



A Method for Measuring Contact Points in Human–Object Interaction Utilizing Infrared Cameras

Jussi Hakala and Jukka Häkkinen *

Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

OPEN ACCESS

Edited by:

Antonios Gasteratos,
Democritus University of Thrace,
Greece

Reviewed by:

Kai Wang,
China United Network
Communications Group, China
Alexander Rassau,
Edith Cowan University, Australia
Ruediger Dillmann,
Karlsruhe Institute of Technology (KIT),
Germany

*Correspondence:

Jukka Häkkinen
jukka.hakkinen@helsinki.fi

Specialty section:

This article was submitted to
Robot and Machine Vision,
a section of the journal
Frontiers in Robotics and AI

Received: 26 October 2021

Accepted: 14 December 2021

Published: 14 February 2022

Citation:

Hakala J and Häkkinen J (2022) A
Method for Measuring Contact Points
in Human–Object Interaction Utilizing
Infrared Cameras.
Front. Robot. AI 8:800131.
doi: 10.3389/frobt.2021.800131

This article presents a novel method for measuring contact points in human–object interaction. Research in multiple prehension-related fields, e.g., action planning, affordance, motor function, ergonomics, and robotic grasping, benefits from accurate and precise measurements of contact points between a subject’s hands and objects. During interaction, the subject’s hands occlude the contact points, which poses a major challenge for direct optical measurement methods. Our method solves the occlusion problem by exploiting thermal energy transfer from the subject’s hand to the object surface during interaction. After the interaction, we measure the heat emitted by the object surface with four high-resolution infrared cameras surrounding the object. A computer-vision algorithm detects the areas in the infrared images where the subject’s fingers have touched the object. A structured light 3D scanner produces a point cloud of the scene, which enables the localization of the object in relation to the infrared cameras. We then use the localization result to project the detected contact points from the infrared camera images to the surface of the 3D model of the object. Data collection with this method is fast, unobtrusive, contactless, markerless, and automated. The method enables accurate measurement of contact points in non-trivially complex objects. Furthermore, the method is extendable to measuring surface contact areas, or patches, instead of contact points. In this article, we present the method and sample grasp measurement results with publicly available objects.

Keywords: grasping, touch, infrared camera, contact point, prehension movements

INTRODUCTION

Our hands are excellent tools for manipulating objects, and we touch and grasp countless objects every day. Prehension movements are divided into three components: moving the hand to the target, setting finger posture for grasping, and aligning the hand so that grasping is possible (Fan et al., 2006). They enable us to select contact points in the object that allow a stable grip and object manipulation (Zatsiorsky and Latash, 2008; Kleinholdermann et al., 2013).

Prehension movements are guided the visual information about the size, shape and density of the object (Cesari and Newell, 1999). This information is used in preliminary planning of hand movement and grasp properties, but also changes that occur during the hand movement are also taken into account (Bridgeman et al., 1979).

While the operations are easy and effortless for humans, they are not easy for robots. Grasping and manipulating previously unseen objects and operating in unstructured, cluttered, and variable environments has proven to be a very difficult task for robots (Cui and Trinkle, 2021). Although great

efforts have been focused on the problem in recent years, it has not been solved (Kleeberger et al., 2020).

There are several ways to approach the problem. One approach is called analytic, in which the grasp is determined by analyzing the shape of the object. This could be done, for example, by analyzing the appearance of objects to determine their grasping affordances (Song et al., 2015; Do et al., 2018).

A second approach is called data-driven, as a dataset is formed to train the robot hand with suitable method, such as deep learning. Examples of the data-driven approach are, for example, a study by Gabellieri et al. (2020), which describe human grasp data collection, where a human operator uses the robot hand to grasp objects and resulting the grasp data is collected with depth cameras and motion tracker. Edmonds et al. (2019) used tactile gloves with force sensors to capture the poses and forces that humans used when opening medicine bottles with safety mechanisms. The data was used to teach a robotic system do conduct the same task.

As grasp data collection is time consuming and expensive, the data has been collected in to databases such as the Columbia Grasp Database, which is based on the GraspIt! Toolkit, and can be used by the research community to develop suitable training methods (Goldfeder et al., 2009).

The critical issue with the data-driven approach is the quality of data that is collected. Human hand trajectory and grasp can be measured in many ways. For example, the interaction can be video-recorded and the interactions manually coded from videos (Cesari and Newell, 1999; Choi and Mark, 2004). However, this is very time consuming and can lead to issues with inter-coder reliability (Kita et al., 1998).

Instead of manual annotation, a computer vision system can detect and track hands from RGB images (Siddharth et al., 2016; Cai et al., 2017; Lyubanenko et al., 2017) or from depth camera images (Oikonomidis et al., 2011). Many of these models perform well in laboratory conditions, but have problems with unconstrained real-life conditions, as large variations in the scenery makes feature extraction difficult (Yang et al., 2015). Attaching visual markers to the hand and tracking the markers (Gentilucci et al., 1992) improves the tracking result, at the cost of making the experimental setup more complex and less natural. However, the most critical problem is hand occlusion, which makes the detection of contact points difficult (Yang et al., 2015; Sridhar et al., 2016) for both manual and computer vision-based methods.

Another category of methods relies on sensors attached to the hand. The sensors in these systems can be electromagnetic (trakSTAR, Ascension Technology Corp., Shelburne, VT, United States; e.g., Chen and Saunders (2018), FASTRAK, Polhemus Corp., Colchester, VT, United States; e.g., Bock and Jüngling (1999), resistive (CyberGlove, Virtual Technologies, Palo Alto, CA, United States; e.g., Ansuini et al., 2007), or infrared (Optotrak, Northern Digital Inc., Waterloo, Ontario, Canada; e.g., Kleinholdermann et al. (2013) and Cressman et al. (2007). Sensors provide accurate hand and joint motion data, but they make the experimental setup less natural and might disrupt hand-object interactions, as both sensors and the related wires may limit movements and change grasping strategies.



FIGURE 1 | The measuring station with a target object.

As all the current methods have limitations, we introduce a contact point measuring method to overcome these limitations. The criteria during the development work were that the method 1) allows for fast data collection without manual annotation; 2) can measure everyday objects with high ecological validity, i.e., without altering their visual appearance or interfering with interaction; and 3) produces accurate data.

Our method is based on thermal imaging, which has been widely used in medicine (Gizińska et al., 2021; Martinez-Jimenez et al., 2021), surveillance (Ivašić-Kos et al., 2019; Duan et al., 2021), and quality control (ElMasry et al., 2020; Khera et al., 2020). The starting point of our application is the temperature of fingers, which is $29.1 \pm 0.6^\circ\text{C}$ at the environmental temperature of $25.4 \pm 0.4^\circ\text{C}$, (Shilco et al., 2019), and which transfers to objects that have been touched. In other words, when a participant touches an object and takes their hand away, there is a short-term heat signature corresponding to the area touched. This heat signature can be used to deduce the grasp type used during the human-object interaction. Furthermore, the heat signature can be used to differentiate between grasp types that look visually similar when the hand is viewed from above, but are actually different as the details of the grip are occluded by the hand.

In the following article we will outline the technical details of the method as well as presenting preliminary data recorded with the system while grasping the YCB object set (Calli et al., 2015).

METHODS

The measuring station (**Figure 1**) consists of four FLIR A65 infrared cameras (FLIR Systems, Wilsonville, OR, United States), one PhoXi 3D Scanner M structured light scanner (Photoneo s.r.o., Bratislava, Slovakia), a control computer, and a frame structure with a flat surface in the middle for placing the target object. For future development, we also included four

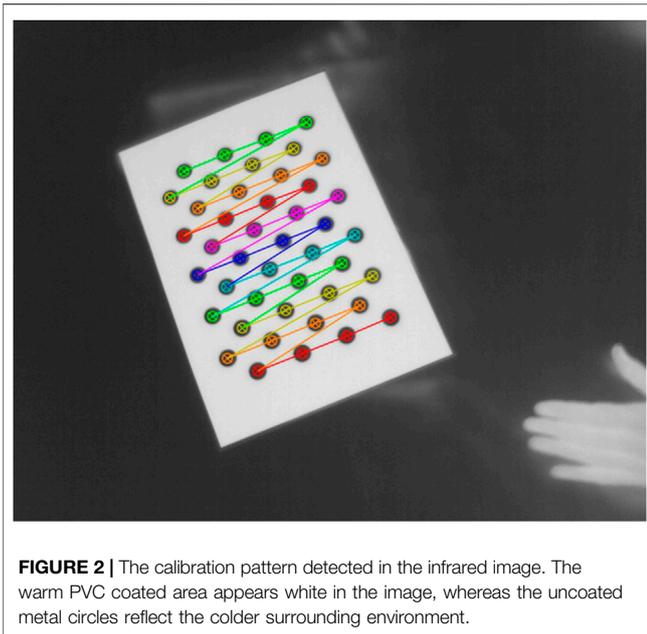


FIGURE 2 | The calibration pattern detected in the infrared image. The warm PVC coated area appears white in the image, whereas the uncoated metal circles reflect the colder surrounding environment.

Xbox One Kinect Sensors (Microsoft Corporation, Redmond, WA, United States) in the configuration.

Calibration of the camera system requires a target where the pattern features are detectable in images recorded with all the cameras in the system. The system consists of four types of cameras that record different bands of the spectrum. The infrared cameras record wavelengths 7.5–13.0 μm (Flir systems, 2016). The 3D scanner utilizes a 638 nm laser for the structured light projection (Photoneo, 2018). We also record the interaction with a Kinect RGB-D camera that uses laser diodes with peak intensity at 850 nm for time-of-flight measurement (Naeemabadi et al., 2018) and separate sensors for the visible part of the spectrum and for near-infrared. To calibrate all cameras internally and externally, we developed a calibration target (Figure 2). The pattern is a 4 by 11 asymmetric circle grid with 20.0 mm spacing and 15.0 mm circle diameter. The pattern was cut by laser from a 70 μm thick matte black PVC film and attached to a 4.0 mm thick unpolished aluminum alloy sheet. We warmed the calibration target with a heating element so that the temperature difference with room temperature was approximately 10°C. The uncoated aluminum circles reflect the surrounding environment in the wavelengths recorded by the devices, whereas the PVC-coated surface is less reflective. The calibration target enables us to utilize the camera calibration implementation in OpenCV (<https://opencv.org/>). We calibrated the camera internal parameters and the camera system externally, so that we obtained a rotation and translation relative to the 3D scanner for each camera.

Data acquisition takes place during and after the interaction. We record the scene with the infrared cameras and one RGB-D camera during the interaction and a short period after the interaction has ended. At the end of the interaction, the subject places the object on the table covered by the cameras and pulls their hand out of the line of sight between the object and

the cameras. Then, the recording ends and we scan the scene with the 3D scanner. Figure 3 illustrates the process flow. Here, we describe the process for a single object, but the method allows for efficient data gathering from multiple objects consecutively. The experiment supervisor needs to take care to place each object in the desired starting location without altering its surface temperature, for example, by wearing insulating gloves or using room-temperature tongs.

To detect the contact points in the infrared images recorded after the interaction, we used the simple blob detection algorithm implemented in OpenCV. The algorithm detects the regions that differ in temperature from the ambient temperature, and outputs the center point coordinates u and v , and radius r of each region. The camera parameters used and the scene configuration determine the required parameters for the algorithm.

For localization, we used the proprietary Photoneo Localization SDK (Photoneo s.r.o., Bratislava, Slovakia). The localization algorithm uses a set of features calculated from a predefined 3D model of the object and searches for matching features in the recorded point cloud. As a result of the localization, we obtain the object rotation and transformation matrices, R and T , relative to the 3D scanner.

To map the contact points on the object surface, we project the contact points from the image coordinates back to the 3D scene, where the rays projected from the camera intersect with the surface of the object 3D model. To solve the visibility of the points and self-occlusions within the object, we use OpenGL (Khronos Group, Beaverton, OR, United States) to obtain the Z-buffer of the object model from each infrared camera location. We transform the object model from the 3D scanner reference frame to the IR camera reference frame and render it. Thus, the resulting Z-buffer provides the Z-depth of the key points in the infrared image. Points outside the object boundary are discarded. The key point coordinates are then transformed to the object reference frame.

As the same contact point may be visible in multiple camera images, we filter the set of contact points. As a measure of point reliability, we use the angle between the camera optical axis and the surface normal at the contact point, because the imperfections in key point detection, calibration, and localization cause an error in projection that increases with the incident angle. Out of the key points from different cameras that are within a 9 mm distance of each other, we retain the point with the smallest incident angle. Finally, we store the remaining key point coordinates X , Y and Z , and radii r , in millimeters in object coordinates, within the database.

We followed the guidelines given by the Finnish Advisory Board on Research Integrity and the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences.

RESULTS

We recorded one grasping event for each of the 77 objects in the YCB object set (Calli et al., 2015; Calli et al., 2017) for which there was a matching 3D model available. Author JHH grasped each

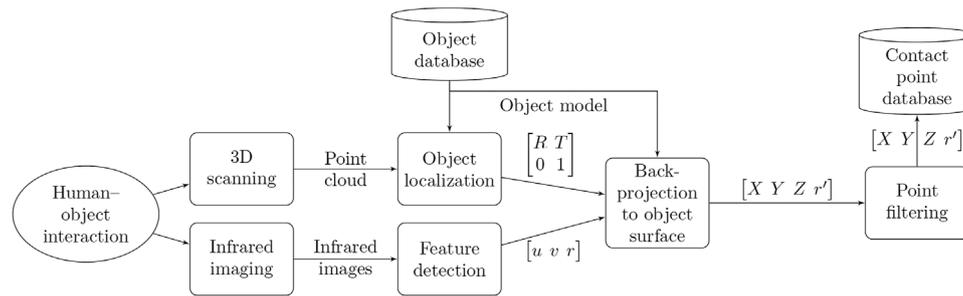


FIGURE 3 | The stages of measuring contact points. After human-object interaction, we record infrared images of the scene and a 3D point cloud. We use the point cloud to localize the object in the scene accurately and a computer-vision algorithm detects the contact points in the infrared images. Next, we project the contact points to the surface of the localized object model, filter, and store the point coordinates and radii.

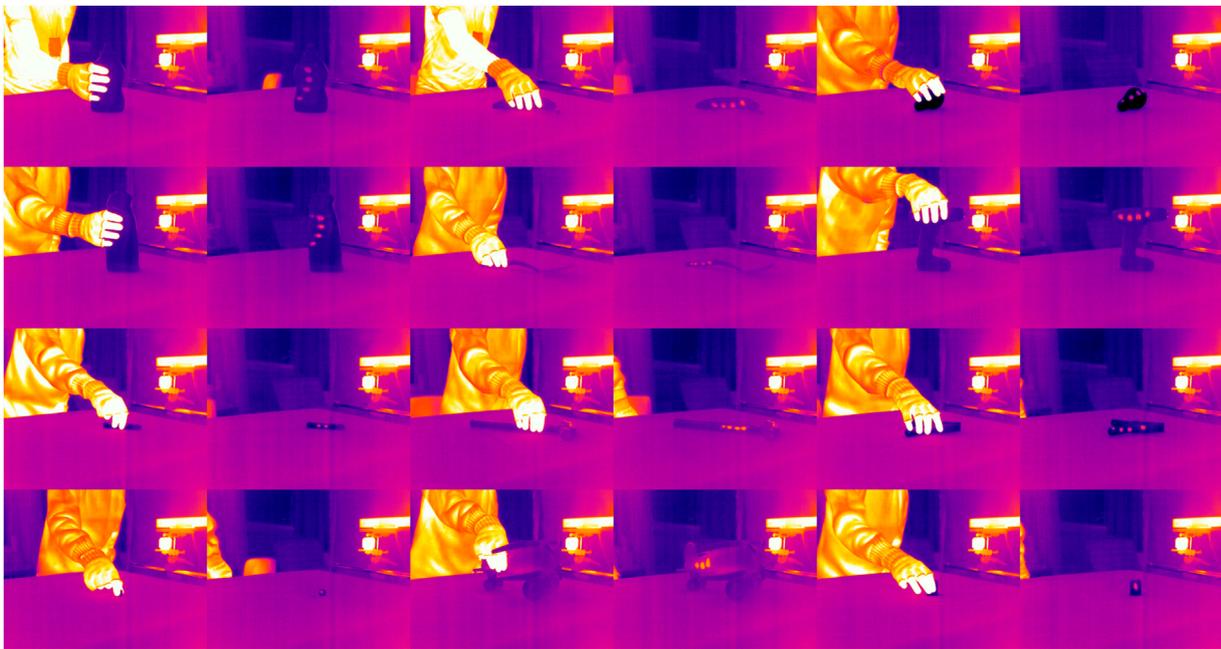


FIGURE 4 | Twelve sample infrared image pairs recorded during the grasping act and after the interaction.

object with his right hand using a natural grasp suitable for lifting the object (**Figure 4**), lifted the object from the table to 10–20 cm height, and set it back down. The grasp type or duration was not restricted; contact duration was often less than 2 s. To limit the skin–object contact areas to the distal phalanx of each finger, the grasper wore an insulating fingerless glove.

Localization succeeded for 67 out of the 77 objects. Localization was unsuccessful possibly due to partially deformed objects (004_sugar_box and 061_foam_brick), non-rigidity of the object (059_chain and 063-a_marbles), and highly reflective surfaces (038_padlock). The localization result of 019_pitcher_base was 180° rotated and thus classified as a failure.

Determining the ground truth for measuring the performance of the method proved challenging, as there is no automated method that could provide the ground truth. We determined the

success of the contact point detection by visually comparing the detected contact points to the IR and RGB camera frames from the grasping event. Thus, our “ground truth” is prone to human error, but we provide the grasping event recordings along with our code¹, so others may confirm our results. Detection of some contact points succeeded in 66 out of the 67 localized objects. **Figures 5, 6** show examples of the detections. Only the highly reflective object 042_adjustable_wrench did not yield any true positive detections. For 62 objects, the method detected all contact points successfully. False positive detections occurred in 16 objects, mainly due to reflections in high-reflectance objects such as 028_skillet_lid, as evident in **Figure 7**. The insulating

¹<https://github.com/jussihh/graspense>

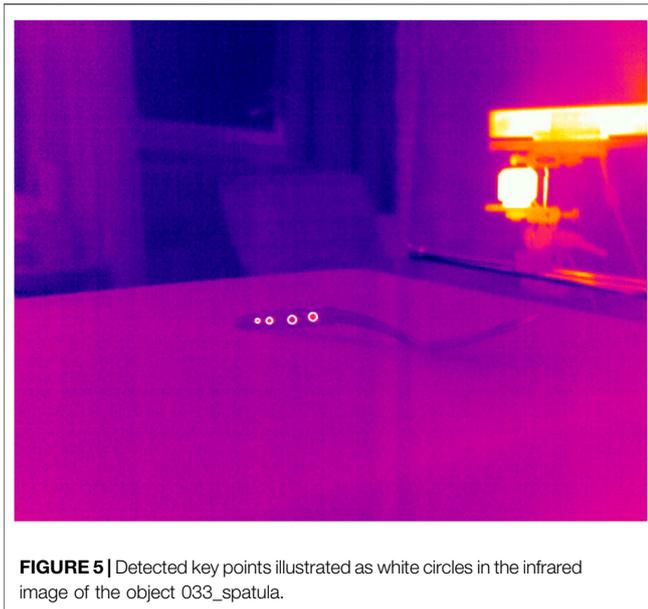


FIGURE 5 | Detected key points illustrated as white circles in the infrared image of the object 033_spatula.

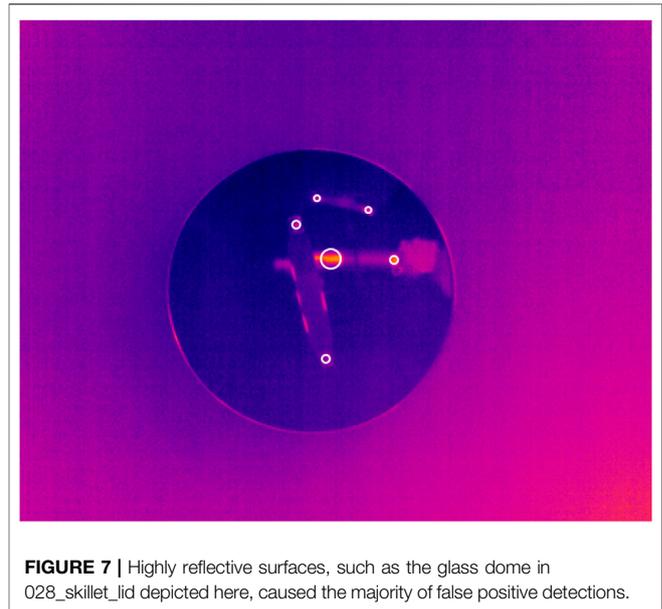


FIGURE 7 | Highly reflective surfaces, such as the glass dome in 028_skillet_lid depicted here, caused the majority of false positive detections.

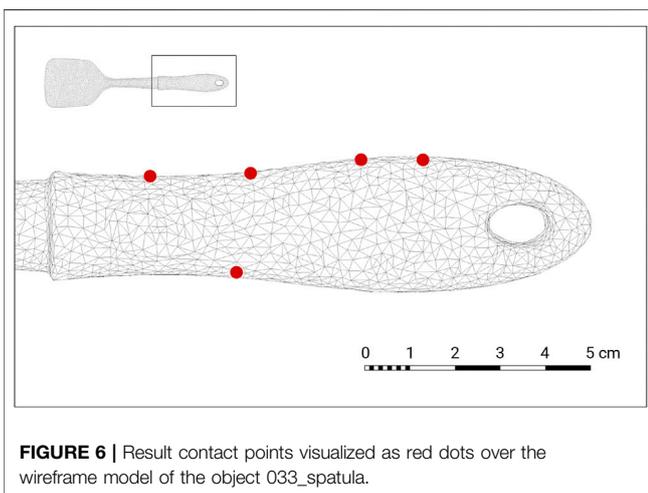


FIGURE 6 | Result contact points visualized as red dots over the wireframe model of the object 033_spatula.

glove covered the second phalanx of the grasper's fingers only partially, which resulted in two contact points from the same finger in three cases. Some contact points went undetected because the surface was not covered by the cameras, such as the underside of the object 029_plate. We computed precision and recall for all finger contacts and used visual inspection of the recordings as the ground truth. For the 67 localized objects, the precision of the system was 0.90 whereas recall was 0.95.

DISCUSSION

In this article, we have described a novel method for measuring contact points in human–object interaction and presented the first test measurement results. We measured grasping with the YCB object set, which includes everyday objects with different sizes and materials. For most of the objects, we were able to

measure contact points successfully. For the 22 plastic toy objects in groups 065, 072, and 073 precision was 0.97 and recall 1.00, which indicates that in applications, where the experimenter can control the object material, the method meets demanding requirements. Highly reflective objects proved difficult for both the 3D scanner and the infrared cameras to measure.

The spatial accuracy of the process is challenging to measure, because the error depends on the properties of the environment, the object, the grasping hand, and grasp type. We could measure the error using an ideal object with a known geometry and temperature difference as a target, such as our calibration target. However, such a measure, likely in fractions of a millimeter, would be highly misleading as it discounts the major sources of error. Thermal conduction over time within the object should be taken into account when considering the spatial accuracy. We are developing a method to measure the spatial accuracy of the entire process; from our experience, we expect the root mean square error to be within a few millimeters with our hardware and software configuration.

The method is efficient and accurate for measuring contact points. For some applications, a more realistic model of the contact area might be more desirable than only the contact area centers and radii that our method currently measures. For example, in robotics, a patch contact model improves grasping compared with a point model (Charusta et al., 2012). Our method is also extendable to measuring the contact areas, by detecting the contact areas in the infrared camera images and mapping them to the model surface. For contact area measurement, the temperature gradient caused by thermal conduction within the object must be included in the model, as the contact areas cool down over time after the contact. This could be achieved by tracking the object and recording the contact duration and surface area temperature over time. However, as the fingers of the grasper occlude the contact area, the exact duration of contact is challenging to measure from unaltered objects. Including

the object surface's material properties and three-dimensional object structure in the model would be useful.

Another improvement we are developing is including a hand and finger tracking method. Tracking would allow us to identify fingers, collect motion trajectory data, and determine the grasp type, which is valuable in some applications (such as grasp planning, e.g., in GraspIt! (Miller and Allen, 2004). To increase the usefulness and improve the performance of the method, prior knowledge about the grasp type and the number of contact points used should be included in data processing, as different grasp types produce different skin-object contact areas. In addition to the contact points and patches, some applications may require additional data. Using, e.g., finite element models for the deformation of fingers might provide useful. Furthermore, measuring the force distribution and friction between the contact surfaces would be very useful in robotics applications. Combining the models for thermal conduction and finger deformation would allow us to approximate the grasp forces. The weights and dimensions of the objects, which might also be useful, are listed in the YCB dataset description (Calli et al., 2015).

In automation and robotics, the method enables the recording of grasping demonstrations with known objects in applications such as warehouse or assembly line automation, if the number of target objects makes this feasible. The method also enables the collection of the vast datasets required for training machine-learning models and the subsequent development of robots that are able to operate in unstructured environments such as households. For example, action-specific contact point and point cloud data recorded from a large number of objects enables the training of a deep neural network to predict contact points in novel objects, a task where manually annotated images are still used (Chu et al., 2018).

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. Infrared and RGB recordings of grasping YCB objects are available at doi.org/10.5281/zenodo.5783717.

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

JHä conceived the initial idea for the measurement system, acquired funding for the project, and supervised it. JHa played a major role in the system planning phase, build the system, and programmed the software. JHa measured and analysed the touch data reported in the article. Both authors contributed to the article and approved the submitted version.

FUNDING

The project was funded by Business Finland project 4331/31/2016 "Automatic measurement of human touch."

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frobt.2021.800131/full#supplementary-material> Supplementary Material: Contains sample data recorded from the YCB object set with the method described in this article. See the included readme.txt for detailed description of the data.

REFERENCES

- Ansuini, C., Santello, M., Tubaldi, F., Massaccesi, S., and Castiello, U. (2007). Control of Hand Shaping in Response to Object Shape Perturbation. *Exp. Brain Res.* 180, 85–96. doi:10.1007/s00221-006-0840-9
- Bock, O., and Jüngling, S. (1999). Reprogramming of Grip Aperture in a Double-Step Virtual Grasping Paradigm. *Exp. Brain Res.* 125, 61–66. doi:10.1007/s002210050658
- Bridgeman, B., Lewis, S., Heit, G., and Nagle, M. (1979). Relation between Cognitive and Motor-Oriented Systems of Visual Position Perception. *J. Exp. Psychol. Hum. Perception Perform.* 5, 692–700. doi:10.1037/0096-1523.5.4.692
- Cai, M., Kitani, K. M., and Sato, Y. (2017). An Ego-Vision System for Hand Grasp Analysis. *IEEE Trans. Human-Mach. Syst.* 47, 524–535. doi:10.1109/THMS.2017.2681423
- 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. Robotics Res.* 36, 261–268. doi:10.1177/0278364917700714
- Calli, B., Singh, A., Walsman, A., Srinivasa, S., Abbeel, P., and Dollar, A. M. (2015). The YCB Object and Model Set: Towards Common Benchmarks for Manipulation Research. International Conference on Advanced Robotics (ICAR), Istanbul, Turkey, July 27–31, 2015–517. doi:10.1109/ICAR.2015.7251504
- Cesari, P., and Newell, K. M. (1999). The Scaling of Human Grip Configurations. *J. Exp. Psychol. Hum. Perception Perform.* 25, 927–935. doi:10.1037/0096-1523.25.4.927
- Charusta, K., Krug, R., Dimitrov, D., and Iliev, B. (2012). Independent Contact Regions Based on a Patch Contact Model. IEEE Int. Conf. Robot Autom., Paul, Minnesota, May 14–15, 2012, 4162–4169. doi:10.1109/ICRA.2012.6225325
- Chen, Z., and Saunders, J. A. (2018). Volitional and Automatic Control of the Hand when Reaching to Grasp Objects. *J. Exp. Psychol. Hum. Perception Perform.* 44, 953–972. doi:10.1037/xhp0000492
- Choi, H. J., and Mark, L. S. (2004). Scaling Affordances for Human Reach Actions. *Hum. Mov. Sci.* 23, 785–806. doi:10.1016/j.humov.2004.08.004
- Chu, F.-J., Xu, R., and Vela, P. A. (2018). Real-World Multiobject, Multigrasp Detection. *IEEE Robot. Autom. Lett.* 3, 3355–3362. doi:10.1109/LRA.2018.2852777
- Cressman, E. K., Franks, I. M., Enns, J. T., and Chua, R. (2007). On-Line Control of Pointing Is Modified by Unseen Visual Shapes. *Conscious. Cogn.* 16, 265–275. doi:10.1016/j.concog.2006.06.003
- Cui, J., and Trinkle, J. (2021). Toward Next-Generation Learned Robot Manipulation. *Sci. Robot* 6, eabd9461. doi:10.1126/scirobotics.abd9461

- 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, Brisbane, Australia, May 21–26, 2018 (ICRA), 5882–5889. doi:10.1109/ICRA.2018.8460902
- Duan, W., Wang, Y., Li, J., Zheng, Y., Ning, C., and Duan, P. (2021). Real-Time Surveillance-Video-Based Personalized thermal comfort Recognition. *Energy and Buildings* 244, 110989. doi:10.1016/j.enbuild.2021.110989
- Edmonds, M., Gao, F., Liu, H., Xie, X., Qi, S., Rothrock, B., et al. (2019). A Tale of Two Explanations: Enhancing Human Trust by Explaining Robot Behavior. *Sci. Robot* 4, eaay4663. doi:10.1126/scirobotics.aay4663
- ElMasry, G., ElGamal, R., Mandour, N., Gou, P., Al-Rejaie, S., Belin, E., et al. (2020). Emerging thermal Imaging Techniques for Seed Quality Evaluation: Principles and Applications. *Food Res. Int.* 131, 109025. doi:10.1016/j.foodres.2020.109025
- Fan, J., He, J., and Tillery, S. I. H. (2006). Control of Hand Orientation and Arm Movement during Reach and Grasp. *Exp. Brain Res.* 171, 283–296. doi:10.1007/s00221-005-0277-6
- Flir Systems (2016). *FLIR A65 FOV 45 (30 Hz, Ver. 2016) Product Documentation*. Publ. No. 75013-0101, Commit 36999. Teledyne FLIR LLC.
- Gabellieri, C., Angelini, F., Arapi, V., Palleschi, A., Catalano, M. G., Grioli, G., et al. (2020). Grasp it Like a Pro: Grasp of Unknown Objects with Robotic Hands Based on Skilled Human Expertise. *IEEE Robot. Autom. Lett.* 5 (2), 2808–2815. doi:10.1109/lra.2020.2974391
- Gentilucci, M., Chieffi, S., Scarpa, M., and Castiello, U. (1992). Temporal Coupling between Transport and Grasp Components during Prehension Movements: Effects of Visual Perturbation. *Behav. Brain Res.* 47, 71–82. doi:10.1016/s0166-4328(05)80253-0
- Gizińska, M., Rutkowski, R., Szymczak-Bartz, L., Romanowski, W., and Straburzyńska-Lupa, A. (2021). Thermal Imaging for Detecting Temperature Changes within the Rheumatoid Foot. *J. Therm. Anal. Calorim.* 145, 77–85. doi:10.1007/s10973-020-09665-0
- Goldfeder, C., Ciocarlie, M., Hao Dang, H., and Allen, P. K. (2009). "The Columbia Grasp Database," in 2009 IEEE international conference on robotics and automation, Japan, May 12–17, 2009 (IEEE), 1710–1716. doi:10.1109/ROBOT.2009.5152709
- Ivašić-Kos, M., Krišto, M., and Pobar, M. (2019). "Human Detection in thermal Imaging Using YOLO," in Proceedings of the 2019 5th International Conference on Computer and Technology Applications, Istanbul Turkey, April 16–17, 2019 (Association for Computing Machinery), 20–24. doi:10.1145/3323933.3324076
- Khera, N., Khan, S. A., and Rahman, O. (2020). Valve Regulated lead Acid Battery Diagnostic System Based on Infrared thermal Imaging and Fuzzy Algorithm. *Int. J. Syst. Assur. Eng. Manag.* 11, 614–624. doi:10.1007/s13198-020-00958-z
- Kita, S., van Gijn, I., and van der Hulst, H. (1998). Movement Phases in Signs and Co-Speech Gestures, and Their Transcription by Human Coders. *Lect. Notes Comput. Sci.* 1371, 23–35. doi:10.1007/BFb0052986
- 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
- Kleinholdermann, U., Franz, V. H., and Gegenfurtner, K. R. (2013). Human Grasp point Selection. *J. Vis.* 13, 23. doi:10.1167/13.8.23
- Lyubanenko, V., Kuronen, T., Eerola, T., Lensu, L., Kälviäinen, H., and Häkkinen, J. (2017). "Multi-camera Finger Tracking and 3D Trajectory Reconstruction for HCI Studies," in International Conference on Advanced Concepts for Intelligent Vision Systems ACIVS 2017: Advanced Concepts for Intelligent Vision Systems, Antwerp, Belgium, September 18–21, 2017. Editors J. Blanc-Talon, R. Penne, W. Philips, D. Popescu, and P. Scheunders. (Cham, Switzerland: Springer), 63–74. doi:10.1016/S0097-8493(02)00077-8
- Martinez-Jimenez, M. A., Loza-Gonzalez, V. M., Kolosovas-Machuca, E. S., Yanes-Lane, M. E., Ramirez-GarciaLuna, A. S., and Ramirez-GarciaLuna, J. L. (2021). Diagnostic Accuracy of Infrared thermal Imaging for Detecting COVID-19 Infection in Minimally Symptomatic Patients. *Eur. J. Clin. Invest.* 51, e13474. doi:10.1111/eci.13474
- Miller, A. T., and Allen, P. K. (2004). GraspIt!. *IEEE Robot. Automat. Mag.* 11 (4), 110–122. doi:10.1109/MRA.2004.1371616
- Naemabadi, M., Dinesen, B., Andersen, O. K., and Hansen, J. (2018). Investigating the Impact of a Motion Capture System on Microsoft Kinect V2 Recordings: A Caution for Using the Technologies Together. *PLoS One* 13, e0204052. doi:10.1371/journal.pone.0204052
- Oikonomidis, I., Kyriazis, N., and Argyros, A. (2011). "Efficient Model-Based 3D Tracking of Hand Articulations Using Kinect," in Proceedings of the 2011 British Machine Vision Conference, DundeeUK, 29 August–2 September 2011 (Durham BMVA Press 2011), 1–11. doi:10.5244/c.25.101
- Photoneo (2018). *PhoXi® 3D Scanner User Manual and Installation Instructions 01/2018, Rev 3*. Bratislava, Slovakia: Photoneo s.r.o..
- Shilco, P., Roitblat, Y., Buchris, N., Hanai, J., Cohensedgh, S., Frig-Levinson, E., et al. (2019). Normative Surface Skin Temperature Changes Due to Blood Redistribution: A Prospective Study. *J. Therm. Biol.* 80, 82–88. doi:10.1016/j.jtherbio.2019.01.009
- SiddharthRangesh, A., Ohn-Bar, E., and Trivedi, M. M. (2016). "Driver Hand Localization and Grasp Analysis: A Vision-Based Real-Time Approach," in IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Brazil, November 1–4, 2016 (IEEE), 2545–2550. doi:10.1109/ITSC.2016.7795965
- Song, H. O., Fritz, M., Goehring, D., and Darrell, T. (2016). Learning to Detect Visual Grasp Affordance. *IEEE Trans. Automat. Sci. Eng.* 13, 798–809. doi:10.1109/TASE.2015.2396014
- Sridhar, S., Mueller, F., Zollhöfer, M., Casas, D., Oulasvirta, A., and Theobalt, C. (2016). "Real-Time Joint Tracking of a Hand Manipulating an Object from RGB-D Input," *Lecture Notes in Computer Science*. B. Leibe, J. Matas, N. Sebe, and M. Welling (Cham: Springer), 9906, 294–310. doi:10.1007/978-3-319-46475-6_19
- Yang, Y., Fermuller, C., Li, Y., and Aloimonos, Y. (2015). "Grasp Type Revisited: A Modern Perspective on a Classical Feature for Vision," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit, Boston, Massachusetts, May 8–10, 2015 (IEEE), 400–408. doi:10.1109/CVPR.2015.7298637
- Zatsiorsky, V. M., and Latash, M. L. (2008). Multifinger Prehension: An Overview. *J. Mot. Behav.* 40, 446–476. doi:10.3200/JMBR.40.5.446-476

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hakala and Häkkinen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.