al., 2017	Robotic tactile perception	Mechatronics
Wang C. et al., 2020	Feature sensing and robotic grasping	Sensors
Caldera et al., 2018	Deep learning methods in grasp detection	Multimodal technologies and interaction
Kroemer et al., 2019	Learning-based methods in robot manipulation	arXiv
Kleeberger et al., 2020	Learning-based robotic grasping	Current robotics reports
Li and Qiao, 2019	Robotic grasping and assembly tasks	IEEE Transactions on mechatronics
Mohammed et al., 2020	Deep reinforcement learning-based grasping	IEEE Access
Zhao W. et al., 2020	Sim-to-real problems of reinforcement learning	arXiv
Billard and Kragic, 2019	Trends and challenges in robot manipulation	Science

uncategorized papers out of proposed framework are described in section End-to-End and Others. Section Learning Modes summarizes the methods from learning modes. In section Applications, it mainly introduces the applications of related approaches from perspectives of end effectors and operating scenarios. Section Challenges and Future Directions provides the challenges and future direction of this field. Section Conclusion is the conclusion of the paper.

PROPOSED TAXONOMY

As listed in **Table 1**, there are already numerous surveys in the field of robotics learning. Some surveys elaborate the perception techniques of robotics, and some others introduce approaches of robot manipulation. Ruiz-del-Solar et al. (2018) and Du et al. (2020) pay attention to the vision methods for robot manipulations. With the exception of visual perception approaches, Luo et al. (2017) and Wang C. et al. (2020) exhibit there are many other perception methods can help improve robot performance. Caldera et al. (2018), Kroemer et al. (2019), Li and Qiao (2019), and Kleeberger et al. (2020) focus on the overview of robot manipulation methods based on deep learning. Mohammed et al. (2020) and Zhao W. et al. (2020) introduce the techniques in robot learning on the basis of reinforcement learning. Billard and Kragic (2019) describes the trends and challenges in robot manipulation.

Unlike the works mentioned above, this paper focuses on the methods of robotic dexterous grasping based on point cloud and deep learning. Compared with previous related reviews, this paper narrows the reviewed works scope through the limitation of inputs and methods, aiming to provide a more detailed description in a specific direction. For the robotic grasping algorithms, the following four classifications have appeared in past researches but not entirely suitable for the topic in this paper. (1) Analytic-based and empirical-based: because this survey pays attention to the use of deep learning, traditional analytic methods are not within the scope of the review, so this classification is not applicable. (2) Task-agnostic and task-specific: this will complicate the classification of reviewed approaches mainly focus on grasping in this paper. Some task-specific papers will outperform under some specific circumstances, but don't have generalization ability, which cannot support a category. (3) Vision-based and vision-free: since this paper is specifically aimed at point cloud-based robotic grasping, most of the networks used are CNN-based, even numerous methods don't explicitly perform object recognition, segmentation, or pose estimation. In other word, the reviewed articles in this paper can be said use visual information explicitly or implicitly. Therefore, this classification is not appropriate. (4) Learning-based and learning-free: this classification is similar to (1). If each stage in the proposed pipeline is not based on learning, this method will not be taken into account in this survey.

In order to elaborate the subject more comprehensively, instead of adopting the existing classification methods, this survey classifies related work from three perspectives: generation-evaluation, learning modes, and applications as shown in **Figure 2**. The proposed generation-evaluation framework is the core part of this paper. In the stage of grasp candidate generation, the methods can be divided into geometry-based sampling and learning-based sampling. In geometry-based sampling, methods are categorized into object-agnostic sampling and object-aware sampling. In object-aware sampling, object detection and segmentation are the most basic methods, object affordance detection and object shape complement are the further improvements. The learning-based sampling can also be divided into object-agnostic sampling and object-aware sampling. Unlike the methods in geometry-based object-aware sampling, there is no more classifications in the branch of learning-based object-aware sampling. In the grasp candidate evaluation stage, methods are split into learning-free and learning-based. This paper will use two separate sections to elaborate the methods of grasp candidate generation and evaluation. The following section will introduce some work that cannot be classified as this framework. They fall into end-to-end learning-based, reinforcement learning-based and others. Two remaining classifications will be briefly explained afterwards.

GRASPING CANDIDATE GENERATION

Grasp candidate generation, also called grasp pose sampling, refers to randomly sampling the parameters (the approaching direction of the end effector, opening size, joint angle, etc.) of

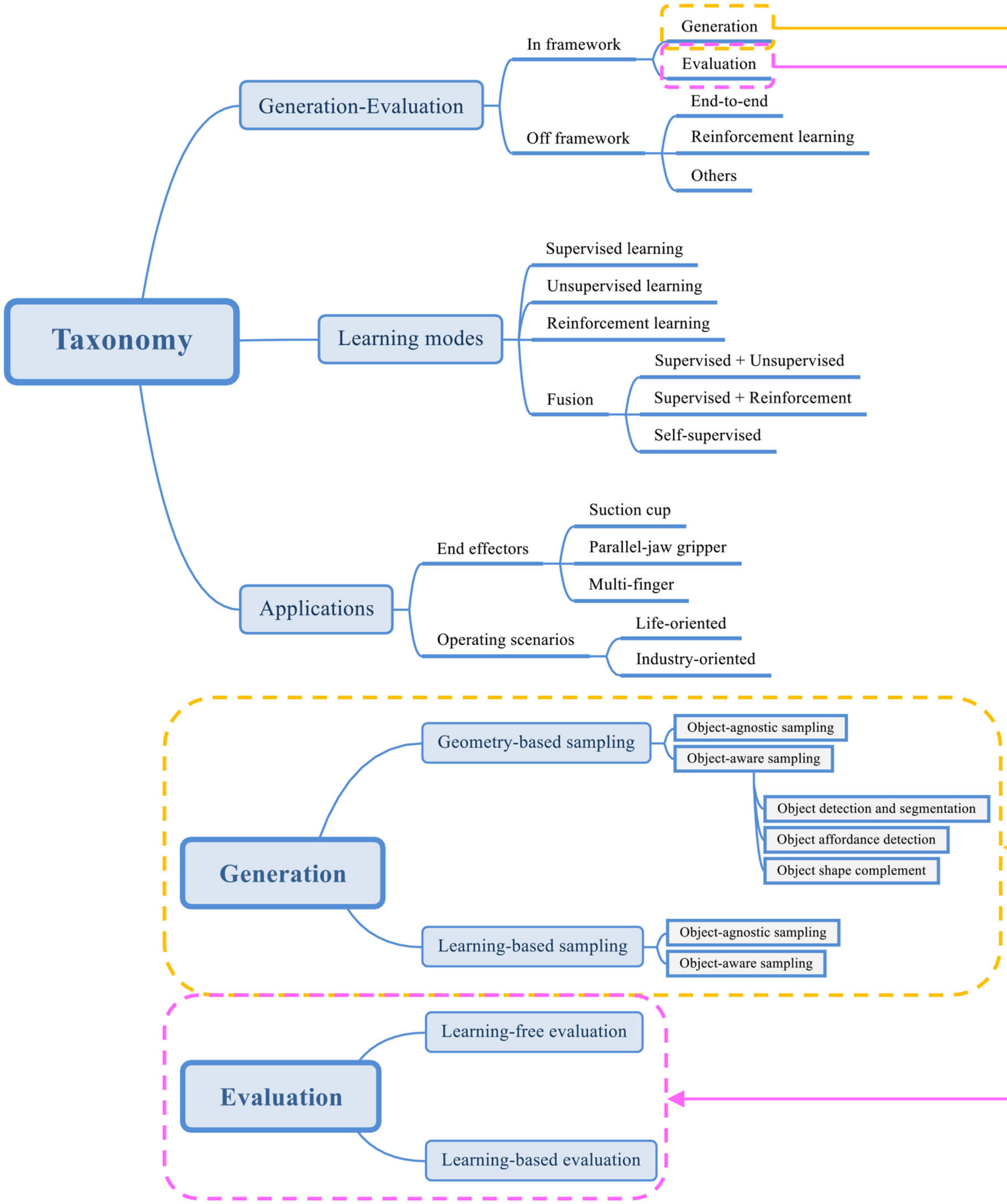


FIGURE 2 | Robotics dexterous grasping methods based on point cloud and deep learning.

the end effector on the target object within a specific range to obtain a large number of possible grasp gestures as shown in **Figure 3**. In order to ensure the optimal or suboptimal grasp posture can be found, the final executed grasp pose is commonly not directly calculated, but a large number of grasp candidates are sampled on numerous points in point cloud with random grasp configurations, so as to promise the reliability of results on the basis of quantitative advantages (Eppner et al., 2019). In general, sampling can be roughly divided into geometry-based sampling and learning-based sampling. In these two categories, approaches

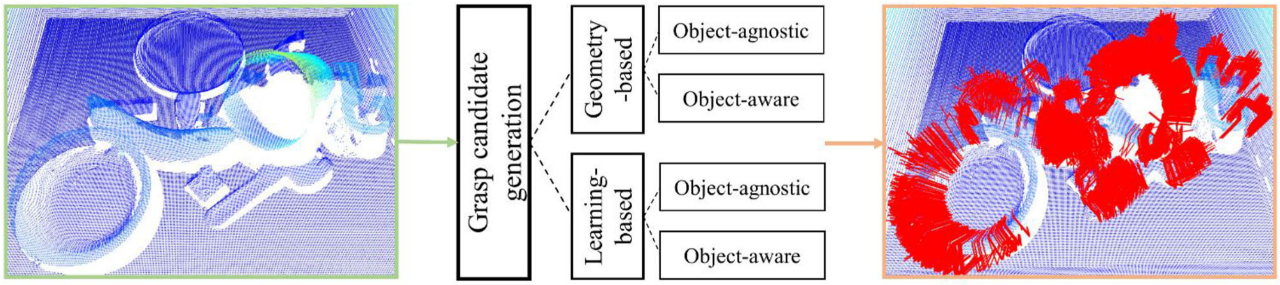


FIGURE 3 | The general pipeline of grasp candidate generation.

can be deeply split into object-agnostic sampling and object-aware sampling. The papers mainly contribute to generation stage are provided in this section.

Geometry-Based Sampling

In the early work, researchers try to apply random sampling to gain grasp candidates, the feasibility of sampling method to generate reliable and reasonable grasping candidates is discovered and verified afterwards (Boularias et al., 2014, 2015). Based on this, with superior interpretability and intuitiveness, the geometry-based grasp pose sampling came into being (Pas and Platt, 2015; Wang and Ling, 2016; Pas et al., 2017). These methods add geometric restrictions to the sampling process, such as the grasp position as close as possible to the center of gravity of the object, or the size of the object in the grasp approaching direction cannot exceed the maximum width of the end effector, so that make the random generation more reasonable. The idea behind this method is to introduce manually calculated features into the hypothesis space of grasping candidates, improves the possibility of finding the optimal or suboptimal grasp posture. Thanks to its generation process is based on the modeling of the grasping model from the real world (Murray et al., 1994), although it is an earlier branch of the generation method, geometry-based sampling is still being adopted by a lot of work.

In general, the current work using geometry-based grasping pose sampling can be divided into two categories based on whether it has the specific information of the object to be grasped, namely, object-agnostic sampling and object-aware sampling. Since the point cloud describes the spatial information of the scene, for the object-agnostic sampling methods, even if there is no specific information of the object, they can also obtain suitable sampling points for generating grasping poses by the spatial information. For the methods of object-aware sampling, they will first extract the specific information of the object through the methods of computer vision, relying on pure point cloud or combining the information of RGB images and depth images, and then perform sampling on these higher-level information.

Object-Agnostic Sampling

Sampling based on physical and geometric constraints could have an impressive performance on robotic grasping, especially in the two parallel-jaw grippers community since mathematical models of robotic grasping was well-defined in the past few years (Murray et al., 1994; Okamura et al., 2000; Prattichizzo et al., 2008; Prattichizzo and Trinkle, 2016). Researchers propose numerous sampling methods by combining the established mathematical theories and task-specific conditions. Since the specific information of the object is not used to help the sampling, this kind of object-agnostic grasping pose sampling must rely on complete grasping mathematical modeling and a large number of physical and geometric constraints in specific tasks to achieve grasping candidate generation. For example, only points higher than the operating plane calculated through the point cloud may be belonging to the object, and then it is possible to result in some feasible grasping poses by sampling at these points. Although the calculation steps are relatively cumbersome and plenty of limitations are based on experience and modeling, in the age when deep learning has not yet been developed, there are still many attempts at this sampling method.

Pas and Platt (2015) and Wang and Ling (2016) define hypotheses contain position and orientation information of the graspable point, as well as its neighborhood points calculated by Taubin quadric fitting (Taubin, 1991; Pas and Platt, 2013). To make the sampling process more flexible, GPD (Pas et al., 2017) selects N points uniformly at random from the region of interest (ROI) of point cloud and then perform grid search on picked points to extract grasp configurations that satisfy the geometric reasoning. Several research works obtain impressive experimental results by adopting this idea (Mahler and Goldberg, 2017; Mahler et al., 2017, 2018; Viereck et al., 2017; Liang et al., 2019). Lou et al. (2020) takes a further step by randomly associate a pose with each sampled point. There are also some methods not using uniformly sampling strategy. Gualtieri et al. (2016) and Kiatos and Malassiotis (2019) calculate a surface normal and an axis of major principal curvature of the object surface in the neighborhood of the sampled point. Grasp candidates are then generated at regular orientations orthogonal to the curvature axis. Zhou et al. (2019) samples the grasping candidates based on the depth descriptor Depth Likelihood Volume (Zhou Z. et al., 2018).

Since the specific information of the target object is not required, the advantages of the object-agnostic sampling method in some aspects are very obvious. These methods do not require object segmentation, thereby avoiding errors caused in the segmentation stage that will affect the accuracy of subsequent grasp candidate generation. At the same time, these methods do not need to know the identity, class and shape of the target object, which also makes it possible to apply the grasp poses sampling algorithms on unknown objects in an open environment. Finally, these methods do not try to register a CAD model of the object to the point cloud, which could be very challenging. However, the strategy of not combining the specific information of the object will also bring inevitable shortcomings to these methods. First of all, these methods have high requirements for input. The quality of the point cloud will immediately affect the reliability of the candidate sampling results. This also indirectly causes the result of sampling on the point cloud extracted by the monocular camera to be much worse than that of the multi-view camera. Since the algorithm does not segment the objects, it can detect "grasps" that treat multiple objects as a single atomic object. This type of error is unusual with small-aperture hands, but one would expect it to become a more significant problem when the hand is physically capable of grasping larger objects. This also makes it reasonable for the algorithm's sampling results to perform poorly in cluttered scenes and collision detection. Moreover, these methods are going to become infeasible for the multi-finger robots. Almost all of techniques mentioned above focus on robots with a vacuum cup or parallel-jaw gripper for their end-effectors. Multi-finger end-effectors will introduce a rich-contact points problem which is more difficult for these methods to handle.

Object-Aware Sampling

Even object-agnostic sampling is able to find numerous grasp candidates, it doesn't use the complete information in point cloud. The drawback causes the low sampling accuracy and time-consuming sampling process. To avoid these defects, researchers propose object-aware sampling which aims to combine the specific information of the object to enhance the reasonability of search space in point cloud (Boularias et al., 2015; Zapata-Impata et al., 2017; Lopes et al., 2018). The search space refers to the points that need to be considered in the point cloud space for the grasping pose sampling algorithm. Since the object-agnostic sampling does not have the pose information of the object, those methods have to calculate all the points in the scene point cloud one by one. This is undoubtedly a brute force search method, and the efficiency of the algorithm itself cannot be ideal. The object-aware sampling method combined with the object pose information will eliminate the points that are impossible to generate grasp candidate from the point cloud based on the corresponding extracted features, that is, reduce the sampling space from the entire point cloud space to a specific space, cut down the number of invalid sampling and searching, improves the efficiency of the algorithm. According to the different acquisition of object pose information, the object-aware sampling method can be divided into three branches: object detection and segmentation, object affordance, and object shape complement. The three methods not only outperform object-agnostic sampling method in general scenarios, but can generate highly reliable pose candidates in their respective applicable environments. The pipelines of the relevant methods are provided in **Figure 4**.

Object Detection and Segmentation

The method based on object detection and segmentation is the earliest one of the three branches. This method first extracts the pose features of the object in the scene by taking the RGB images or the depth images of the scene or directly using the point cloud as input, and then segment the point cloud space to obtain a smaller and more reliable search space based on these features. Due to the introduction of the specific information of the object, this type of sampling algorithm has a significant improvement in the performance of the cluttered environment and collision detection. At the same time, with the rapid development of object detection and segmentation algorithms, YOLO (Redmon et al., 2016) and other efficient and easy-to-deploy backbone networks are widely used in this sampling method, and there is still a lot of work around its ideas.

By adopting sampling method in GPD proposed by Pas et al. (2017), Lopes et al. (2018), Schnaubelt et al. (2019), Bui et al. (2020), Chen et al. (2020), and Deng et al. (2020) sample the grasp points in point cloud for candidates generation. Lopes et al. (2018) find the largest planar surfaces which is infeasible for grasping by using RANSAC (Fischler and Bolles, 1981) and isolates the closest object to the camera from the rest of the scene to obtain object segmentation based on min-cut (Golovinskiy and Funkhouser, 2009). This work compares the experiments before and after reducing the point cloud search space, and proves that the grasping success rate has increased from 45 to 90%. Although the object's pose information is extracted with the help of object segmentation methods that are not based on deep learning, the impressive results show that the improvement is considerable. This shows that reducing the point cloud search space is a very reasonable and correct direction. Subsequently, more grasp candidate sampling methods integrated with object segmentation based on deep learning have been used in the development of this branch. Schnaubelt et al. (2019) segments the depth image by using Maskfusion (Runz et al., 2018) combined with increased noise robustness (Ückermann et al., 2012) and Bui et al. (2020) extracts object segmentation from point cloud with region of interest (ROI) obtained from YOLOv3 (Redmon and Farhadi, 2018). Deng et al. (2020) detects and segments the object from RGB-D images based on PoseCNN (Xiang et al., 2017), then a sampling method in Eppner et al. (2019) is adopted to generate 100 candidates for assessment and execution. Chen et al. (2020) utilizes object segmentation for mask-guide to improve the precision of sampling. Lin and Cong (2019), Lin et al. (2019), Sun and Lin (2020), and Yu S. et al. (2020) follow the same idea in GPD with additional physical or geometric constraints. Lin and Cong (2019) and Yu S. et al. (2020) adopt variant of PointNet (Qi et al., 2017a) and RANSAC for object segmentation respectively, then sample the grasp candidates with integrating physical and geometric constraints. In specific, the former work mainly considers the mechanical constraints in physics other than paying more attention to spatial constraints in the latter work. Lin et al. (2019) and Sun and Lin (2020) achieve object estimation via PPR-net (Dong et al., 2019) and Mask R-CNN (He et al., 2017) correspondingly followed by sampling the candidates based on using the closest ring of the centroid of the object. The main contribution and grasping accuracy are provided in **Table 2**.

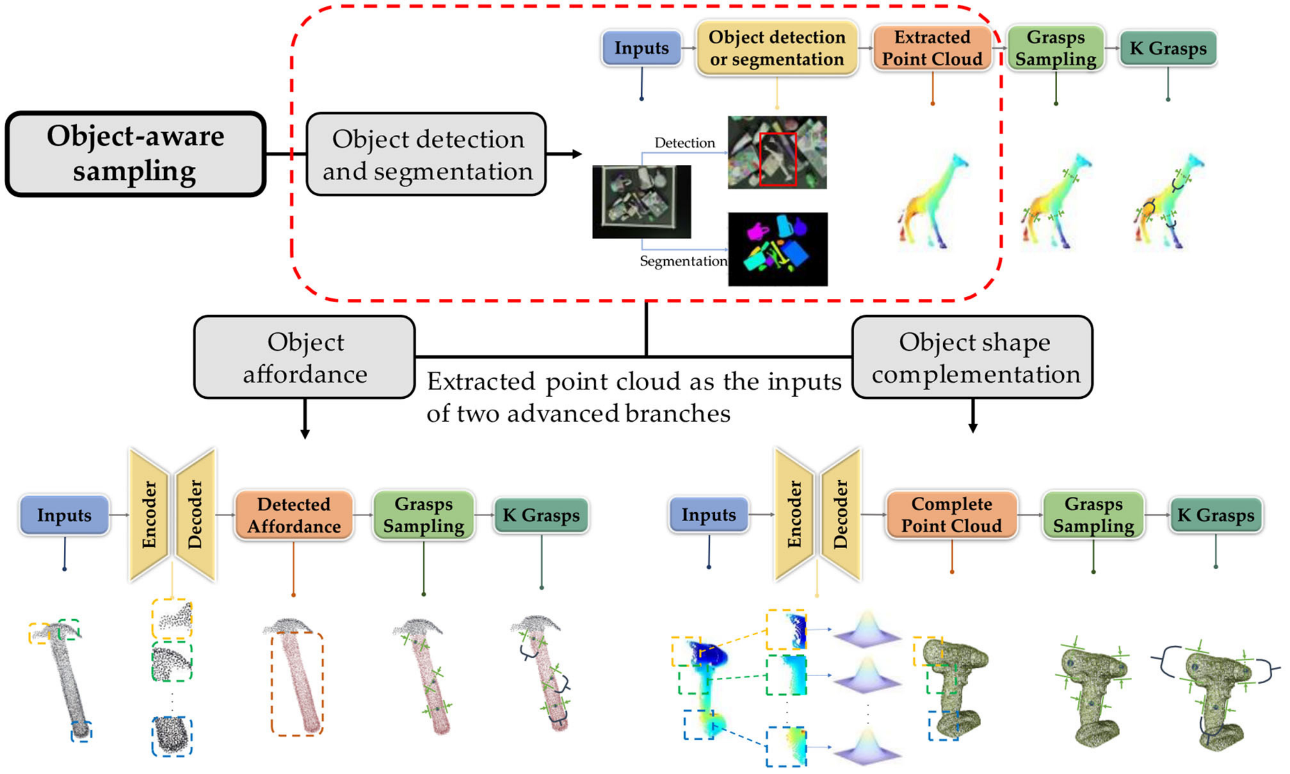


FIGURE 4 | Entire pipelines of three classifications in object-aware sampling. Object detection and segmentation is the most basic method. Inputs are commonly RGB images, which are detected or segmented by networks to extract the object point clouds. Extracted point cloud can either be utilized to sample the grasp candidates immediately or fed into object affordance or shape complementation methods. Object affordance methods take extracted point cloud as inputs to obtain the affordance of object to reduce the sampling search space. On the contrary, object shape complementation aims to acquire the entire object point cloud to improve the grasp candidate generation confidence (The hammer point cloud is from YCB datasets).

Object Affordance

Affordance is introduced by Gibson (2014), which describes how likely the agent is capable to execute an action based on its surrounding environment. In robotics community, affordance, as a new physical and geometric property of objects, refers to the part of the objects with high probability to be operable. Specifically, affordance refers to the most likely part of the objects to make the grasping successful determined based on the knowledge of human grasping habits. Image that, people will always hold the handle of the hammer instead of the hammerhead when picking up a hammer, or hold the apple in the hand instead of grasping the apple stem when picking up an apple. Grasping pose sampling based on object affordance is an advanced method developed from object detection and segmentation methods. This method will determine the operable part with high possibility on the detected object. Since the search space in the point cloud is further reduced, the sampling results are more reasonable than that of grasp candidate sampling on the entire object. Concretely, if the target is a knife, if grasp sampling is performed on the entire object, those candidate grasps located on the blade will inevitably cause damage to the end effector of robots or other objects in the environment. Based on several methods to learn and understand object affordance proposed by prior art (Varadarajan and Vincze, 2012; Koppula et al., 2013; Katz et al., 2014; Zhu et al., 2014, 2015; Do et al., 2018), there has also been a lot of work in the grasp candidate generation based on the direction of object affordance.

Inspired by sampling integrated with traditional affordance detection methods (Pas and Platt, 2016; Kanoulas et al., 2017; Liu C. et al., 2019), diverse deep learning-based affordance-based sampling techniques are proposed. Qian et al. (2020) employs ResNet101 (He et al., 2017) with feature pyramid network (FPN) (Lin et al., 2017) to perform affordance detection and applies the sampling method proposed in Pas et al. (2017) with refined local reference frame computation. Instead, Fang K. et al. (2020) finds object affordance implicitly based on Mar et al. (2017) with a multi-dimensional continuous action space and uniformly samples grasps from the object surface using antipodality constraints (Mahler et al., 2017). Manuelli et al. (2019) detects keypoints of object affordance together with local dense geometric information instead of segmenting entire affordance, the reduced search space is able to guarantee a high-quality grasp candidates sampling (Gualtieri et al., 2016; Mahler et al., 2016, 2019).

Affordance learning is an advanced variant of basic segmentation method. Different forms of affordance such

TABLE 2 | The summary of geometry-base object-aware grasp candidate generation.

Method category	Work	Method backbone	Success rate (%)	End-effector	Environment				Simulation /reality
					Object arrangement	Object number	Object shape	Test novel object	
Object detection and segmentation	Less is More (Lopes et al., 2018)	RANSAC	90	–	Single object	1	Irregular	No	R
	(Schnaubelt et al., 2019)	Maskfusion	–	–	Cluttered	5	Irregular	No	R
	RED (Chen et al., 2020)	Mask-RCNN + PointNet	84 (S) 82 (R)	Parallel-jaw gripper	Cluttered	7	Irregular	Yes	S/R
	(Bui et al., 2020)	YOLOv3	–	Parallel-jaw gripper	Single object	1	Regular	No	S/R
	(Deng et al., 2020)	PoseCNN	86.7	Parallel-jaw gripper	Cluttered	–	Irregular	Yes	R
	(Lin and Cong, 2019)	PointNet	90	Parallel-jaw gripper	Cluttered	5	Irregular	No	R
	(Yu S. et al., 2020)	RANSAC + VGG	–	Parallel-jaw gripper	Single object	1	Irregular	No	R
	(Lin et al., 2019)	PPR-net	78	Parallel-jaw gripper	Cluttered	30	Regular	No	R
	(Sun and Lin, 2020)	Mask R-CNN	71.1	Parallel-jaw gripper	Single object	–	Regular	No	R
Object affordance	(Qian et al., 2020)	ResNet101 + FPN	95	Parallel-jaw gripper	Single object	1	Regular	No	R
	TOG-Net (Fang K. et al., 2020)	SOM	80	Parallel-jaw gripper	Single object	1	Irregular	No	R
	kPAM (Manuelli et al., 2019)	Integral human pose regression	–	Parallel-jaw gripper	Single object	1	Regular	No	R
Object shape complement	(Varley et al., 2017)	CNN	93.33	Three fingers	Cluttered	–	Irregular	No	R
	(Lundell et al., 2019)	CNN	59	Parallel-jaw gripper	Cluttered	10	Irregular	No	R
	(Yan et al., 2019)	CNN	61	Parallel-jaw gripper	Cluttered	–	Irregular	Yes	R
	(Torii and Hashimoto, 2018)	DNN	85.6	Parallel-jaw gripper	Cluttered	–	Regular	No	S
	(Liu and Cao, 2020)	CNN	94.06	Parallel-jaw gripper	Cluttered	–	Irregular	Yes	R

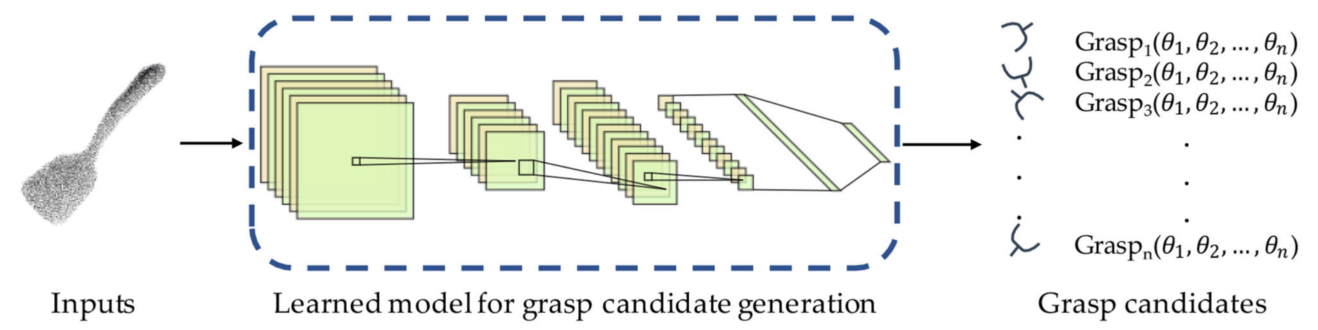


FIGURE 5 | Learning-based sampling (The spatula point cloud is from YCB datasets).

as semantic labels (Zhu et al., 2014), spatial maps (Jiang et al., 2012), and motion trajectories (Zhu et al., 2015) are suitable to diverse tasks, which is in a position to further reduce the search space and improve the confidence of candidates generation.

Object Shape Complement

Shape complementation is another improved variant of segmentation method. Unlike affordance learning attempts to understand the graspable components from detected objects, shape complement pays more efforts on "looking" the entire target object more completely. As another advanced branch developed from object detection and segmentation methods, shape complement and object affordance have completely different thoughts. Grasp candidate sampling based on object affordance is to reduce the sampling space to the area where the grasp is most likely to succeed, while shape completion is to try to minimize the occurrence of unreasonable sampling by obtaining more information about the shape of the object. In particular, for the point cloud captured by a monocular camera, it is impossible to outline the shape of the object where the light cannot reach. For symmetric objects, the effect of shape complement may not be so obvious, but for asymmetric objects, this type of methods is particularly conducive.

Varley et al. (2017) proposes a convolutional neural network which takes voxelized partial mesh of object as input and output the complemented shape. After a few post-preprocessing, GraspIt! (Miller and Allen, 2004) is used to generate grasp candidates. Lundell et al. (2019) improves the network architecture based on the method in Dai et al. (2017) by adding Monte-Carlo (MC)-Dropout (Gal and Ghahramani, 2016), an advanced dropout layer (Srivastava et al., 2014), into both training and run-time step to generate a set of shape samples. Grasp candidates sampling by GraspIt! is employed on the mean of shape samples. Yan et al. (2019) reconstructs object point cloud by integrating the segmentation via Mask R-CNN and several encoder-decoder modules (Fan et al., 2017) inspired by single-view 3D-object reconstruction (Jiang et al., 2018). Differ from complementing point cloud of object, Liu and Cao (2020) and Torii and Hashimoto (2018) leverage object primitives to simplify object shape under the detection output of convolutional neural network. Compared with GPD, their experiment results increased by 10.56 and 18%, respectively.

Although existing shape complement methods commonly accompany with high uncertainty, being aware of object shape is capable of incredibly facilitating the accuracy, robustness, and confidence of grasp proposals generation.

Object-aware methods aim to reduce the search space by being aware of the specific object, which improve the sampling performance. As shown in **Table 2**, the summary of geometry-based object-aware sampling method is listed.

Learning-Based Sampling

Geometry-based sampling methods generate candidate grasp poses by changing the grasp configuration randomly under physical and geometric constraints in specific task, however, sampling a number of grasps poses proposals is computationally expensive. Furthermore, sampling the rotational or translational dimension possibly produce some unstable and unreasonable grasp configurations since the conditions in the grasp modeling are only artificially extracted. More recently, deep learning techniques improve performances in many traditional analytic tasks greatly based on more powerful feature extraction abilities compared to human handcraft. On the strength of deep learning, some researchers move from physical and geometric reasoning-based sampling to deep learning-based sampling as example in **Figure 5**. Different from the geometry-based sampling method, the learning-based sampling method will complete the learning based on a dataset during training. The learning-based methods can be divided into supervised learning or unsupervised learning according to the learning model. For sampling methods based on supervised learning, it takes scene point clouds or RGB images as input to obtain sampling results directly or extract appropriate grasp points first and then place the grasp posture on these points. It compares the prediction result with the ground truth to calculate the difference in sampling point determination, the difference in hand posture prediction. The sampling method based on unsupervised learning learns the distribution of sampling from the training data by the generative model, reducing the KL divergence or JS divergence to reduce the distribution difference between the data generated by the model and the training data. The trained model will be used for sample the grasp poses. Similar to geometry-based sampling, learning-based sampling methods fall into object-agnostic and object-aware group.

Object-Agnostic Sampling

Object-agnostic methods take the point cloud as input and generate the proposals by learned models without detecting the object in the point cloud. Compared with geometry-based object-agnostic sampling, although the learning-based object-agnostic sampling method also searches in the entire scene point cloud, it has the following improvements on the thoughts: (1) Geometry-based sampling determines whether the point is feasible to generate grasp pose only based on the spatial information in the point cloud (three coordinate values of x, y, z), while learning-based sampling can learn more high-level information through the models or networks to help improve accuracy. (2) the generalization of geometry-based sampling is not strong, especially for cluttered scenes, it often regards two very close objects as one, which leads to unreasonable results of grasp sampling. The learning-based sampling method is able to significantly improve this problem through the benefits of more extracted features. (3) Geometry-based sampling is essentially a simple random combination of various parameters. Although it has an advantages in amount, it also caused a sharp increase in the number of negative samples. In contrast, the sampling process of learning-based generation method is actually obtained by the prediction of the model, and the sampling result will contain more information about the effect of high-level features. With the development of deep learning, some researchers try in the direction of learning-based object-agnostic sampling methods using neural networks.

Jiang et al. (2020) proposes a deep convolutional neural network (DCNN) to predict the set of grasp points from the input depth-image. Inspired by Varley et al. (2015) that obtains grasps on pixels, Morrison et al. (2018a) presents a Generative Grasping CNN (GG-CNN) which generates candidates immediately on pixelwise. GG-CNN treats each pixel of image liberally without any hypothetical searching space, which may assure a higher probability of finding a global optimal grasp pose. Guan et al. (2019) adopt Fully Convolutional Neural Network (FCNN) (Long et al., 2015) to take four channels images synthesized by an RGB image and depth image as input and output three maps contain all information of potential grasp poses.

Theoretically, the learning-based object-agnostic sampling method should be better than the geometry-based grasp candidate generation method, but there is not much work in this direction. Moreover, in addition to some advantages in the efficiency of algorithm operation, there is no remarkable improvement in other aspects. This is mainly due to the limitations of this method: (1) The training dataset is difficult to generate. Since the grasping pose needs to be sampled by the prediction of the model, the ground truth label of the sample is difficult to represent. This results in the difficulty of model training. Although the careful design of the loss function and network structure can slightly improve this shortcoming, it may also cause a decrease in efficiency during training and testing. And some work is done by simplifying the objects in the dataset to complete the training, but this causes the generalization of the model to be very poor. (2) Although the model is used to complete the sampling process, due to the lack of specific information of the object, the essence of the algorithm is to perform a brute force search in the entire point cloud space. (3) The neural network-based model can indeed extract higher-level features, but because of the difficulty of ground truth representation and the lack of object specific information, the extracted features may not be too satisfactory, which leads to the performance of the algorithm is not outstanding.

Object-Aware Sampling

Geometry-based object-aware sampling first utilizes computer vision techniques to localize and segment object, then samples the candidates based on the reduced searching space. This method has shown the reliability and reasonability of generated candidates is improved observably, however, adopting handcraft constraints in sampling step may cause generating some unstable grasp poses and computationally expensive. To further address these issues, learning-based object-aware generation modes are proposed by researchers. These methods acquire several grasp candidates with the help of trained model after localizing and segmenting object.

Mousavian et al. (2019) and Murali et al. (2020) employ a sampler ground on variational autoencoders (VAE) (Kingma and Welling, 2013). The sampler's architecture is similar to GANs (Goodfellow et al., 2014), which takes in PointNet++ (Qi et al., 2017b) as encoder and decoder, aiming to generate several grasp candidates and determine how likely they are successful. Yu H. et al. (2019) doesn't localize and segment the object explicitly. The author presents regression network and refine network to regress an optimal grasp region, and sample and sort grasp candidates correspondingly. Zhao B. et al. (2020) uses two neural networks to segment the point cloud and generate grasp proposals correspondingly. Fang H.-S. et al. (2020) presents an end-to-end grasp pose prediction network given N point coordinates as input. Inspired by anchor-based progress in 2D object detection (Ren et al., 2015; Liu et al., 2016), Wu et al. (2020) adopts PointNet++ as backbone to build up a Grasp Proposal Network (GPNet) to generate a set of grasps. The generated proposals are pruned via two physical schemes which are removing grid corners not locate on the object surface and the contact points antipodal constraint (Chen and Burdick, 1993). Li Y. et al. (2020) proposes a Deep Residual U-Nets on the basis of residual modules (He et al., 2016) to predict the graspable region of object, which is followed by a K-means (Lloyd, 1982) model clusters the graspable point cloud and the center of each cluster is leveraged as a grasp point. Ardón et al. (2019) employs Markov logic networks (MLN) (Richardson and Domingos, 2006) for knowing the relationship between diverse objects and a pre-trained Res-Net (He et al., 2016) is utilized to accomplish object perception and feature extraction for querying grasp affordances by Gibbs sampling (Kim and Nelson, 1999). The main thought back Ardón et al. (2019) is sampling several grasp affordances and evaluate them, the affordance with highest possibility will be selected and corresponding grasp configuration is calculated. Inspired by leveraging rectangle represent grasp part (Jiang et al., 2011; Lenz et al., 2015), Vohra et al. (2019) and Yu Q. et al. (2020) sample numerous rectangles to characterize candidate graspable parts and gain the optimal grasp pose by filtering and scoring candidates. After catching sight of high efficiency and easy implementability of pixelwise sampling (Morrison et al., 2018a), Yu Y. et al. (2020) first detects the object via SSD (Liu et al., 2016), the detection results will be checked if the target object is occluded through clustered point cloud from K-means and an image inpainting and recognition network (IRNet) which is inspired by Yu et al. (2018) combined with light-weighted recognition network MobileNet (Howard et al., 2017). The detection and confirmation output are feed into a deep grasping guidance network (DgGNet) to generate and qualify the grasp in each pixel.

As shown in the table, the table compares the learning-based object-aware methods with GPD and GG-CNN, which are representative work of the geometry-based object-agnostic and learning-based object-agnostic sampling method. Although there is no universally applicable benchmark, based on the comparison of the success rate in the same grasp operating environment, it can be seen that the learning-based object-aware sampling methods have a significant improvement in the final grasp success rate in a cluttered environment. At the same time, the grasp success rate will be further improved for the single object.

Benefitting from the progress of several data-driven methodologies, learning-based object-aware models have the highest potential to fulfill grasp proposals selection. However, this type of method also has some drawbacks. One is that the sampling results obtained through prediction usually only consider the pre-shape of the end effector, which probably leads to unavoidable collisions during the motion planning process. The other is for multi-finger end effectors (excluding suction cups and parallel-jaw gripper), it is usually difficult for the model to predict the contact point of each finger, which will make the model perform satisfactorily in a cluttered environment tougher.

Applicable Scene

Generally speaking, these grasp candidate generation methods can be tried on all tasks. But because the ideas of these methods are not the same, for some specific tasks, some methods will theoretically perform better than others.

Geometry Based or Learning Based?

Geometry-based grasp candidate generation is to extract the constraints in the mathematical modeling of the grasp to sample the possible feasible grasp poses in the point cloud. For scenes that are not very cluttered, this method is sufficient to sample a lot of reasonable grasp pose candidates, but if the operating environment is too cluttered, this type of method will be easy to wrongly judge two close objects as an object for sampling. Moreover, geometry-based grasp sampling requires additional conditions for collision detection. Although poor results may not necessarily occur, the efficiency of the algorithm will be greatly reduced.

The learning-based grasp candidate generation is based on a trained model, especially a neural network, which takes point cloud as input, and obtains the sampling result of the grasp pose according to the prediction of its output. Although this method is less interpretable and intuitive than the geomery-based method, the neural network is able to extract richer features in the hidden layer to help sampling, thereby reducing the computational difficulty of collision detection. However, the common problem of the learning-based sampling model is it often requires a lot of data to train a robust model. Collecting data and fabricate a dataset is expensive, which leads to the preparation process of the method time-consuming.

As shown in the **Table 3**, according to some experimental conditions, the recommendation of which grasp generation method to use is listed. This table only makes recommendations for the specific stage of grasp candidate generation. If combined with the evaluation stage, which is detailed in section Grasp Candidate Evaluation, there is no guarantee that geometry-based sampling will perform worse than learning-based sampling.

Object Agnostic or Object Aware?

As far as the current development in the field of robotic grasp is concerned, the object-agnostic sampling method is highly unrecommended. Object-agnostic sampling was proposed in the age when visual detection methods were not effective, but the lack of object information has a great influence on the generation of reliable grasp poses. Therefore, no matter what the task is, it is indispensable to add object information to the algorithm.

In the method of object-aware sampling, it can be subdivided into three branches: object detection and segmentation, object affordance, and shape complement. If it is only for a single object and its shape is regular, then the difference between these three methods will not be too obvious. But in reality, the operating environment of robots is not so ideal and simple. As the earliest developed branch, the method based on object detection and segmentation has a high degree of applicability. Regardless of the operation scenario, the pose information of the object can be extracted through the detection and segmentation of the grasped object, thereby helping the generation of grasping candidates. This kind of method is highly adaptable and can be used as a preliminary attempt in various tasks. Object affordance is to reduce the entire object in the sampling space to a more reliable local area. For simple regular objects, this type of method may not make much sense. However, this method is particularly important for objects that have a large deviation of the center of gravity or damage to the end effector of the robot. The shape complement method can be used as a solution to the poor quality of the input point cloud. The generation of the point cloud depends largely on the light conditions in the experimental scene. Sometimes the point cloud of some objects is very sparse due to poor lighting or the scene is too cluttered. In this case, direct sampling is not advisable. The shape complement is to restore the original shape of the object, so as to improve the information of the object and help the generation of the grasp candidate.

As shown in the **Table 4**, according to the experimental conditions, the recommendation of which grasp generation method to use is listed.

GRASP CANDIDATE EVALUATION

To execute the optimal grasp, a necessary step is to evaluate the generated grasp candidates previously. Evaluation is a latter portion aims to score grasp success probability or classify graspability of grasp proposals. By considering whether the

TABLE 3 | Geometry-based and learning-based recommendation under different conditions.

Condition	Recommendation reason	Recommendation
Single object environment	Easy to sample grasp pose	Geometry-based
Collision-regardless	More constraints required to detect collision	
Hard to generate dataset	No training process based on large-scale dataset	
Cluttered environment	Sample grasp pose based on advanced features	Learning-based
collision-concern	No need to build hand-craft collision detection constraints	
easy to generate dataset	Suitable for training a model	

TABLE 4 | Object-aware sampling branches recommendation.

Condition	Recommendation reason	Recommendation
Preliminary attempts	Highly adaptable	Object detection and segmentation
Regular object	Easy to detect and segment	
Harmless irregular object	Easy to detect and segment, no need to consider unsafe grasp pose	
Regular object	Able to detect more reasonable grasp	Object affordance
Irregular object	Able to detect where to grasp	
Harmful object	Able to detect a safe grasp	
Poor lighting condition	Restore object shape	Shape complement
Sparse point cloud	Restore object shape	
Irregular object	Filter unreasonable grasp pose with symmetric shape assumption	

approach is data-driven or not, evaluation methods can be divided into learning-free and learning-based. Learning-free determine each grasp a good or bad one based on geometry information or control system flow. On the other hand, learning-based attempts to acquire a model to perform as an evaluator trained from datasets.

Learning-Free Candidate Evaluation

Learning-free candidate evaluation determine each grasp a good or bad one mainly based on geometry information. Since approaches in this group don't utilize learned models to carry out the assessment, they put efforts into converting proposals evaluation to an optimization problem. The thoughts behind this transformation is inspired by the optimal control theory in control system or reinforcement learning fields.

Zapata-Impata et al. (2017) presents a function to assess grasp configurations stability by considering the distance, direction and geometric shape of the grasp. On the contrary, Mahler et al. (2016) prefers to rank each grasp candidate based on physical conditions, especially force closure (Weisz and Allen, 2012; Kim et al., 2013; Laskey et al., 2015; Mahler et al., 2015). Following previous works solved the problem via Monte-Carlo integration (Kehoe et al., 2012; Weisz and Allen, 2012) or Multi-Armed Bandits (MAB) (Laskey et al., 2015), Mahler et al. (2016) takes in latter method to find the best grasp. Adopting the thoughts back of reinforcement learning, Manuelli et al. (2019) transforms the evaluation to an optimization problem which is used to find the desired robot action. The constraints of the optimization are established based on geometry, especially the distance of it.

Since learning-free approach normally perform the grasp selection under several assumptions and constraints which are simulated by geometry, causing the lack of flexibility and generality. There are only few works in mainstream take in this evaluation technique. In contrast, learning-based assessment has a wider variety.

Learning-Based Candidate Evaluation

Learning-based are widely used in evaluation step among numerous works. Due to assessing the grasp quality based on trained model from large datasets, learning-based methods are more robust and generalized than learning-free one with prior knowledge instead of complex analysis. Learning-based candidate evaluation pipeline example is provided in **Figure 6**. The learning-based grasp evaluation method is essentially a model that completes the binary classification task. The general pipeline at this stage is to first extract representation of grasp pose and grasp part, then use the learned model to finish evaluation. Whether it is based on SVM in the early days and later based on CNN, the essence of model training has not changed, which is to reduce the classification loss. From hinge loss to cross-entropy or other classification loss, the difficulty at this stage is not how to design the loss function, but how to effectively represent the grasp and use it as the input of the model for evaluation.

At early of the first, Le et al. (2010), Jiang et al. (2011), and Pas and Platt (2015) evaluate the grasp points by utilizing support vector machine (SVM) (Boser et al., 1992). SVM-based approaches are able to classify the suitable grasps with a good result in some simple cases or trivial problems. With the

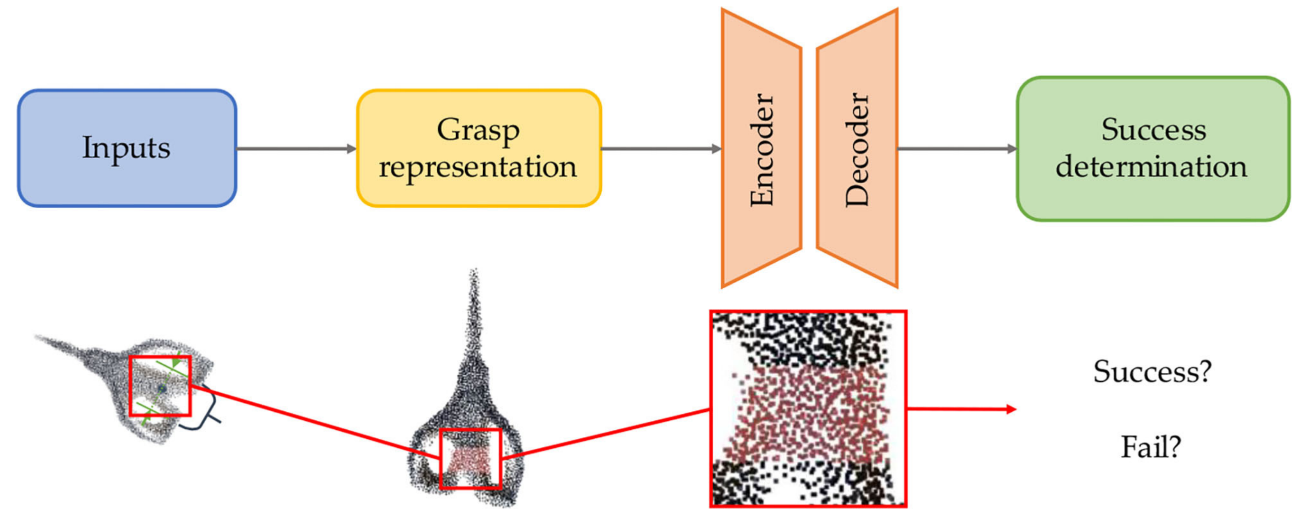


FIGURE 6 | Learning-based candidate evaluation (The scissors point cloud is from YCB datasets).

complexity growth of grasp scenarios, SVM, or other traditional data-driven techniques are decreasingly robust. To ensure the evaluation methods' capability of determining or scoring grasp candidates with noisy inputs, more researchers move their attention to deep learning-based evaluation methods.

Compared with traditional data-driven techniques, deep learning-based evaluation is more precise and robust. Kappler et al. (2016) first indicates the feasibility of evaluating based on CNN. Inspired by this, Gualtieri et al. (2016), Wang and Ling (2016), and Pas et al. (2017) use LeNet (LeCun et al., 1998) to classify the grasp proposals and achieve an impressive performance. Even LeNet architecture is shallow and are not robust to noisy grasp proposals, CNN-based classifier's potentials of evaluating grasps attracts large amounts of interests in robotic community. Following prior arts (Lenz et al., 2015; Gualtieri et al., 2016; Mahler et al., 2016), Mahler et al. (2017) and Mahler et al. (2018) propose a Grasp Quality Convolutional Neural Network (GQ-CNN) to evaluate grasp and suction task respectively. Depending on flexibility of modification, comprehensibility of architecture and simplicity of implementation, GQ-CNN becomes a wide preference among several works. Jaśkowski et al. (2018) utilizes a new CNN architecture and add batch normalization (Ioffe and Szegedy, 2015) to refine GQ-CNN. Mahler and Goldberg (2017) models bin picking on the basis of Partially Observable Markov Decision Process (POMDP) (Astrom, 1965) and fine-tunes GQ-CNN with a new dataset to evaluate actions instead of grasp configurations, which improves the generalization of GQ-CNN. Satish et al. (2019) further enhances GQ-CNN by designing a FC-GQ-CNN through fully convolutional network. Fang K. et al. (2020) proposes a Task-Oriented Grasping Network (TOG-Net) by making progress on GQ-CNN via residual network layers (He et al., 2016) to obtain task-agnostic grasp quality, conditioned task-oriented grasp quality and manipulation action. Although GQ-CNN achieve an impressive performance, it currently requests high-quality depth sensors to obtain desirable point cloud, which limits the deployment in many cases.

Except from GQ-CNN, other works also propose some novel evaluators to figure out grasp candidate assessment. Following 3D CNN predictor designed in Choi et al. (2018), Lou et al. (2020) passes voxelized point cloud of each grasp candidate into networks and fortify a reachability predictor to strengthen selected grasp robustness. Inspired by the work in Varley et al. (2015) and Lu et al. (2020) utilizes a patches-CNN to gain the information from different patches in images to calculate a suitable grasp. Van der Merwe et al. (2020) takes a further step by way of signed distance functions (SDF) to earn object reconstruction. The extracted point cloud embedding is set as the input to the success probability predictor (Lu et al., 2020) extended by collision-free strategy (Zhou and Hauser, 2017; Lu and Hermans, 2019; Lu et al., 2020). Qian et al. (2020) modifies fully connected layer by a novel pooling layer in R-FCN (Dai et al., 2016) which is able to perceive object localization change precisely. Yu Q. et al. (2020) classifies grasp rectangles via a 7-layer CNN. Although these variants of 2D CNN achieve impressive performances, a common drawback is it's hard for them to handle imperfect observed point cloud and extract sufficient and stable information.

Profiting from studies on point cloud (Rusu and Cousins, 2011; Guo Y. et al., 2020; Lu and Shi, 2020), PointNet and PointNet++ are two extraordinary and widespread models which make a further promotion on evaluation networks. Liang et al. (2019), Mousavian et al. (2019), and Yan et al. (2019) use PointNet to directly take point cloud as input and output the grasp candidate evaluation. The points within the closing area of the gripper are utilized to represent the grasp. Immediately transform grasp to the points in corresponding area addresses unstable prediction results from imperfect local observed point cloud and carries out an acceleration of evaluation process. Stimulated by PointNetGPD (Liang et al., 2019), Singh et al. (2018) tries to replace PointNet with PointNet++, KD-Networks (Klokov and Lempitsky, 2017) and Dynamic Graph Convolutional Neural Networks (DGCNN) (Wang Y. et al., 2019) to obtain a better result. By considering parallel structure utilized in some works ignores grasp candidate generation errors result in unreliable evaluation, Grasp Proposal Networks (GPNet) (Wu et al., 2020) adopts GraspNet (Mousavian et al., 2019) and designs a structure which is able to allow generator and evaluator to be trained jointly. Fang H.-S. et al. (2020) utilizes PointNet++ as their ApproachNet to obtain suitable grasps. In contrast with traditional 2D-input CNN, methods based on PointNet or PointNet++ are capable to handle noisy inputs and assess grasp candidate probability and stability with higher confidence.

Approaches presented by other arts are not in line with CNN-based methods but still achieve good performance. Enlightened by wide & deep model in recommender system (Cheng et al., 2016), Context-Aware Grasping Engine (CAGE) (Liu et al., 2020) treats discovering a feasible grasp configuration as a recommendation problem. Based on the prior knowledge, the model predicts suitable grasp by finding grasp configuration in the similar situation. On the contrary, Wu et al. (2019) utilizes a reinforcement learning pipeline, which assesses and refines the action taken in each time stamp based on reward function via policy gradient (Sutton et al., 1999).

END-TO-END AND OTHERS

Except from those works can be categorized into grasp generation or evaluation part, there still has numerous arts out of this framework. By adopting the advantages of end-to-end learning, some researchers attempt to concatenate two parts to enable training the network jointly. Instead of using supervised learning, some works empower grasping ability of robots through interacting with environments on the basis of reinforcement learning. Other proposed models are based on approaches not in mainstream.

End-to-End Learning

With the development of deep learning, end-to-end learning gradually becomes one of the most outstanding learning mode, which is seemingly natural consequence of deep neural architectures blurring the classic boundaries between learning machine and other processing components by casting a possibly complex processing pipeline into the coherent and flexible modeling language of neural networks (Glasmachers, 2017). Deep learning based on end-to-end mode is capable of training and generating a more powerful model via a holistic object function.

S^4G (Qin et al., 2020) proposes a single-shot grasp proposal network based on PointNet++ which assigns each point in point cloud a grasp configuration and its quality score. Non-maximum suppression (NMS) and weighted random sampling are applied to the output to select a grasp to be executed. Based on the network in (Choi et al., 2018), Liu M. et al. (2019) improves its performance through introducing a new combined loss which is composed of consistency loss and collision loss. These two losses aim to resolve grasp pose ambiguity and penalizes the penetrations respectively. Yu Y. et al. (2019) preprocesses the input via utilizing FPN with ResNet50 and K-means by taking RGB images and point cloud as inputs correspondingly to extract the multi-scale masks of target object. Then a DrGNet takes masks as inputs to perform depthwise separable convolution. The encoded results from DrGNet are refined by RefineNet (Nekrasov et al., 2018) and sSE (Roy et al., 2018) to obtain a desirable grasp. Provided by object mask, grayscale and depth images as input, Tosun et al. (2020) trains grasp proposal network (GPNet) (Tosun et al., 2019) and shape reconstruction network (SRNet) (Mitchell et al., 2019) parallelly to acquire grasp proposal and reconstructed point cloud. The embeddings from GPNet and SRNet are combined to refine the detected grasp. PointNetRGPE (Wang Z. et al., 2020) first predicts the corresponded class number from object point cloud data, which is used to fuse with point coordinates to pass into grasping pose estimation network. The network has three sub-networks based on PointNet to acquire the translation, rotation and rotation sign of grasp pose. Other than generate the grasp configuration parameters, GraspCNN (Xu et al., 2019) treats the grasp estimation as an object detection problem. It takes RGB images as input and outputs an oriented diameter circle. The circle and oriented diameter indicate the grasp area and gripper open width and closing orientation respectively. Obtained circle on RGB image is calculated to project into the point cloud.

Instead of designing end-to-end framework based on supervised learning, some arts propose their end-to-end models through the methodologies of reinforcement learning. Zeng et al. (2018a) attempts to jointly train two FCNs in Q-learning framework to extract the visual state representations and obtain the push and grasp from the policy. Compared with the demands of a large amounts of manually labeled data in supervised learning, approach presented in Zeng et al. (2018a) is totally self-supervised. Wu et al. (2019) takes depth images as input and obtain 10 2D maps after flowing in an FPN. Ten maps represent current state of end-effector and objects and are utilized to sample the action. Policy is learned based on policy gradient during the action execution. Manuelli et al. (2019) segments and detects the keypoints of object via Mask R-CNN and pose estimation network (Sun et al., 2018). The grasp planner (Gualtieri et al., 2016; Zeng et al., 2018b; Mahler et al., 2019) is selected to generate a pose and refined by proposed optimization method.

End-to-end models are capable of learning a complete function maps from visual inputs to grasp poses. Although there is not much work based on an end-to-end approach, this direction has great potential for development. Since it takes the point cloud as input, and then directly outputs the grasp pose, this omits the consideration of the features extracted from the connection in the multi-step method. At the same time, the efficiency of the end-to-end algorithm is usually better than that of the multi-stage model, which makes it possible for the robot to operate in real time in a dynamic environment.

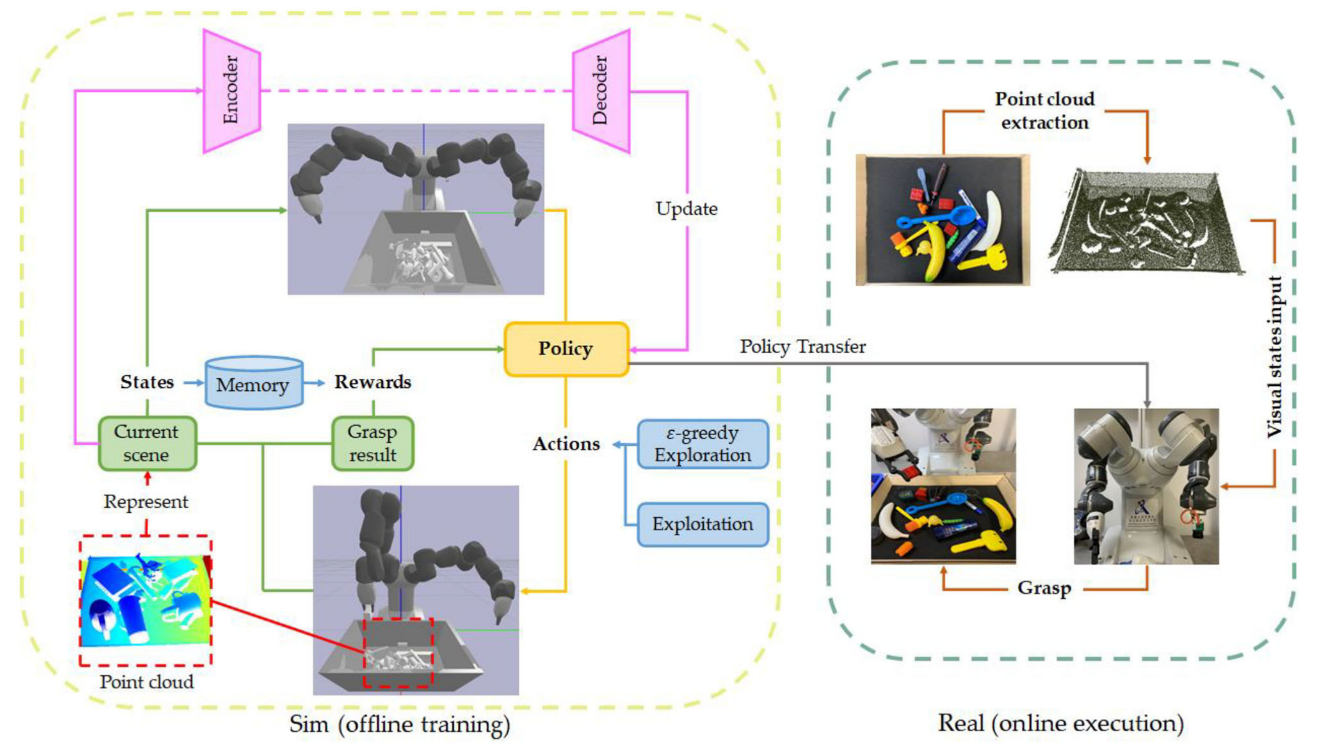


FIGURE 7 | The deep reinforcement learning framework of robotic grasp learning based on point cloud. In order to reduce the cost of trial and error, the current robot grasping based on reinforcement learning is to first train the model in a simulation environment and then migrate to the real robot (Sim to real).

Reinforcement Learning

Reinforcement learning is another interesting solution of grasp pose detection. Although the work on reinforcement learning has been mentioned in the end-to-end method, what mentioned earlier is end-to-end learning based on the reinforcement learning framework. Generally speaking, reinforcement learning methods do not necessarily adopt end-to-end thinking, and vice versa. As shown in **Figure 7**, RL approaches enable robot to interact with environments to study the policy maps from visual inputs to actions. In contrast with supervised learning, trial and error thoughts back of RL capacitate robots to learn self-exploration ability, which makes robots have higher dexterity. The reason why reinforcement learning makes robots more dexterous than supervised learning is because the training ideas of the two are completely different. Supervised learning is to update the model parameters through sample and label pairs until the loss function is minimized. But the biggest drawback of this training method is that its dataset usually has only one label per sample. That is, for a grasp point, it can only correspond to one grasp pose, which actually adds a lot of restrictions to the model. Image that a robot wants to grasp a block or a ball, the same scene can have hundreds of thousands of grasp poses for humans because the deviation of the hand and joint angle has no effect on the successful grasp of such objects. However, there is only one correct answer in supervised dataset. On the contrary, reinforcement learning is to obtain the policy of grasping objects through trial and error. Since the robot may have explored many grasping possibilities during training, and the reward function value of each grasp pose is not bad, the algorithm will add these possibilities to the policy function, so that the possible answers for the operation become more.

Ficuciello et al. (2019) first take point cloud as input, an object recognition module is utilized to accomplish object detection and pose estimation. Acquired features are feed into a pretrained neural network to obtain the robot grasp initial configuration. Then a RL loop is used to refine the initialized parameters by assigning the executed grasp a cost to update the policy. Inspired by Deep Q-learning network (DQN) (Mnih et al., 2015), Gualtieri and Platt (2018) trains a CNN to learn Q-function and utilizes gradient Monte Carlo (Sutton and Barto, 2018) to update the rule. At each time stamp, it generate several grasp candidates based on hierarchical sampling and then one pose will be chosen by the learned policy. Chen et al. (2020) adopts RL to obtain an appropriate viewpoint based on the mask-guided award to perform GPD module. Rather than deploying RL strategy on grasp planner, this model focuses on acquire a better view sight to improve the grasp accuracy.

RL-based approaches empower robot ability of self-exploration from trial and error, successfully trained models are capable of planning grasping dexterously. However, methods based on reinforcement learning, especially when used for robot operations, will have a serious problem – the exploration space of the algorithm will become extremely large, or the grasp poses that require trial and error are innumerable. This will make the learning efficiency become very slow, and because the exploration space is too large, the positive and negative samples will be extremely imbalanced. The robot may not be able to increase the reward function value after many trial and error, and the policy cannot be updated. Besides, due to the necessity of a mass of sampling during training, the data collection process is also time-consuming. In addition, it still needs more time to transfer the model in simulation to real world since the cost of each trial may be extremely high.

Others

Aside from end-to-end learning and reinforcement learning, some other works achieve the goal through some unique and creative methods.

Zhu et al. (2019) first obtains the graspable area via performing ellipse fitting method on segmented mask from Mask R-CNN. Then RANSAC is utilized to acquire the orientation of the grasp from the pixel mask and point cloud. Instead of generating the grasp pose parameters, Shao et al. (2020) attempts to predict the grasp contact points. It not only extracts the feature of object point cloud, but also pays attention to gripper properties. An unsupervised autoencoder adopts the structure of PointNet to learn a low-dimensional latent space of gripper representation and construct robotic hand representation from URDF file. Then the gripper representation and object point cloud features extracted by PointNet++ are combined to feed into a proposed Point Set Selection Network (PSSN) to generate correspond number of contact points based on beam search. Kokic et al. (2017) employs two CNNs, one for affordance detection and another one for classification and orientation estimation. Extracted parameters are used to compute a grasp by Haustein et al. (2017).

Approaches proposed in this group can achieve an impressive performance in some special cases, however, generalization of these methods are not approving.

LEARNING MODES

There are numerous types of machine learning algorithms, which aim to solve problems in different situations or under diverse demands. In the robot grasp learning field, learning modes can be mainly divided into supervised learning, unsupervised learning and reinforcement learning even some arts may be outside of these three categories. Nowadays, due to the challenge of grasp pose detection with high dexterity, it facilitates the fusion of different learning modes to integrate the merits of each other to improve the model performance. The descriptions of three main modes are provided first, and then the fusion models are discussed based on three primitive modes.

Supervised Learning

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs (Russell and Norvig, 2002). The most prominent property of supervised learning is the datasets used to train the models have the labels for each sample. Supervised learning models are the most widely used since they are simple to implement, easy to train and suitable for most tasks. In robot grasp learning task, supervised learning methods can be used for candidate either generation or evaluation. Many end-to-end learning approaches are also based on the thoughts back of supervised learning.

Initially, supervised learning models without deep learning techniques are only used for grasp candidate evaluation since extensive demands of handcraft feature preprocessing cause low performance of candidate generation. Pas and Platt (2015) uses SVM to assess the quality of grasp proposals and acquires an incredible performance. With the development of deep learning, supervised learning models based on neural networks outperform and replace traditional techniques in many tasks. Gualtieri et al. (2016), Wang and Ling (2016), and Pas et al. (2017) try to evaluate the grasp pose candidates based on LeNet for the first time. Astonished by the power of neural network, deep learning models attract dramatically substantial interests from researchers. Mahler and Goldberg (2017) and Mahler et al. (2017, 2018) propose CNN-based evaluators with more complex architecture.

Inspired by several state-of-arts works in computer vision (Ren et al., 2015; He et al., 2016, 2017; Liu et al., 2016; Redmon et al., 2016), researchers begin to move the attention to grasp candidate generation. Some works attempt to use object detection and instance segmentation approaches to reduce the search space in the point cloud (Lopes et al., 2018; Schnaubelt et al., 2019; Bui et al., 2020) or regress the grasp part using rectangle bounding boxes directly (Vohra et al., 2019; Yu Q. et al., 2020). Specifically, models designed to process point cloud furtherly accelerate the progress of candidate generation (Lin and Cong, 2019; Yu S. et al., 2020) and evaluation (Singh et al., 2018; Mousavian et al., 2019; Yan et al., 2019; Fang H.-S. et al., 2020; Qian et al., 2020). Moreover, end-to-end learning models mentioned in section End-to-End Learning follow the supervised learning framework as well.

By adopting supervised learning methods, especially those in deep learning field, feature extraction of robot grasp learning has transformed from handcraft to learning-based. Supervised learning methods is capable of learning the models to accomplish specific tasks with only needs of regarding datasets. However, training an efficient supervised learning model requires a large amount of data. Data collection is very expensive and time-consuming in many cases which cause the model not able to learn enough knowledge to perform as expect, which is known as underfitting. In addition, incomplete training set also probably result in overfitting of models due to samples cannot represent the entire rules to be learned. For end-to-end learning, incorrect-design architecture has higher potentials to be overfitting. Besides, generalization of grasp learning is still a challenge. Model trained on one case is commonly hard to be transferred to other cases.

Unsupervised Learning

In contrast with supervised learning, unsupervised learning takes unlabeled data as input aims to find the internal relationship of samples which allows for modeling of probability densities over inputs (Hinton et al., 1999). Two of the main techniques in unsupervised learning are dimensionality reduction and clustering. Clustering is used to group or segment the datasets to numerous clusters, which is adequate for processing the point cloud. Some works adopt K-means as clustering method to segment the point cloud as their desirable features for downstream.

Different from classical unsupervised learning approaches, VAE and GAN are the product of the growth of deep learning. Both VAE and GAN are generative models, which are to model the real data distribution from the training data, and then use the learned model and distribution to generate and model new data in turn. They are similar in two respects. One is that random noise is used in the data generation mode (such as Gaussian distribution is commonly used), and the other is that when modeling the distribution, it is necessary to measure the difference between the distribution of noise and training data. The difference between the two is essentially that the distribution measurement criteria are different (that is, the loss is different). VAE uses a more explicit measurement method, assuming that the training data is generated by another distribution, and directly measures the KL divergence of the training data and noise. From this, the ELBO theory, reparameterization trick and so on have been developed. The GAN cleverly avoids the direct measurement of the distribution difference, but lets the neural network learn this distance through confrontation. When the discriminator cannot distinguish between the two distributions, it is considered that the two distributions are consistent. The emergence of VAE and GAN makes neural networks also usable in tasks based on random sampling. Some works have adopted this strategy in the stage of grasp candidate generation. Inspired by GAN, Mousavian et al. (2019) and Murali et al. (2020) utilize VAE to sample multiple grasp proposals to speed up candidate generation.

Robot grasp learning is hard to be accomplished only based on unsupervised learning methods since study on unlabeled data is not trivial to acquire sufficient knowledge to drive robots perform dexterous grasp poses as human. However, unsupervised learning can efficiently segment point cloud or sample grasp candidates with satisfying results even using fairly simple models.

Reinforcement Learning

Differs from demands of labeled and unlabeled data for supervised and unsupervised learning respectively, reinforcement learning (RL) aims to learn a policy maps from agent's states to actions and maximizes reward by interacting with environment. The core problem of RL is to find a balance between exploration and exploitation. The motivation of RL is letting the agent cumulate knowledge by trial and error. As mentioned in section Reinforcement Learning, there are few RL works achieve robot grasp goal based on point cloud and deep learning. The reason is the difficulty of transferring point cloud to visual state representation during robot exploration.

Fusion

More recently, increasingly number of researchers begin to fuse different modes to adopt and integrate their advantages. Fused models enable employing the most appropriate learning types in each part of architecture, which allows a great promotion of grasp learning. It is very tough to model and accomplish a task only based on single learning modes. In addition, sometimes single mode perhaps complicates the problem and reduces the solving efficiency.

As mentioned in section Unsupervised Learning, it is hard to design a pipeline only based on unsupervised learning. K-means (Yu Y. et al., 2019, 2020; Li Y. et al., 2020) or GAN-analogous (Mousavian et al., 2019; Murali et al., 2020) methods only serves one functional part in entire architecture. On the other hand, like the discussion in section Reinforcement Learning, it is significant to acquire valid state representation to obtain a powerful reinforcement learning model. By integrating the visual perception, it is intuitive to utilize the input RGB images or point cloud to extract state information. Object localization and segmentation are almost achieved by numerous supervised learning-based CNN. Therefore, reinforcement learning techniques are naturally fused with supervised learning. However, large-scale datasets are commonly demanded in supervised learning to obtain a high-performance model, which request time-consuming data collection and annotation.

Self-supervised learning is a new mode in machine learning. Instead of paying expensive cost of gathering and labeling data, as a subset of unsupervised learning, self-supervised learning aims to study the labels represented by features based on provided unlabeled data. In robot learning, it's pretty costly to collect datasets and annotate each sample manually. In addition, defining, labeling and representing ground-truth for each sample may be challenging and wrong annotation can cause the model hard to train. Self-supervised learning is able to extract the feature-based labels during studying to avoid inaccuracy of handcraft. Zeng et al. (2018a), Deng et al. (2020) and Fang K. et al. (2020) adopt self-supervised framework to learn the model with a few data to pretrain or without any data to train from scratch. The results are surprisingly good and exhibit the great potential to employ self-supervised learning into robot learning research. Moreover, since training process is on the basis of unannotated data, the model performance can become cumulatively better with growth of learning time. Self-supervision is looked forward to promoting the continual learning of robotics.

APPLICATIONS

There are numerous applications of robot grasping based on point cloud and deep learning. Generally, as shown in **Figure 8**, these methods can be categorized from two different concepts. From the perspective of end effector, approaches can be divided into suction cup, two fingers (parallel-jaw gripper) and multi fingers. In terms of operating scenarios, approaches fall into life-oriented and industry-oriented.

End Effector

Due to grasping mechanisms and DOFs vary from single suction cup (one finger) to five-finger hand, end effector plays an important role in grasp learning algorithm design. With more fingers assembled on end effector, the dexterity of

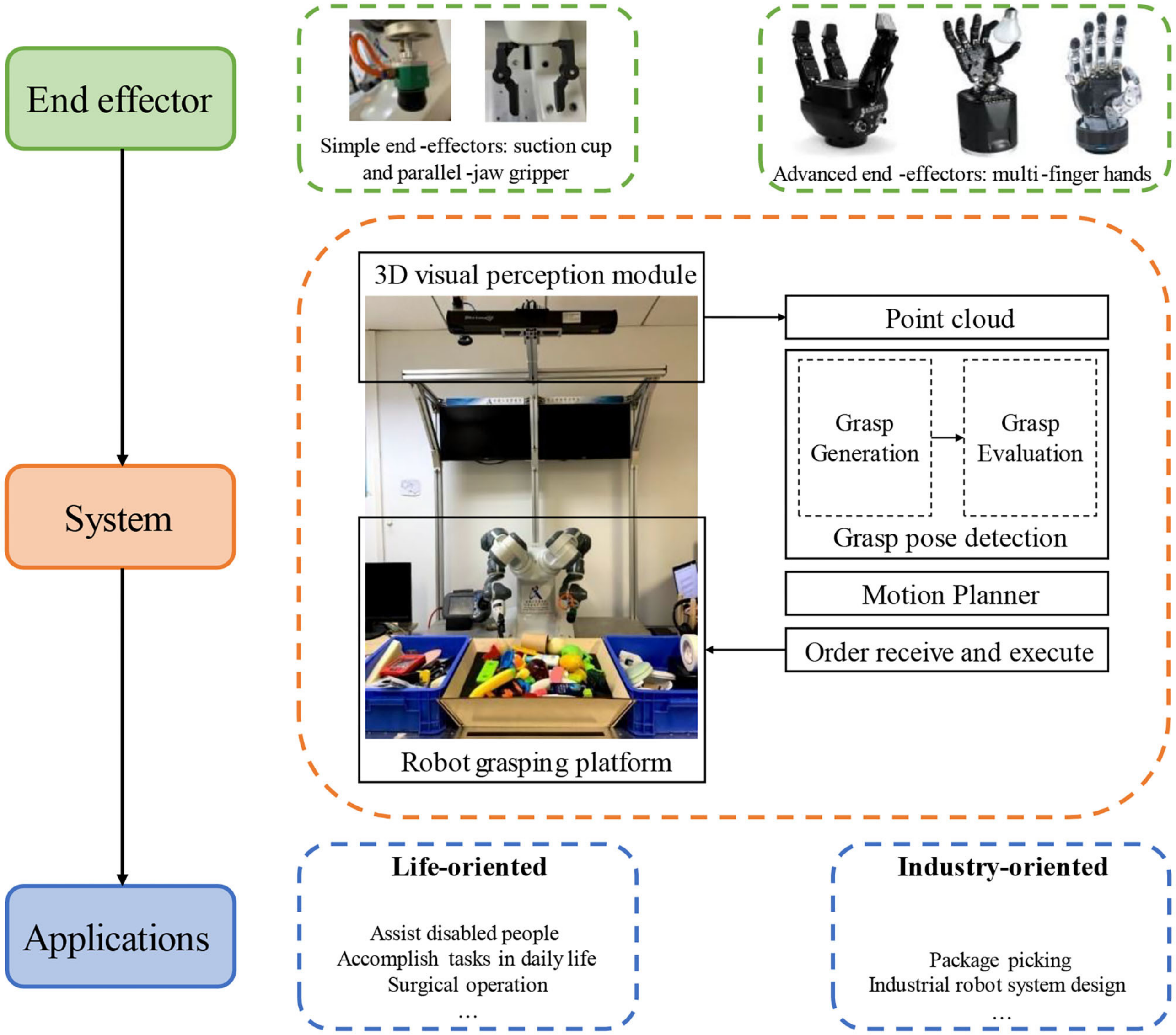


FIGURE 8 | End effector, system and applications for robotics dexterous grasping. End effector is divided into simple end-effectors and advanced end-effectors. The former group contains suction cup and parallel-jaw gripper, the latter class indicates those multi-finger hands. Grasping system is first designed, deployed, and matured on simple end-effectors, then transferred and improved on advanced end-effectors. Developed systems are applied in different scenarios, life-oriented, or industry-oriented.

grasp increases dramatically to allow robot to accomplish more complicated tasks.

Suction Cup

The suction manipulation based on vacuum cup has an unparalleled advantage over other multi-finger operation, which just request detect one feasible contact point to perform object picking. Single contact operation is not only convenient, and there is no need to worry too much about the collision between the end effector and other objects, especially in cluttered environment. Jiang et al. (2020), Mahler et al. (2018), and Mahler et al. (2019) pay efforts to pick the object through suction and obtain 98, 95, and 97.5% success rate respectively.

Parallel-Jaw Gripper

Robots equipped with parallel-jaw gripper are more favored in research and real life than humans grasp an object using movable joints index finger and thumb. Compared with suction manipulation, parallel-jaw gripper is able to perform more dexterous operations even if it demands more consideration of collision. The main reasons are grippers are easy to model in physical simulation environment and sample the grasp poses in point cloud. By knowing the grasp configuration and location in operating space, it is straightforward to calculate the contact points via geometrically symmetric shape to detect the collision. Pas and Platt (2015), Gualtieri et al. (2016), Mahler et al. (2017), Pas et al. (2017), and Liang et al. (2019) are fairly representative arts in the field of two-finger grasping.

Multi Fingers

There are merely few works focus on multi-finger grasp learning, not only is it difficult to model in the simulation environment, but also because of its overhigh DOFs and contact points make the sampling-based candidate generation strategy more sophisticated. Despite this, there are still some works focus on multi-finger grasping have achieved satisfactory results. Guan et al. (2019), Lin and Cong (2019), Liu C. et al. (2019), Wu et al. (2019), Shao et al. (2020), and Yu Y. et al. (2019), Yu Y. et al. (2020) utilize three-finger hand as their end effectors and Ficuciello et al. (2019) and Yu Q. et al. (2020) adopt five-finger to design grasp learning algorithm.

Operating Scenarios

To categorize approaches in terms of operating scenarios, they can generally be divided into life-oriented and industry-oriented groups. Life-oriented methods aim to design a robot system can serve or help people to accomplish tasks in their daily life. Industry-oriented methods are designed to complete those dangerous and arduous tasks in industry.

Life-Oriented

Grasp is one of the most primitive and core manipulations of robot, many advanced operations are variant of grasping. With the growth of robot grasp learning, more and more tasks can be finished by robot in people's daily life, which improves quality of people's life. Gualtieri et al. (2017) and Zhang and Demiris (2020) propose robot systems to assist disabled people grab objects and dress cloths. Llopart et al. (2017), Zhou et al. (2019), Yang et al. (2020) and Zeng et al. (2020) aim to learn grasping capability to accomplish opening doors, grabbing glasses, picking objects from human's hands and throwing arbitrary objects. More interestingly, Parhar et al. (2018), Guo N. et al. (2020), and Kang et al. (2020) utilizes robot grasp ability to help completing crops harvesting in the farm. Zhang et al. (2020) enables robot grasping to be controlled by a smartphone. Hu et al. (2019) creatively integrates robotic grasp with surgery, which may help doctors during surgical operation. Nishikawa et al. (2019) and Schnaubelt et al. (2019) let the robot learn the knowledge to aid rescuers clean up disaster scene.

Industry-Oriented

Some works are aimed at industrial scenarios and propose algorithms for robot grasping learning to further promote productivity. Tian et al. (2017), Antonova et al. (2018), Li et al. (2018), Li H. et al. (2020), Song et al. (2019), Bui et al. (2020), and Liu et al. (2020) present some system design idea based on robot grasping that can be used in industrial production. Amazon Picking Challenge (APC) is a competition to provide a challenge problem to robotics community (Wurman and Romano, 2015) that spawns numerous excellent works based on the combination of point cloud and deep learning. Hernandez et al. (2016), Zeng et al. (2017), Morrison et al. (2018b), and Matsumoto et al. (2020) are some typical picking systems designed in top teams.

CHALLENGES AND FUTURE DIRECTIONS

Robot grasp learning based on 3D point cloud provides more potentials to estimate and execute more precise grasp pose on the target objects. Deep learning growth also brings more promising strategies on grasping system design. Applications mentioned in section Applications exhibit the great probability of dexterous-grasping ability can be closely integrated with tasks in a variety of industries. However, due to sparse and unstructured properties of point cloud, training difficulty and low generalization of deep learning and ambiguity of dexterous grasping definition, the challenges of robot grasp learning are provided. Some future directions are also discussed based on these unresolved problems. Challenges and directions can mainly be divided into three groups, improving perception and manipulation abilities of robots, promoting the intelligence of robots and enabling abilities of transferring.

Easy-Vision and Complicated-Dexterity

To achieve successful grasp goal, proposed methods heretofore commonly utilize multi viewpoints input and simplified grasp definition. Suppose deploying floor mopping robots, it is difficult to install a camera in each room. Instead, the camera will be equipped on robot which is able to barely provide partial observation of many objects. Hence, easy vision becomes increasingly useful. "Easy" means robots only take partial observation of objects from single viewpoint to detect the grasp pose. Concerning the grasp definition, lots of works are on ground of multiple hypothesizes and posture restrictions, which are not the true sense of 6-DOF grasping. Moreover, as mentioned in section End Effector, most arts focus on grasp learning with parallel-jaw gripper, which cannot be said as genuine dexterous grasping.

Single-View Grasping

Considering training a formidable grasping models, numerous works capacitate robot's visual perception ability by offering visual training data collected from multi-viewpoints. In spite of this can achieve a high grasping accuracy in special cases, these approaches have an important assumption which is robot has a complete knowledge of the environments to perform grasp poses. As a result of impossibility and hardness of providing entire object observation in diverse task scenarios, single-view grasp learning plays an increasingly important role. Shape complementation (Watkins-Valls et al., 2019; Van der Merwe et al., 2020) or taking partial observation as input immediately (Yan et al., 2019; Qian et al., 2020; Qin et al., 2020) are tried by researchers. Even sophisticating the model design and training, single-view grasp learning enhances algorithm robustness and reduces data collection cost.

6-DOF Grasping

To accomplish grasping task in a simple way, many works stipulate the end effector can only grasp objects along the axis perpendicular to workspace. However, these restrictions make performed grasps lose the so-called 6-DOF. In addition, these analogous 6-DOF grasps are hard to achieve a satisfying accuracy in cluttered environment. 6-DOF grasping should be flexible that is capable of approaching the detected grasp point from any directions. With the progress of grasping learning models and robot hardware, the problem of quasi-6-DOF grasping has been gradually improved (Gualtieri and Platt, 2018; Lin and Cong, 2019; Mousavian et al., 2019; Lou et al., 2020; Murali et al., 2020).

Multi-Finger Grasping

The ultimate goal of robot grasp learning is to give the robot an anthropomorphic grasping ability, and it is a particularly important step to evolve the end effector from parallel-jaw gripper to five fingers. Thanks to the simple characteristics of modeling and physical analysis of parallel-jaw gripper, many works have achieved incredible results on the problem of two-finger grasping. However, as for multi-finger robots, especially five-finger robots, effective work is still lacking (Ficuciello et al., 2019; Yu Q. et al., 2020). High hands DOF and computational complexity caused by multi-contact points and multi-joints remain to the difficult enhancement of five-finger grasping.

Fusion, Self-Exploration, and Continual Learning

Promoting the intelligence of robots is an appealing and core field in the future directions. Learning algorithms give robot intelligence, and the quality of the algorithms is the most prominent dependent factor for accurately and dexterously grasping objects. In order to design more complementary algorithms, fusion has become the current general trend. Fusion not only includes the integration of multiple learning modes, but also the integration between multiple modalities. Furthermore, enabling robots to have the ability of self-exploration and continual learning with a few prior knowledges has also become a hot Research Topic.

Fusion

As mentioned in section Fusion, the fusion of learning modes adopts the most appropriate learning modes at each part of the method architecture to improve model performance and reduce learning costs. In the contrast with multi-mode fusion, multimodal fusion not only relies on point cloud itself to improve the ability of grasp learning, it also utilizes language or tactile sense to enrich the features extracted in the learning process so that the robot has more grasping knowledge (Sung et al., 2017; Zhou Y. et al., 2018; Abi-Farraj et al., 2019; Kumar et al., 2019; Ottenhaus et al., 2019; Wang T. et al., 2019; Watkins-Valls et al., 2019).

Self-Exploration

Self-exploration refers to the method in which the robot learns to grasp through interacting with environments. Self-exploration breaks away from traditional approaches of relying on supervision but turns the problem into learning a policy maps from states to actions via trial and error. Reinforcement learning is currently the most powerful tool for self-exploration. As mentioned in section Reinforcement Learning, some works have been carried out around RL and achieved impressive results. Nevertheless, because the reward function is difficult to design, the deep reinforcement learning training requires a large-scale dataset support and the generalization performance is poor, there is still a lot of room for development of robot self-exploration based on RL.

Continual Learning

Continual learning (CL) is a machine learning paradigm where the data distribution and learning objective change through time, or where all the training data and objective criteria are never available at once (Hadsell et al., 2020; Lesort et al., 2020). Even for a grasping task, it is currently tough to grasp different types of objects in different scenes by a single robot. Continual learning allows the robot to learn new manipulation knowledge while not forgetting what has learned before. This ability makes it possible for a robot to complete multiple tasks. However, due to the demands for a large amount of memory to store the learned knowledge, and to adapt to the changes in the data distribution at any time, there is not much work in related fields as present. Moreover, continual learning and self-exploration complement each other. As the number of trial and error increases, robots will learn more and more knowledge. How to ensure that new knowledge doesn't overlap, conflict and cover the previous knowledge will require continual learning strategy.

Sim-to-Real and Generalization

Designing, training, and deploying a grasp learning model is often time-consuming and laborious. Researchers expect to apply transfer learning strategies (Pan and Yang, 2009) to the learned model, so that similar grasping tasks don't need to be restarted from scratch. At present, there are three crucial problems in robot learning have not been well-solved. First is the problem of sample efficiency. Because the data for training robots is difficult to collect and a lot of collected data will not be helpful to the promotion of model learning due to unrepresentative. The other is the generalization ability of the model. Since the robot's end effectors are diverse and grasp learning is for specific tasks, the generalization ability is mostly poor. Finally, sim-to-real is also a challenge. Unlike traditional artificial intelligence tasks, robot learning commonly involves training and testing algorithms in a physical simulation environment and then transplanting to real robots. However, because the simulation environments are based on the ideal physical situation, it is quite different from the reality. Moreover, the migration of visual algorithms obtained in the simulation environments to reality will have cross-domain problems, which will cause the performance of the algorithms drop significantly after migration.

Sample Efficiency

In robot learning, especially those methods based on reinforcement learning, sample efficiency is particularly significant due to the particularity of the data. The low sample efficiency will not only increase the cost of data collecting and model training, but also cause the model to easily underfit. For the off-policy reinforcement learning approaches, since the policy is improved based on training samples, models with high sample efficiency can quickly learn appropriate policy from the data (Gualtieri and Platt, 2018; Zeng et al., 2018a; Zeng, 2019).

Model Generalization and Transfer

The methods mentioned in sections Grasping Candidate Generation and Grasp Candidate Evaluation based on CNN to generate and evaluate grasp candidates show the potentials of generalization (Mahler et al., 2017, 2018; Yan et al., 2019; Chen et al., 2020; Shao et al., 2020). However, for the proposed approaches based on reinforcement learning are hard to be transferred to other tasks. Even if there are already some works on the integration of transfer learning and reinforcement learning (Tirinzoni et al., 2018; Ammanabrolu and Riedl, 2019; Gamrian and Goldberg, 2019; Liu Y. et al., 2019; Xu and Topcu, 2019), the work related to the robot grasping is lacking. The essence of transfer learning is the registration problem at the task level, including not only the task itself, but also its input and output. If robot grasping can be defined from a higher level, if will make transfer learning possible.

Sim-to-Real

Simulation-based training provides data at low-cost, but involves inherent mismatches with real-world settings (Zhao W. et al., 2020). At present, domain randomization and domain adaptation are widely used in sim-to-real problems. In order to not degrading the performance, these methods attempt to make the data distribution from the simulation environments and real-world environments more similar to each other. Nonetheless, the low interpretability of domain randomization approaches and non-real hypothesis of domain adaption still make the sim-to-real hard to solve. Imitation learning, meta-learning and knowledge distillation are also supposed to have probability to facilitate the solution, however, more time is requested to achieve the goals.

CONCLUSION

The current researches on robot dexterous grasp learning based on point cloud and deep learning can be divided into grasp candidate generation and grasp candidate evaluation. On the basis of this effective and reliable two-stage algorithm model, this survey proposes a more generalized learning framework. Most of the work can be summarized as a substantial contribution to one of these two stages. For work that does not belong to this framework, aside from the end-to-end model, the most prominent part is reinforcement learning framework. Although reinforcement learning is not summarized in the main framework of this review, sampling grasping actions from policy and assessing grasping actions from reward function can actually be regarded as the idea of proposed framework. But in order to respect the original motivation of the authors, RL-based approaches are not categorized. This survey aims to provide valuable insights and inspiration ground of sufficient bibliographical contents. Although there are still numerous challenges and limitations, methods with point cloud and deep learning have proven their potentials in promoting the improvement of robot dexterous grasping.

AUTHOR CONTRIBUTIONS

HD and PW brought up the core concept and architecture of this manuscript. HD and YH collected the majority of references. HD wrote the paper. GX and WW processed the objects point clouds of figures in the paper. XS refines some details of expression in section Object-Aware Sampling. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the Strategic Priority Research Program of Chinese Academy of Science under Grant XDB32050100 and in part by the National Natural Science Foundation of China under Grants 91748131, 62006229, and 61771471.

REFERENCES

Abi-Farraj, F., Pacchierotti, C., Arenz, O., Neumann, G., and Giordano, P. R. (2019). A haptic shared-control architecture for guided multi-target robotic grasping. *IEEE Trans. Haptics* 13, 270–285. doi: 10.1109/TOH.2019.2913643

Aldoma, A., Marton, Z.-C., Tombari, F., Wohlkinger, W., Potthast, C., Zeisl, B., et al. (2012). Tutorial: Point cloud library: Three-dimensional object recognition and 6 dof pose estimation. *IEEE Robot. Autom. Magaz.* 19, 80–91. doi: 10.1109/MRA.2012.2206675

Ammanabrolu, P., and Riedl, M. O. (2019). Transfer in deep reinforcement learning using knowledge graphs. *arXiv [Preprint].* arXiv:1908.06556. doi: 10.18653/v1/D19-5301

Antonova, R., Kokic, M., Stork, J. A., and Kragic, D. (2018). Global search with bernoulli alternation kernel for task-oriented grasping informed by simulation. *arXiv [Preprint].* arXiv:1810.04438.

Ardón, P., Pairet, È., Petrick, R. P., Ramamoorthy, S., and Lohan, K. S. (2019). Learning grasp affordance reasoning through semantic relations. *IEEE Robot. Autom. Lett.* 4, 4571–4578. doi: 10.1109/LRA.2019.2933815

Astrom, K. J. (1965). Optimal control of Markov processes with incomplete state information. *J. Math. Anal. Appl.* 10, 174–205. doi: 10.1016/0022-247X(65)90154-X

Bauza, M., Alet, F., Lin, Y.-C., Lozano-Pérez, T., Kaelbling, L. P., Isola, P., et al. (2019). Omnipush: accurate, diverse, real-world dataset of pushing dynamics with rgb-d video. *arXiv [Preprint].* arXiv:1910.00618. doi: 10.1109/IROS40897.2019.8967920

Billard, A., and Kragic, D. (2019). Trends and challenges in robot manipulation. *Science* 364:eaat8414. doi: 10.1126/science.aat8414

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (Pittsburgh, PA), 144–152. doi: 10.1145/130385.130401

Bottarel, F., Vezzani, G., Pattacini, U., and Natale, L. (2020). GRASPA 1.0: GRASPA is a robot arm grasping performance benchmark. *IEEE Robot. Autom. Lett.* 5, 836–843. doi: 10.1109/LRA.2020.2965865

Boularias, A., Bagnell, J. A., and Stentz, A. (2014). "Efficient optimization for autonomous robotic manipulation of natural objects," in *AAAI* (Quebec City, QC), 2520–2526.

Boularias, A., Bagnell, J. A., and Stentz, A. (2015). "Learning to manipulate unknown objects in clutter by reinforcement," in *Twenty-Ninth AAAI Conference on Artificial Intelligence* (Austin, TX).

Bui, H.-D., Nguyen, H., La, H. M., and Li, S. (2020). A deep learning-based autonomous robot manipulator for sorting application. *arXiv [Preprint]*. arXiv:2009.03565. doi: 10.1109/IRC.2020.00055

Caldera, S., Rassau, A., and Chai, D. (2018). Review of deep learning methods in robotic grasp detection. *Multimodal Technol. Interact.* 2:57. doi: 10.3390/mti2030057

Calli, B., Singh, A., Bruce, J., Walsman, A., Konolige, K., Srinivasa, S., et al. (2017). Yale-CMU-Berkeley dataset for robotic manipulation research. *Int. J. Robot. Res.* 36, 261–268. doi: 10.1177/0278364917700714

Calli, B., Singh, A., Walsman, A., Srinivasa, S., Abbeel, P., and Dollar, A. M. (2015a). "The ycb object and model set: Towards common benchmarks for manipulation research," in *2015 International Conference on Advanced Robotics (ICAR)* (Istanbul: IEEE), 510–517. doi: 10.1109/ICAR.2015.7251504

Calli, B., Walsman, A., Singh, A., Srinivasa, S., Abbeel, P., and Dollar, A. M. (2015b). Benchmarking in manipulation research: the YCB object and model set and benchmarking protocols. *IEEE Robot. Autom. Mag.* 22, 36–52. doi: 10.1109/MRA.2015.2448951

Chen, C.-S., Chen, P.-C., and Hsu, C.-M. (2016). Three-dimensional object recognition and registration for robotic grasping systems using a modified viewpoint feature histogram. *Sensors* 16:1969. doi: 10.3390/s16111969

Chen, I.-M., and Burdick, J. W. (1993). Finding antipodal point grasps on irregularly shaped objects. *IEEE Trans. Robot. Autom.* 9, 507–512. doi: 10.1109/70.246063

Chen, X., Ye, Z., Sun, J., Fan, Y., Hu, F., Wang, C., et al. (2020). Transferable active grasping and real embodied dataset. *arXiv [Preprint]*. arXiv:2004.13358. doi: 10.1109/ICRA40945.2020.9197185

Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., et al. (2016). "Wide and deep learning for recommender systems," in *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems* (Boston, MA), 7–10. doi: 10.1145/2988450.2988454

Choi, C., Schwarting, W., DelPreto, J., and Rus, D. (2018). Learning object grasping for soft robot hands. *IEEE Robot. Autom. Lett.* 3, 2370–2377. doi: 10.1109/LRA.2018.2810544

Ciocarlie, M., Goldfeder, C., and Allen, P. (2007). "Dexterous grasping via eigengrasps: a low-dimensional approach to a high-complexity problem," in *Robotics: Science and Systems Manipulation Workshop-Sensing and Adapting to the Real World* (Atlanta, GA: Georgia Tech).

Ciocarlie, M. T., and Allen, P. K. (2009). Hand posture subspaces for dexterous robotic grasping. *Int. J. Robot. Res.* 28, 851–867. doi: 10.1177/0278364909105606

Dai, A., Ruizhongtai Qi, C., and Nießner, M. (2017). "Shape completion using 3d-encoder-predictor cnns and shape synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 5868–5877. doi: 10.1109/CVPR.2017.693

Dai, J., Li, Y., He, K., and Sun, J. (2016). "R-fcn: Object detection via region-based fully convolutional networks," in *Advances in Neural Information Processing Systems* (Barcelona), 379–387.

Deng, X., Xiang, Y., Mousavian, A., Eppner, C., Bretl, T., and Fox, D. (2020). "Self-supervised 6d object pose estimation for robot manipulation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)* (Paris: IEEE), 3665–3671. doi: 10.1109/ICRA40945.2020.9196714

Depierre, A., Dellandréa, E., and Chen, L. (2018). "Jacquard: a large scale dataset for robotic grasp detection," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid: IEEE), 3511–3516. doi: 10.1109/IROS.2018.8593950

Do, T.-T., Nguyen, A., and Reid, I. (2018). "Affordancenet: an end-to-end deep learning approach for object affordance detection," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 1–5. doi: 10.1109/ICRA.2018.8460902

Dong, Z., Liu, S., Zhou, T., Cheng, H., Zeng, L., Yu, X., et al. (2019). "PPR-Net: point-wise pose regression network for instance segmentation and 6d pose estimation in bin-picking scenarios," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Macau: IEEE), 1773–1780. doi: 10.1109/IROS40897.2019.8967895

Du, G., Wang, K., Lian, S., and Zhao, K. (2020). Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review. *Artif. Intellig. Rev.* 54, 1677–1734. doi: 10.1007/s10462-020-09888-5

Dyrstad, J. S., Bakken, M., Grøtli, E. I., Schulerud, H., and Mathiassen, J. R. (2018). "Bin picking of reflective steel parts using a dual-resolution convolutional neural network trained in a simulated environment," in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (Kuala Lumpur: IEEE), 530–537. doi: 10.1109/ROBIO.2018.8664766

Eppner, C., Mousavian, A., and Fox, D. (2019). A billion ways to grasp: an evaluation of grasp sampling schemes on a dense, physics-based grasp data set. *arXiv [Preprint]*. arXiv:1912.05604.

Fan, H., Su, H., and Guibas, L. J. (2017). "A point set generation network for 3d object reconstruction from a single image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 605–613. doi: 10.1109/CVPR.2017.264

Fang, H.-S., Wang, C., Gou, M., and Lu, C. (2020). "GraspNet-1Billion: a large-scale benchmark for general object grasping," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Seattle, WA: IEEE), 11444–11453. doi: 10.1109/CVPR42600.2020.01146

Fang, K., Zhu, Y., Garg, A., Kurenkov, A., Mehta, V., Fei-Fei, L., et al. (2020). Learning task-oriented grasping for tool manipulation from simulated self-supervision. *Int. J. Robot. Res.* 39, 202–216. doi: 10.1177/0278364919872545

Ficuciello, F., Migliozzi, A., Laudante, G., Falco, P., and Siciliano, B. (2019). Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework. *Sci. Robot.* 4:eaao4900. doi: 10.1126/scirobotics.aao4900

Finn, C., and Levine, S. (2017). "Deep visual foresight for planning robot motion," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 2786–2793. doi: 10.1109/ICRA.2017.7989324

Fischinger, D., Weiss, A., and Vincze, M. (2015). Learning grasps with topographic features. *Int. J. Robot. Res.* 34, 1167–1194. doi: 10.1177/0278364915577105

Fischler, M. A., and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24, 381–395. doi: 10.1145/358669.358692

Gal, Y., and Ghahramani, Z. (2016). "Dropout as a bayesian approximation: representing model uncertainty in deep learning," in *International Conference on Machine Learning* (New York, NY), 1050–1059.

Gamrian, S., and Goldberg, Y. (2019). "Transfer learning for related reinforcement learning tasks via image-to-image translation," in *International Conference on Machine Learning* (Long Beach, CA: PMLR), 2063–2072.

Gibson, J. J. (2014). *The Ecological Approach to Visual Perception: Classic Edition*. New York, NY; London: Psychology Press. doi: 10.4324/9781315740218

Glasmachers, T. (2017). Limits of end-to-end learning. *arXiv [Preprint]*. arXiv:1704.08305.

Goldfeder, C., Ciocarlie, M., Dang, H., and Allen, P. K. (2009). "The columbia grasp database," in *2009 IEEE International Conference on Robotics and Automation* (Kobe: IEEE), 1710–1716. doi: 10.1109/ROBOT.2009.5152709

Golovinskiy, A., and Funkhouser, T. (2009). "Min-cut based segmentation of point clouds," in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops* (Kyoto: IEEE), 39–46. doi: 10.1109/ICCVW.2009.5457721

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). "Generative adversarial nets," in *Advances in Neural Information Processing Systems* (Montreal, QC), 2672–2680.

Gualtieri, M., Kuczynski, J., Shultz, A. M., Ten Pas, A., Platt, R., and Yanco, H. (2017). "Open world assistive grasping using laser selection," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 4052–4057. doi: 10.1109/ICRA.2017.7989465

Gualtieri, M., and Platt, R. (2018). Learning 6-dof grasping and pick-place using attention focus. *arXiv [Preprint]*. arXiv:1806.06134.

Gualtieri, M., Ten Pas, A., Saenko, K., and Platt, R. (2016). "High precision grasp pose detection in dense clutter," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Daejeon: IEEE), 598–605. doi: 10.1109/IROS.2016.7759114

Guan, H., Li, J., and Yan, R. (2019). "An efficient robotic grasping pipeline base on fully convolutional neural network," in *2019 5th International Conference on Control, Automation and Robotics (ICCAR)* (Beijing: IEEE), 172–176. doi: 10.1109/ICCAR.2019.8813502

Guo, N., Zhang, B., Zhou, J., Zhan, K., and Lai, S. (2020). Pose estimation and adaptable grasp configuration with point cloud registration and geometry understanding for fruit grasp planning. *Comput. Electron. Agric.* 179:105818. doi: 10.1016/j.compag.2020.105818

Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., and Bennamoun, M. (2020). Deep learning for 3d point clouds: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* doi: 10.1109/TPAMI.2020.3005434

Hadsell, R., Rao, D., Rusu, A. A., and Pascanu, R. (2020). Embracing change: continual learning in deep neural networks. *Trends Cogn. Sci.* 24, P1028–P1040. doi: 10.1016/j.tics.2020.09.004

Haustein, J. A., Hang, K., and Kragic, D. (2017). "Integrating motion and hierarchical fingertip grasp planning," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 3439–3446. doi: 10.1109/ICRA.2017.7989392

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision* (Venice), 2961–2969. doi: 10.1109/ICCV.2017.322

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 770–778. doi: 10.1109/CVPR.2016.90

Hernandez, C., Bharatheesha, M., Ko, W., Gaiser, H., Tan, J., van Deurzen, K., et al. (2016). "Team delft's robot winner of the amazon picking challenge 2016," in *Robot World Cup* (Leipzig: Springer), 613–624. doi: 10.1007/978-3-319-68792-6_51

Hinton, G. E., Sejnowski, T. J., and Poggio, T. A. (1999). *Unsupervised Learning: Foundations of Neural Computation*. Cambridge; London: MIT Press. doi: 10.7551/mitpress/7011.001.0001

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv [Preprint]*. arXiv:1704.04861.

Hu, J., Sun, Y., Li, G., Jiang, G., and Tao, B. (2019). Probability analysis for grasp planning facing the field of medical robotics. *Measurement* 141, 227–234. doi: 10.1016/j.measurement.2019.03.010

Ioffe, S., and Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv [Preprint]*. arXiv:1502.03167.

James, S., Davison, A. J., and Johns, E. (2017). Transferring end-to-end visuomotor control from simulation to real world for a multi-stage task. *arXiv [Preprint]*. arXiv:1707.02267.

Jaśkowski, M., Swiatkowski, J., Zajac, M., Klimek, M., and Potiuk, J., Rybicki, P., et al. (2018). Improved GQ-CNN: deep learning model for planning robust grasps. *arXiv [Preprint]*. arXiv:1802.05992.

Jiang, L., Shi, S., Qi, X., and Jia, J. (2018). "Gal: Geometric adversarial loss for single-view 3d-object reconstruction," in *Proceedings of the European Conference on Computer Vision (ECCV)* (Munich), 802–816. doi: 10.1007/978-3-030-01237-3_49

Jiang, P., Ishihara, Y., Sugiyama, N., Oaki, J., Tokura, S., Sugahara, A., et al. (2020). Depth image–based deep learning of grasp planning for textureless planar-faced objects in vision-guided robotic bin-picking. *Sensors* 20:706. doi: 10.3390/s20030706

Jiang, Y., Lim, M., and Saxena, A. (2012). Learning object arrangements in 3d scenes using human context. *arXiv [Preprint]*. arXiv:1206.6462.

Jiang, Y., Moseson, S., and Saxena, A. (2011). "Efficient grasping from rgbd images: Learning using a new rectangle representation," in *2011 IEEE International Conference on Robotics and Automation* (Shanghai: IEEE), 3304–3311. doi: 10.1109/ICRA.2011.5980145

Johns, E., Leutenegger, S., and Davison, A. J. (2016). "Deep learning a grasp function for grasping under gripper pose uncertainty," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Daejeon: IEEE), 4461–4468. doi: 10.1109/IROS.2016.7759657

Kang, H., Zhou, H., Wang, X., and Chen, C. (2020). Real-time fruit recognition and grasping estimation for robotic apple harvesting. *Sensors* 20:5670. doi: 10.3390/s20195670

Kanoulas, D., Lee, J., Caldwell, D. G., and Tsagarakis, N. G. (2017). Visual grasp affordance localization in point clouds using curved contact patches. *Int. J. Humanoid Robot.* 14:1650028. doi: 10.1142/S0219843616500286

Kappler, D., Bohg, J., and Schaal, S. (2015). "Leveraging big data for grasp planning," in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA: IEEE), 4304–4311. doi: 10.1109/ICRA.2015.7139793

Kappler, D., Schaal, S., and Bohg, J. (2016). "Optimizing for what matters: the top grasp hypothesis," in *2016 IEEE International Conference on Robotics and Automation (ICRA)* (Stockholm: IEEE), 2167–2174. doi: 10.1109/ICRA.2016.7487367

Katz, D., Venkatraman, A., Kazemi, M., Bagnell, J. A., and Stentz, A. (2014). Perceiving, learning, and exploiting object affordances for autonomous pile manipulation. *Auton. Robots* 37, 369–382. doi: 10.1007/s10514-014-9407-y

Kehoe, B., Berenson, D., and Goldberg, K. (2012). "Toward cloud-based grasping with uncertainty in shape: estimating lower bounds on achieving force closure with zero-slip push grasps," in *2012 IEEE International Conference on Robotics and Automation* (St Paul, MN: IEEE), 576–583. doi: 10.1109/ICRA.2012.6224781

Kiatos, M., and Malassiotis, S. (2019). "Grasping unknown objects by exploiting complementarity with robot hand geometry," in *International Conference on Computer Vision Systems* (Thessaloniki: Springer), 88–97. doi: 10.1007/978-3-030-34995-0_8

Kim, C.-J., and Nelson, C. R. (1999). *State-Space Models With Regime Switching: Classical and Gibbs-Sampling Approaches With Applications*. Cambridge; London: MIT Press Books 1.

Kim, J., Iwamoto, K., Kuffner, J. J., Ota, Y., and Pollard, N. S. (2013). Physically based grasp quality evaluation under pose uncertainty. *IEEE Trans. Robot.* 29, 1424–1439. doi: 10.1109/TRO.2013.2273846

Kingma, D. P., and Welling, M. (2013). Auto-encoding variational bayes. *arXiv [Preprint]*. arXiv:1312.6114.

Kleeberger, K., Bormann, R., Kraus, W., and Huber, M. F. (2020). A survey on learning-based robotic grasping. *Curr. Robot. Rep.* 1, 239–249. doi: 10.1007/s43154-020-00021-6

Klokov, R., and Lempitsky, V. (2017). "Escape from cells: deep kd-networks for the recognition of 3d point cloud models," in *Proceedings of the IEEE International Conference on Computer Vision* (Venice), 863–872. doi: 10.1109/ICCV.2017.99

Kokic, M., Stork, J. A., Haustein, J. A., and Kragic, D. (2017). "Affordance detection for task-specific grasping using deep learning," in *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)* (Birmingham: IEEE), 91–98. doi: 10.1109/HUMANOIDS.2017.8239542

Koppula, H. S., Gupta, R., and Saxena, A. (2013). Learning human activities and object affordances from rgb-d videos. *Int. J. Robot. Res.* 32, 951–970. doi: 10.1177/0278364913478446

Kroemer, O., Niekum, S., and Konidaris, G. (2019). A review of robot learning for manipulation: challenges, representations, and algorithms. *arXiv [Preprint]*. arXiv:1907.03146.

Kumar, V., Herman, T., Fox, D., Birchfield, S., and Tremblay, J. (2019). Contextual reinforcement learning of visuo-tactile multi-fingered grasping policies. *arXiv [Preprint]*. arXiv:1911.09233.

Kumra, S., and Kanan, C. (2017). "Robotic grasp detection using deep convolutional neural networks," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Vancouver, BC: IEEE), 769–776. doi: 10.1109/IROS.2017.8202237

Laskey, M., Mahler, J., McCarthy, Z., Pokorny, F. T., Patil, S., Van Den Berg, J., et al. (2015). "Multi-armed bandit models for 2d grasp planning with uncertainty," in *2015 IEEE International Conference on Automation Science and Engineering (CASE)* (Gothenburg: IEEE), 572–579. doi: 10.1109/CoASE.2015.7294140

Le, Q. V., Kamm, D., Kara, A. F., and Ng, A. Y. (2010). "Learning to grasp objects with multiple contact points," in *2010 IEEE International Conference on Robotics and Automation* (Anchorage, AK: IEEE), 5062–5069. doi: 10.1109/ROBOT.2010.5509508

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791

Lenz, I., Lee, H., and Saxena, A. (2015). Deep learning for detecting robotic grasps. *Int. J. Robot. Res.* 34, 705–724. doi: 10.1177/0278364914549607

Lesort, T., Lomonaco, V., Stoian, A., Maltoni, D., Filliat, D., and Díaz-Rodríguez, N. (2020). Continual learning for robotics: definition, framework, learning strategies, opportunities and challenges. *Inform. Fusion* 58, 52–68. doi: 10.1016/j.inffus.2019.12.004

Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J., and Quillen, D. (2018). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *Int. J. Robot. Res.* 37, 421–436. doi: 10.1177/0278364917710318

Li, H., Tan, J., and He, H. (2020). *MagicHand: Context-Aware Dexterous Grasping Using an Anthropomorphic Robotic Hand.* Paris: IEEE. doi: 10.1109/ICRA40945.2020.9196538

Li, P., DeRose, B., Mahler, J., Ojea, J. A., Tanwani, A. K., and Goldberg, K. (2018). "Dex-net as a service (dnaas): a cloud-based robust robot grasp planning system," in *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)* (Munich: IEEE), 1420–1427. doi: 10.1109/COASE.2018.8560447

Li, R., and Qiao, H. (2019). A survey of methods and strategies for high-precision robotic grasping and assembly tasks—some new trends. *IEEE ASME Trans. Mechatron.* 24, 2718–2732. doi: 10.1109/TMECH.2019.2945135

Li, Y., Schomaker, L., and Kasaei, S. H. (2020). Learning to grasp 3D objects using deep residual U-nets. *arXiv [Preprint].* arXiv:2002.03892. doi: 10.1109/RO-MAN47096.2020.9223541

Liang, H., Ma, X., Li, S., Görner, M., Tang, S., Fang, B., et al. (2019). "Pointnetgpd: detecting grasp configurations from point sets," in *2019 International Conference on Robotics and Automation (ICRA)* (Montreal, QC: IEEE), 3629–3635. doi: 10.1109/ICRA.2019.8794435

Lin, H.-I., and Cong, M. N. (2019). "Inference of 6-DOF robot grasps using point cloud data," in *2019 19th International Conference on Control, Automation and Systems (ICCAS)* (Chongqing: IEEE), 944–948. doi: 10.23919/ICCAS47443.2019.8971464

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017). "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 2117–2125. doi: 10.1109/CVPR.2017.106

Lin, Y., Zeng, L., Dong, Z., and Fu, X. (2019). "A vision-guided robotic grasping method for stacking scenes based on deep learning," in *2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* (Chongqing: IEEE), 91–96. doi: 10.1109/IMCEC46724.2019.8983819

Liu, C., Fang, B., Sun, F., Li, X., and Huang, W. (2019). Learning to grasp familiar objects based on experience and objects' shape affordance. *IEEE Trans. Syst. Man Cybern. Syst.* 49, 2710–2723. doi: 10.1109/TSMC.2019.2901955

Liu, H., and Cao, C. (2020). "Grasp pose detection based on point cloud shape simplification," in *IOP Conference Series: Materials Science and Engineering* (Wuhan: IOP Publishing), 012007. doi: 10.1088/1757-899X/717/1/012007

Liu, M., Pan, Z., Xu, K., Ganguly, K., and Manocha, D. (2019). Generating grasp poses for a high-dof gripper using neural networks. *arXiv [Preprint].* arXiv:1903.00425. doi: 10.1109/IROS40897.2019.8968115

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). "Ssd: single shot multibox detector," in *European Conference on Computer Vision* (Amsterdam: Springer), 21–37. doi: 10.1007/978-3-319-46448-0_2

Liu, W., Daruna, A., and Chernova, S. (2020). *CAGE: Context-Aware Grasping Engine.* (Paris: IEEE). doi: 10.1109/ICRA40945.2020.9197289

Liu, W., Sun, J., Li, W., Hu, T., and Wang, P. (2019). Deep learning on point clouds and its application: a survey. *Sensors* 19:4188. doi: 10.3390/s19194188

Liu, Y., Hu, Y., Gao, Y., Chen, Y., and Fan, C. (2019). Value function transfer for deep multi-agent reinforcement learning based on N-step returns in *IJCAI* (Macao), 457–463. doi: 10.24963/ijcai.2019/65

Llopart, A., Ravn, O., and Andersen, N. A. (2017). "Door and cabinet recognition using convolutional neural nets and real-time method fusion for handle detection and grasping," in *2017 3rd International Conference on Control, Automation and Robotics (ICCAR)* (Nagoya: IEEE), 144–149. doi: 10.1109/ICCAR.2017.7942676

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Trans. Inform. Theory* 28, 129–137. doi: 10.1109/TIT.1982.1056489

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA), 3431–3440. doi: 10.1109/CVPR.2015.7298965

Lopes, V., Alexandre, L. A., and Fernandes, M. (2018). ""Less is more": Simplifying point clouds to improve grasping performance", in *2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)* (Torres Vedras: IEEE), 256–260. doi: 10.1109/ICARSC.2018.8374192

Lou, X., Yang, Y., and Choi, C. (2020). *Learning to Generate 6-DoF Grasp Poses With Reachability Awareness.* Paris: IEEE. doi: 10.1109/ICRA40945.2020.9197413

Lu, H., and Shi, H. (2020). Deep learning for 3D point cloud understanding: a survey. *arXiv [Preprint].* arXiv:2009.08920.

Lu, Q., Chenna, K., Sundaralingam, B., and Hermans, T. (2020). "Planning multi-fingered grasps as probabilistic inference in a learned deep network," in *Robotics Research*, eds N. M. Amato, G. Hager, S. Thomas, and M. Torres-Torriti (Puerto Varas: Springer), 455–472. doi: 10.1007/978-3-030-28619-4_35

Lu, Q., and Hermans, T. (2019). Modeling grasp type improves learning-based grasp planning. *IEEE Robot. Autom. Lett.* 4, 784–791. doi: 10.1109/LRA.2019.2893410

Lundell, J., Verdoja, F., and Kyrki, V. (2019). Robust grasp planning over uncertain shape completions. *arXiv [Preprint].* arXiv:1903.00645. doi: 10.1109/IROS40897.2019.8967816

Luo, S., Bimbo, J., Dahiya, R., and Liu, H. (2017). Robotic tactile perception of object properties: a review. *Mechatronics* 48, 54–67. doi: 10.1016/j.mechatronics.2017.11.002

Mahler, J., and Goldberg, K. (2017). "Learning deep policies for robot bin picking by simulating robust grasping sequences," in *Conference on Robot Learning* (Mountain View, CA), 515–524.

Mahler, J., Liang, J., Niyaz, S., Laskey, M., Doan, R., Liu, X., et al. (2017). Dex-net 2.0: deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics. *arXiv [Preprint].* arXiv:1703.09312. doi: 10.15607/RSS.2017.XIII.058

Mahler, J., Matl, M., Liu, X., Li, A., Gealy, D., and Goldberg, K. (2018). "Dex-Net 3.0: computing robust vacuum suction grasp targets in point clouds using a new analytic model and deep learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 1–8. doi: 10.1109/ICRA.2018.8460887

Mahler, J., Matl, M., Satish, V., Danielczuk, M., DeRose, B., McKinley, S., et al. (2019). Learning ambidextrous robot grasping policies. *Sci. Robot.* 4:eaau4984. doi: 10.1126/scirobotics.aau4984

Mahler, J., Patil, S., Kehoe, B., Van Den Berg, J., Ciocarlie, M., Abbeel, P., et al. (2015). "Gp-gpis-opt: Grasp planning with shape uncertainty using gaussian process implicit surfaces and sequential convex programming," in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA: IEEE), 4919–4926. doi: 10.1109/ICRA.2015.7139882

Mahler, J., Pokorny, F. T., Hou, B., Roderick, M., Laskey, M., Aubry, M., et al. (2016). "Dex-net 1.0: a cloud-based network of 3d objects for robust grasp planning using a multi-armed bandit model with correlated rewards," in *2016 IEEE International Conference on Robotics and Automation (ICRA)* (Stockholm: IEEE), 1957–1964. doi: 10.1109/ICRA.2016.7487342

Manuelli, L., Gao, W., Florence, P., and Tedrake, R. (2019). kpam: Keypoint affordances for category-level robotic manipulation. *arXiv [Preprint].* arXiv:1903.06684.

Mar, T., Tikhanoff, V., Metta, G., and Natale, L. (2017). "Self-supervised learning of tool affordances from 3D tool representation through parallel SOM mapping," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 894–901. doi: 10.1109/ICRA.2017.7989110

Matsumoto, E., Saito, M., Kume, A., and Tan, J. (2020). "End-to-end learning of object grasp poses in the Amazon Robotics Challenge," in *Advances on Robotic Item Picking*, eds A. Causo, J. Durham, K. Hauser, K. Okada, and A. Rodriguez (Berlin; Heidelberg: Springer), 63–72. doi: 10.1007/978-3-030-35679-8_6

Miller, A. T., and Allen, P. K. (2004). Graspit! a versatile simulator for robotic grasping. *IEEE Robot. Autom. Magaz.* 11, 110–122. doi: 10.1109/MRA.2004.1371616

Mitchell, E., Engin, S., Isler, V., and Lee, D. D. (2019). Higher-order function networks for learning composable 3D object representations. *arXiv [Preprint].* arXiv:1907.10388.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533. doi: 10.1038/nature14236

Mohammed, M. Q., Chung, K. L., and Chyi, C. S. (2020). Review of deep reinforcement learning-based object grasping: techniques, open challenges and recommendations. *IEEE Access* 8, 178450–178481. doi: 10.1109/ACCESS.2020.3027923

Monica, R., and Aleotti, J. (2020). Point cloud projective analysis for part-based grasp planning. *IEEE Robot. Autom. Lett.* 5, 4695–4702. doi: 10.1109/LRA.2020.3003883

Moreno, P., Hornstein, J., and Santos-Victor, J. (2011). *Learning to Grasp From Point Clouds*. Lisboa: Instituto Superior Tecnico and Instituto de Sistemas e Robotica.

Morrison, D., Corke, P., and Leitner, J. (2018a). Closing the loop for robotic grasping: a real-time, generative grasp synthesis approach. *arXiv [Preprint].* arXiv:1804.05172. doi: 10.15607/RSS.2018.XIV.021

Morrison, D., Tow, A. W., Mctaggart, M., Smith, R., Kelly-Boxall, N., Wade-Mccue, S., et al. (2018b). "Cartman: the low-cost cartesian manipulator that won the amazon robotics challenge," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 7757–7764. doi: 10.1109/ICRA.2018.8463191

Mousavian, A., Eppner, C., and Fox, D. (2019). "6-dof graspnet: variational grasp generation for object manipulation," in *Proceedings of the IEEE International Conference on Computer Vision* (Seoul), 2901–2910. doi: 10.1109/ICCV.2019.00299

Murali, A., Mousavian, A., Eppner, C., Paxton, C., and Fox, D. (2020). "6-dof grasping for target-driven object manipulation in clutter," in *2020 IEEE International Conference on Robotics and Automation (ICRA)* (Paris: IEEE), 6232–6238. doi: 10.1109/ICRA40945.2020.9197318

Murray, R. M., Li, Z., Sastry, S. S., and Sastry, S. S. (1994). *A Mathematical Introduction to Robotic Manipulation.* Boca Raton, FL: CRC Press.

Nekrasov, V., Shen, C., and Reid, I. (2018). Light-weight refinenet for real-time semantic segmentation. *arXiv [Preprint].* arXiv:1810.03272.

Nishikawa, K., Ohya, J., Ogata, H., Hashimoto, K., Matsuzawa, T., Imai, A., et al. (2019). Automatic estimation of the position and orientation of the drill to be grasped and manipulated by the disaster response robot based on analyzing depth camera information. *Electron. Imaging* 2019, 452–457. doi: 10.2352/ISSN.2470-1173.2019.7.IRIACV-452

Okamura, A. M., Smaby, N., and Cutkosky, M. R. (2000). "An overview of dexterous manipulation," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)* (San Francisco, CA: IEEE), 255–262.

Ottenhaus, S., Renninghoff, D., Grimm, R., Ferreira, F., and Asfour, T. (2019). "Visuo-haptic grasping of unknown objects based on gaussian process implicit surfaces and deep learning," in *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)* (Toronto, ON: IEEE), 402–409. doi: 10.1109/Humanoids43949.2019.9035002

Pan, S. J., and Yang, Q. (2009). A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22, 1345–1359. doi: 10.1109/TKDE.2009.191

Parhar, T., Baweja, H., Jenkins, M., and Kantor, G. (2018). "A deep learning-based stalk grasping pipeline," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 1–5. doi: 10.1109/ICRA.2018.8460597

Pas, A. T., Gualtieri, M., Saenko, K., and Platt, R. (2017). Grasp pose detection in point clouds. *Int. J. Rob. Res.* 36, 1455–1473. doi: 10.1177/0278364917735594

Pas, A. T., and Platt, R. (2013). Localizing grasp affordances in 3-D points clouds using taubin quadric fitting. *arXiv [Preprint].* arXiv:1311.3192.

Pas, A. T., and Platt, R. (2015). Using geometry to detect grasps in 3d point clouds. *arXiv [Preprint].* arXiv:1501.03100.

Pas, A. T., and Platt, R. (2016). "Localizing handle-like grasp affordances in 3d point clouds," in *Experimental Robotics*, eds M. A. Hsieh, O. Khatib, and V. Kumar (Marrakech; Essaouira: Springer), 623–638. doi: 10.1007/978-3-319-23778-7_41

Prattichizzo, D., and Trinkle, J. C. (2016). "Grasping," in *Springer Handbook of Robotics*, eds B. Siciliano and O. Khatib (Berlin; Heidelberg: Springer), 955–988. doi: 10.1007/978-3-319-32552-1_38

Prattichizzo, D., Trinkle, J. C., Siciliano, B., and Khatib, O. (2008). "Springer handbook of robotics," in *Grasping*, eds B. Siciliano and O. Khatib (Berlin; Heidelberg: Springer), 671–700. doi: 10.1007/978-3-540-30301-5_29

Qi, C. R., Su, H., Mo, K., and Guibas, L. J. (2017a). "Pointnet: deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 652–660.

Qi, C. R., Yi, L., Su, H., and Guibas, L. J. (2017b). "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," in *Advances in Neural Information Processing Systems* (Long Beach, CA), 5099–5108.

Qian, K., Jing, X., Duan, Y., Zhou, B., Fang, F., Xia, J., et al. (2020). Grasp pose detection with affordance-based task constraint learning in single-view point clouds. *J. Intellig. Robot. Syst.* 100, 145–163. doi: 10.1007/s10846-020-01202-3

Qin, Y., Chen, R., Zhu, H., Song, M., Xu, J., and Su, H. (2020). "S4g: amodal single-view single-shot se (3) grasp detection in cluttered scenes," in *Conference on Robot Learning* (Osaka: PMLR), 53–65.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 779–788. doi: 10.1109/CVPR.2016.91

Redmon, J., and Farhadi, A. (2018). Yolov3: an incremental improvement. *arXiv [Preprint].* arXiv:1804.02767.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). "Faster r-cnn: towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems* (Montreal, QC), 91–99.

Richardson, M., and Domingos, P. (2006). Markov logic networks. *Mach. Learn.* 62, 107–136. doi: 10.1007/s10994-006-5833-1

Roy, A. G., Navab, N., and Wachinger, C. (2018). "Concurrent spatial and channel 'squeeze and excitation'in fully convolutional networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Granada: Springer), 421–429. doi: 10.1007/978-3-030-00928-1_48

Ruiz-del-Solar, J., Loncomilla, P., and Soto, N. (2018). A survey on deep learning methods for robot vision. *arXiv [Preprint].* arXiv:1803.10862.

Runz, M., Buffier, M., and Agapito, L. (2018). "Maskfusion: Real-time recognition, tracking and reconstruction of multiple moving objects," in *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Munich: IEEE), 10–20. doi: 10.1109/ISMAR.2018.00024

Russell, S., and Norvig, P. (2002). *Artificial Intelligence: A Modern Approach, 2nd Edn*. Prentice Hall

Rusu, R. B., Bradski, G., Thibaux, R., and Hsu, J. (2010). "Fast 3d recognition and pose using the viewpoint feature histogram," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Taipei: IEEE), 2155–2162. doi: 10.1109/IROS.2010.5651280

Rusu, R. B., and Cousins, S. (2011). "3d is here: Point cloud library (pcl)," in *2011 IEEE International Conference on Robotics and Automation* (Shanghai: IEEE), 1–4. doi: 10.1109/ICRA.2011.5980567

Satish, V., Mahler, J., and Goldberg, K. (2019). On-policy dataset synthesis for learning robot grasping policies using fully convolutional deep networks. *IEEE Robot. Autom. Lett.* 4, 1357–1364. doi: 10.1109/LRA.2019.2895878

Saut, J.-P., Sahbani, A., El-Khoury, S., and Perdereau, V. (2007). "Dexterous manipulation planning using probabilistic roadmaps in continuous grasp subspaces," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems* (San Diego, CA: IEEE), 2907–2912. doi: 10.1109/IROS.2007.4399090

Schmidt, P., Vahrenkamp, N., Wächter, M., and Asfour, T. (2018). "Grasping of unknown objects using deep convolutional neural networks based on depth images," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 6831–6838. doi: 10.1109/ICRA.2018.8463204

Schnaubelt, M., Kohlbrecher, S., and von Stryk, O. (2019). "Autonomous assistance for versatile grasping with rescue robots," in *2019 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (Würzburg: IEEE), 210–215. doi: 10.1109/SSRR.2019.8848947

Schwarz, M., Lenz, C., García, G. M., Koo, S., Periyasamy, A. S., Schreiber, M., et al. (2018). "Fast object learning and dual-arm coordination for cluttered stowing, picking, and packing," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 3347–3354. doi: 10.1109/ICRA.2018.8461195

Shao, L., Ferreira, F., Jorda, M., Nambiar, V., Luo, J., Solowjow, E., et al. (2020). UniGrasp: learning a unified model to grasp with multifingered robotic hands. *IEEE Robot. Autom. Lett.* 5, 2286–2293. doi: 10.1109/LRA.2020.2969946

Singh, N., Blum, Z., and Renjith, N. (2018). *Point Cloud Grasp Classification for Robot Grasping.* Stanford, CA: Stanford University.

Song, K.-T., Chang, Y.-H., and Chen, J.-H. (2019). "3D vision for object grasp and obstacle avoidance of a collaborative robot," in *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)* (Hong Kong: IEEE), 254–258. doi: 10.1109/AIM.2019.8868694

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.

Sun, G.-J., and Lin, H.-Y. (2020). "Robotic grasping using semantic segmentation and primitive geometric model based 3D pose estimation," in *2020 IEEE/SICE International Symposium on System Integration (SII)* (Honolulu, HI: IEEE), 337–342. doi: 10.1109/SII46433.2020.9026297

Sun, X., Xiao, B., Wei, F., Liang, S., and Wei, Y. (2018). "Integral human pose regression," in *Proceedings of the European Conference on Computer Vision (ECCV)* (Munich), 529–545. doi: 10.1007/978-3-030-01231-1_33

Sung, J., Lenz, I., and Saxena, A. (2017). "Deep multimodal embedding: Manipulating novel objects with point-clouds, language and trajectories," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 2794–2801. doi: 10.1109/ICRA.2017.7989325

Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Cambridge; London: MIT Press.

Sutton, R. S., McAllester, D., Singh, S., and Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. *Adv. Neural Inform. Process. Syst.* 12, 1057–1063.

Taubin, G. (1991). Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Trans. Pattern Anal. Mach. Intellig.* 11, 1115–1138. doi: 10.1109/34.103273

Tian, N., Matl, M., Mahler, J., Zhou, Y. X., Staszak, S., Correa, C., et al. (2017). "A cloud robot system using the dexterity network and berkeley robotics and automation as a service (brass)," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 1615–1622. doi: 10.1109/ICRA.2017.7989192

Tirinzoni, A., Sessa, A., Pirotta, M., and Restelli, M. (2018). Importance weighted transfer of samples in reinforcement learning. *arXiv [Preprint].* arXiv:1805.10886.

Torii, T., and Hashimoto, M. (2018). "Model-less estimation method for robot grasping parameters using 3D shape primitive approximation," in *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)* (Munich: IEEE), 580–585. doi: 10.1109/COASE.2018.8560417

Tosun, T., Mitchell, E., Eisner, B., Huh, J., Lee, B., Lee, D., et al. (2019). Pixels to plans: learning non-prehensile manipulation by imitating a planner. *arXiv [Preprint].* arXiv:1904.03260. doi: 10.1109/IROS40897.2019.8968224

Tosun, T., Yang, D., Eisner, B., Isler, V., and Lee, D. (2020). Robotic grasping through combined image-based grasp proposal and 3D reconstruction. *arXiv [Preprint].* arXiv:2003.01649.

Ückermann, A., Elbrechter, C., Haschke, R., and Ritter, H. (2012). "3D scene segmentation for autonomous robot grasping," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Vilamoura-Algarve: IEEE), 1734–1740. doi: 10.1109/IROS.2012.6385692

Van der Merwe, M., Lu, Q., Sundaralingam, B., Matak, M., and Hermans, T. (2020). "Learning continuous 3d reconstructions for geometrically aware grasping," in *2020 IEEE International Conference on Robotics and Automation (ICRA)* (Paris: IEEE), 11516–11522. doi: 10.1109/ICRA40945.2020.9196981

Varadarajan, K. M., and Vincze, M. (2012). "Afrob: The affordance network ontology for robots," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Vilamoura-Algarve: IEEE), 1343–1350. doi: 10.1109/IROS.2012.6386232

Varley, J., DeChant, C., Richardson, A., Ruales, J., and Allen, P. (2017). "Shape completion enabled robotic grasping," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Vancouver, BC: IEEE), 2442–2447. doi: 10.1109/IROS.2017.8206060

Varley, J., Weisz, J., Weiss, J., and Allen, P. (2015). "Generating multi-fingered robotic grasps via deep learning," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Hamburg: IEEE), 4415–4420. doi: 10.1109/IROS.2015.7354004

Viereck, U., Pas, A. T., Saenko, K., and Platt, R. (2017). Learning a visuomotor controller for real world robotic grasping using simulated depth images. *arXiv [Preprint].* arXiv:1706.04652.

Vohra, M., Prakash, R., and Behera, L. (2019). "Real-time grasp pose estimation for novel objects in densely cluttered environment," in *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (New Delhi: IEEE), 1–6. doi: 10.1109/RO-MAN46459.2019.8956438

Wang, C., Zhang, X., Zang, X., Liu, Y., Ding, G., Yin, W., et al. (2020). Feature sensing and robotic grasping of objects with uncertain information: a review. *Sensors* 20:3707. doi: 10.3390/s20133707

Wang, H.-Y., and Ling, W.-K. (2016). "Robotic grasp detection using deep learning and geometry model of soft hand," in *2016 IEEE International Conference on Consumer Electronics-China (ICCE-China)* (Guangzhou: IEEE), 1–6. doi: 10.1109/ICCE-China.2016.7849757

Wang, T., Yang, C., Kirchner, F., Du, P., Sun, F., and Fang, B. (2019). Multimodal grasp data set: a novel visual–tactile data set for robotic manipulation. *Int. J. Adv. Robot. Syst.* 16:1729881418821571. doi: 10.1177/1729881418821571

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., and Solomon, J. M. (2019). Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph.* 38, 1–12. doi: 10.1145/3326362

Wang, Z., Xu, Y., He, Q., Fang, Z., Xu, G., and Fu, J. (2020). Grasping pose estimation for SCARA robot based on deep learning of point cloud. *Int. J. Adv. Manuf. Technol.* 108, 1217–1231. doi: 10.1007/s00170-020-05257-2

Watkins-Valls, D., Varley, J., and Allen, P. (2019). "Multi-modal geometric learning for grasping and manipulation," in *2019 International Conference on Robotics and Automation (ICRA)* (Montreal, QC: IEEE), 7339–7345. doi: 10.1109/ICRA.2019.8794233

Weisz, J., and Allen, P. K. (2012). "Pose error robust grasping from contact wrench space metrics," in *2012 IEEE International Conference on Robotics and Automation* (St Paul, MN: IEEE), 557–562. doi: 10.1109/ICRA.2012.6224697

Wu, B., Akinola, I., and Allen, P. K. (2019). Pixel-attentive policy gradient for multi-fingered grasping in cluttered scenes. *arXiv [Preprint].* arXiv:1903.03227.

Wu, C., Chen, J., Cao, Q., Zhang, J., Tai, Y., Sun, L., et al. (2020). Grasp proposal networks: an end-to-end solution for visual learning of robotic grasps. *arXiv [Preprint].* arXiv:2009.12606.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., et al. (2015). "3d shapenets: a deep representation for volumetric shapes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA), 1912–1920.

Wurman, P. R., and Romano, J. M. (2015). The amazon picking challenge 2015. *IEEE Robot. Autom. Magaz.* 22, 10–12. doi: 10.1109/MRA.2015.2452071

Xiang, Y., Schmidt, T., Narayanan, V., and Fox, D. (2017). Posecnn: a convolutional neural network for 6d object pose estimation in cluttered scenes. *arXiv [Preprint].* arXiv:1711.00199. doi: 10.15607/RSS.2018.XIV.019

Xu, Y., Wang, L., Yang, A., and Chen, L. (2019). GraspCNN: real-time grasp detection using a new oriented diameter circle representation. *IEEE Access* 7, 159322–159331. doi: 10.1109/ACCESS.2019.2950535

Xu, Z., and Topcu, U. (2019). Transfer of temporal logic formulas in reinforcement learning. *arXiv [Preprint].* arXiv:1909.04256.

Yan, X., Khansari, M., Hsu, J., Gong, Y., Bai, Y., Pirk, S., et al. (2019). Data-efficient learning for sim-to-real robotic grasping using deep point cloud prediction networks. *arXiv [Preprint].* arXiv:1906.08989.

Yang, W., Paxton, C., Cakmak, M., and Fox, D. (2020). Human grasp classification for reactive human-to-robot handovers. *arXiv [Preprint].* arXiv:2003.06000. doi: 10.1109/IROS45743.2020.9341004

Yu, H., Lai, Q., Liang, Y., Wang, Y., and Xiong, R. (2019). "A cascaded deep learning framework for real-time and robust grasp planning," in *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (Dali: IEEE), 1380–1386.

Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., and Huang, T. S. (2018). "Generative image inpainting with contextual attention," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT), 5505–5514. doi: 10.1109/CVPR.2018.00577

Yu, Q., Shang, W., Zhao, Z., Cong, S., and Li, Z. (2020). "Robotic grasping of unknown objects using novel multilevel convolutional neural networks: from parallel gripper to dexterous hand," in *IEEE Transactions on Automation Science and Engineering* (New York, NY).

Yu, S., Zhai, D.-H., Wu, H., Yang, H., and Xia, Y. (2020). "Object recognition and robot grasping technology based on RGB-D data," in *2020 39th Chinese Control Conference (CCC)* (Shenyang: IEEE), 3869–3874.

Yu, Y., Cao, Z., Liang, S., Geng, W., and Yu, J. (2020). A novel vision-based grasping method under occlusion for manipulating robotic system. *IEEE Sensors J*. 20, 10996–11006. doi: 10.1109/JSEN.2020.2995395

Yu, Y., Cao, Z., Liang, S., Liu, Z., Yu, J., and Chen, X. (2019). "A grasping CNN with image segmentation for mobile manipulating robot," in *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (Dali: IEEE), 1688–1692.

Zapata-Impata, B. S., Mateo Agulló, C., Gil, P., and Pomares, J. (2017). "Using geometry to detect grasping points on 3D unknown point cloud," in *Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics (Madrid)*, 154–161. doi: 10.5220/0006470701540161

Zeng, A. (2019). *Learning Visual Affordances for Robotic Manipulation* (doctoral dissertation). Princeton University.

Zeng, A., Song, S., Lee, J., Rodriguez, A., and Funkhouser, T. (2020). Tossingbot: Learning to throw arbitrary objects with residual physics. *IEEE Trans. Robot*. 36, 1307–1319. doi: 10.15607/RSS.2019.XV.004

Zeng, A., Song, S., Welker, S., Lee, J., Rodriguez, A., and Funkhouser, T. (2018a). "Learning synergies between pushing and grasping with self-supervised deep reinforcement learning," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid: IEEE), 4238–4245. doi: 10.1109/IROS.2018.8593986

Zeng, A., Song, S., Yu, K.-T., Donlon, E., Hogan, F. R., Bauza, M., et al. (2018b). "Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD: IEEE), 1–8. doi: 10.1109/ICRA.2018.8461044

Zeng, A., Yu, K.-T., Song, S., Suo, D., Walker, E., Rodriguez, A., et al. (2017). "Multi-view self-supervised deep learning for 6d pose estimation in the amazon picking challenge," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 1386–1383. doi: 10.1109/ICRA.2017.7989165

Zhang, F., and Demiris, Y. (2020). "Learning grasping points for garment manipulation in robot-assisted dressing," in *2020 IEEE International Conference on Robotics and Automation (ICRA)* (Paris: IEEE), 9114–9120. doi: 10.1109/ICRA40945.2020.9196994

Zhang, H., Ichnowski, J., Avigal, Y., Gonzales, J., Stoica, I., and Goldberg, K. (2020). "Dex-Net AR: distributed deep grasp planning using an augmented reality application and a smartphone camera," in *IEEE International Conference on Robotics and Automation (ICRA)* (Paris: IEEE). doi: 10.1109/ICRA40945.2020.9197247

Zhang, Q., Qu, D., Xu, F., and Zou, F. (2017). "Robust robot grasp detection in multimodal fusion," in *MATEC Web of Conferences: EDP Sciences* (Chengdu), 00060. doi: 10.1051/matecconf/201713900060

Zhao, B., Zhang, H., Lan, X., Wang, H., Tian, Z., and Zheng, N. (2020). REGNet: REgion-based grasp network for single-shot grasp detection in point clouds. *arXiv [Preprint]*. arXiv:2002.12647.

Zhao, W., Queralta, J. P., and Westerlund, T. (2020). Sim-to-real transfer in deep reinforcement learning for robotics: a survey. *arXiv [Preprint]*. arXiv:2009.13303. doi: 10.1109/SSCI47803.2020.9308468

Zhou, Y., Chen, M., Du, G., Zhang, P., and Liu, X. (2018). Intelligent grasping with natural human-robot interaction. *Indus. Robot* 45, 44–53. doi: 10.1108/IR-05-2017-0089

Zhou, Y., and Hauser, K. (2017). "6dof grasp planning by optimizing a deep learning scoring function," in *Robotics: Science and Systems (RSS) Workshop on Revisiting Contact-Turning a Problem into a Solution* (Cambridge, MA), 6.

Zhou, Z., Pan, T., Wu, S., Chang, H., and Jenkins, O. C. (2019). Glassloc: plenoptic grasp pose detection in transparent clutter. *arXiv [Preprint]*. arXiv:1909.04269. doi: 10.1109/IROS40897.2019.8967685

Zhou, Z., Sui, Z., and Jenkins, O. C. (2018). "Plenoptic monte carlo object localization for robot grasping under layered translucency," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid: IEEE), 1–8. doi: 10.1109/IROS.2018.8593629

Zhu, S., Zheng, X., Xu, M., Zeng, Z., and Zhang, H. (2019). "A robotic semantic grasping method for pick-and-place tasks," in *2019 Chinese Automation Congress (CAC)* (Hangzhou: IEEE), 4130–4136. doi: 10.1109/CAC48633.2019.8996328

Zhu, Y., Fathi, A., and Fei-Fei, L. (2014). "Reasoning about object affordances in a knowledge base representation," in *European Conference on Computer Vision*, eds D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Zurich: Springer), 408–424. doi: 10.1007/978-3-319-10605-2_27

Zhu, Y., Zhao, Y., and Chun Zhu, S. (2015). "Understanding tools: task-oriented object modeling, learning and recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA), 2855–2864. doi: 10.1109/CVPR.2015.7298903

Zito, C., Ortenzi, V., Adjigble, M., Kopicki, M., Stolkin, R., and Wyatt, J. L. (2019). Hypothesis-based belief planning for dexterous grasping. *arXiv [Preprint]*. arXiv:1903.05517.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership