



Automatic Measurement of Head and Facial Movement for Analysis and Detection of Infants' Positive and Negative Affect

Zakia Hammal^{1*}, Jeffrey F. Cohn^{1,2}, Carrie Heike³ and Matthew L. Speltz⁴

¹Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, ²Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA, ³Seattle Children's Hospital, Seattle, WA, USA, ⁴University of Washington School of Medicine, Seattle, WA, USA

OPEN ACCESS

Edited by:

Javier Jaen,
Universitat Politècnica de València,
Spain

Reviewed by:

Andrej Košir,
University of Ljubljana, Slovenia
Guillaume Chanel,
University of Geneva, Switzerland

*Correspondence:

Zakia Hammal
zakia_hammal@yahoo.fr

Specialty section:

This article was submitted to
Human-Media Interaction,
a section of the
journal *Frontiers in ICT*

Received: 23 June 2015

Accepted: 08 October 2015

Published: 02 December 2015

Citation:

Hammal Z, Cohn JF, Heike C and
Speltz ML (2015) Automatic
Measurement of Head and Facial
Movement for Analysis and Detection
of Infants' Positive and Negative
Affect.
Front. ICT 2:21.
doi: 10.3389/fict.2015.00021

Previous work in automatic affect analysis (AAA) has emphasized static expressions to the neglect of the dynamics of facial movement and considered head movement only a nuisance variable to control. We investigated whether the dynamics of head and facial movements apart from specific facial expressions communicate affect in infants, an under-studied population in AAA. Age-appropriate tasks were used to elicit positive and negative affect in 31 ethnically diverse infants. 3D head and facial movements were tracked from 2D video. Head angles in the horizontal (pitch), vertical (yaw), and lateral (roll) directions were used to measure head movement; and the 3D coordinates of 49 facial points to measure facial movements. Strong effects were found for both head and facial movements. Angular velocity and angular acceleration of head pitch, yaw, and roll were higher during negative relative to positive affect. Amplitude, velocity, and acceleration of facial movement were higher as well during negative relative to positive affect. A linear discriminant analysis using head and facial movement achieved a mean classification rate of positive and negative affect equal to 65% ($Kappa = 0.30$). Head and facial movements individually and in combination were also strongly related to observer ratings of affect intensity. Our results suggest that the dynamics of head and facial movements communicate affect at ages as young as 13 months. These interdisciplinary findings from behavioral science and computer vision deepen our understanding of communication of affect and provide a basis for studying individual differences in emotion in socio-emotional development.

Keywords: head movement, facial movement, positive and negative affect, infants

INTRODUCTION

Within the past 5 years, there have been dramatic advances in automatic affect analysis (AAA). From person-specific feature detection and recognition of facial expressions in posed behavior in controlled settings (Zeng et al., 2012), AAA has progressed to person-independent feature detection and recognition of facial expression in spontaneous behavior in diverse settings (Valstar et al., 2013, 2015; Sariyanidi et al., 2015). These include therapy interviews, psychology research, medical settings, and webcam recordings in homes (Cohn and De la Torre, 2015). Great strides have been made

in detection of both holistic expressions and anatomically based facial action units (AUs) in a range of applications.

In almost all work to date, the focus has been on detection or recognition of static expressions. This emphasis on static expressions has its origins in the rich descriptions of Sir Charles Bell (1844), Duchenne (1990), and Darwin in the nineteenth century (Darwin, 1872). Dependent on illustrations and photographic plates, lacking the means to quantify facial expression on a time basis, and informed by theoretical models that emphasized static expressions, they naturally considered static representations paramount. Later work in "basic emotions" by Ekman (1972), Ekman et al. (2002), and Izard (1971); Izard (1994) in the twentieth century furthered this focus. The advent of video technology and advances in image and video processing together with Ekman's anatomically based Facial Action Coding System [FACS: (Ekman and Friesen, 1978; Ekman et al., 2002)] made movement more accessible to analysis. Yet, its relative neglect continued. Annotating intensity variation from video is arduous; and computational approaches almost exclusively follow the lead of behavioral science in emphasizing static facial expressions and AUs.

When AAA has considered temporal information, it typically has done so to improve detection of facial expressions or AUs. HMMs, for instance, have been used to temporally segment AUs by establishing a correspondence between their onset, peak, and offset and an underlying latent state. Valstar and Pantic (2007) used a combination of SVM and HMM to temporally segment and recognize AUs. Koelstra and colleagues (Koelstra et al., 2010) used Gentle-Boost classifiers on motion from a non-rigid registration combined with an HMM. Similar approaches include a non-parametric discriminant HMM (Shang and Chan, 2009) and partially observed hidden conditional random fields (Chang et al., 2009). In related work, Cohen et al. (2003) used Bayesian networks to classify the six universal expressions from video. Naive-Bayes classifiers and Gaussian tree-augmented naive Bayes (TAN) classifiers were also used to learn dependencies among different facial motion features. In a series of papers, Qiang and his colleagues (Tong et al., 2007, 2010; Li and Ji, 2013; Wang et al., 2013) used dynamic Bayesian networks to detect AUs. In each of these cases, the goal was improved AUs or facial expressions detection [for a more complete review, please see Zeng et al. (2012) and Sariyanidi et al. (2015)].

Emphasis on static holistic expressions and AUs, however, ignores a critical aspect of affect communication. People do not communicate solely by the exchange of static face displays. Communication is multimodal and dynamic. Faces and heads move and the dynamics of that movement convey communicative intent and emotion. The meaning of an expression can differ markedly depending on its dynamics and on the dynamics of modalities with which it is "packaged" (Beebe and Gerstman, 1980). Smiling coordinated with contraction of the sphincter muscle around the eyes, which raises the cheeks and creates wrinkles lateral to the eye corners, communicates enjoyment when accompanied by little or upward head motion but embarrassment when head motion is downward and to the side (Keltner, 1995). Spontaneous but not posed smiles evidence ballistic timing (Cohn and Schmidt, 2004). Such differences

are sufficiently powerful to discriminate between spontaneous and posed expressions with over 90% accuracy (e.g., Cohn and Schmidt, 2004; Valstar et al., 2006).

Dynamics communicates longer-duration affective states as well. Mergl and colleagues using motion capture (Mergl et al., 2005) found that the onset velocity of lip corners in laughter and posed expressions systematically varied between depressed and non-depressed participants. Girard and colleagues (Girard et al., 2014), in a longitudinal study of adult patients in treatment for depression, found that head motion velocity increased as they recovered from depression. Dibeklioglu and colleagues (Dibeklioglu et al., 2015) reported related findings for both head and face. In addition to communicating affect, head motion serves regulatory functions. Head-nodding can signal agreement and head-turning disagreement (Knapp and Hall, 2010). Head motion signals attention to a partner or a shared target, and depending on its pose and dynamics head motion invites or discourages speaker switching. If the dynamics of head and facial movement is absent, communication falters. Dynamics of head and face movement independently of specific facial expressions convey affective information and serves the pragmatics of human communication.

Little attention outside of behavioral science has addressed the dynamics of affective information apart from expression detection. An exception was Busso and colleagues (Busso et al., 2007) who found that head movement in avatars could provide emotion-specific information. A pending challenge for AAA is to address the dynamics of head and facial movement in emotion communication.

To understand the beginnings of non-verbal communication, we explored whether AAA could reveal the extent to which the dynamics of head and facial movement of infants communicate affective meaning independent of the morphology of facial expression. Additionally, we wanted to inform subsequent research into how the dynamics of non-verbal behavior in infants may be related to change with development in normative and high-risk samples. Both theory and data suggest that affect communication, manifested by non-verbal behaviors (such as head and face), plays a critical role in infant's social, emotional, and cognitive development (Campos et al., 1983; Tronick, 1989; Izard et al., 2011; Messinger et al., 2012). Differences in how infants respond to positive and negative emotion inductions are predictive of developmental outcomes that include attachment security (Cohn et al., 1991), and behavioral problems (Moore et al., 2001). Existing evidence suggests that head movement, facial expression, and attention may be closely coordinated (Michel et al., 1992). With the possible exception of Messinger and colleagues (Messinger et al., 2009, 2012), automatic analysis of affect related behavior has only been studied in adults. We investigated whether AAA was sufficiently advanced to reveal the dynamics of emotion communication in infants.

This paper is the first to investigate automatic analysis of head and facial movement in infants during positive and negative affect. Automatic analysis of head and facial movement in infants is challenging for several reasons. The relative shape of infant faces differs markedly from that of adults. Compared to adults, infants have smoother skin, fewer wrinkles, and

often very faint eyebrows. Occlusion, due to hands in mouth and extreme head movement, is another common problem. To overcome these challenges, we used a newly developed technique to track and align 3D features from 2D video (Jeni et al., 2015). This allowed us to ask to what extent head and facial movements communicate infants' affect. Affect was represented as a continuous bipolar dimension ranging from intense negative to intense positive (Messinger et al., 2009; Baker et al., 2010). To elicit positive and negative affect, we used two age-appropriate emotion inductions, Bubble Task and Toy-Removal Task (Goldsmith and Rothbart, 1999). We then used quantitative measurements to compare the spatiotemporal dynamics of head and facial movements during negative and positive affect.

We found strong evidence that head and facial movements communicate affect. Velocity and acceleration of head and facial movements were greater during negative affect compared with positive affect and observers' time-varying ratings of infant affect were strongly related to head and facial dynamics. The dynamics of head and facial movement together accounted for about a third of the variance in behavioral ratings of infant affect.

EXPERIMENTAL SET-UP AND HEAD AND FACIAL LANDMARKS TRACKING

We investigated the extent to which non-verbal behavior apart from specific facial expressions communicate affect in infants at ages as young as 13 months. We asked whether the temporal dynamics of head and facial movements communicate information about positive and negative affect. We used an experimental paradigm to elicit positive and negative affect. We hypothesized that infants' head and facial movements during positive and negative affect would systematically differ.

Participants

Participants were 31 ethnically diverse infants ($M = 13.1$, $SD = 0.52$) recruited as part of a multi-site study involving children's hospitals in Seattle, Chicago, Los Angeles, and Philadelphia. Participants in this study were primarily from the Seattle site. Two infants were African-American, 2 Asian-American, 6 Hispanic-American, 17 European-American, 3 Multiracial, and 1 unknown. Ten were girls. Twelve infants were mildly affected with craniofacial microsomia (CFM). CFM is congenital condition associated with varying degrees of facial asymmetry. Case-control differences between CFM and unaffected infants will be a focus of future research as more infants are ascertained. The study involved human subjects. Parents all gave informed consent for their infant to participate in the experiment. Additionally, informed consent was obtained for all images that appear in this publication.

Observational Procedures

Infants were seated in a highchair in front of a table with an experimenter and their mother across from them. The experimenter sat to the mothers' left out of camera view and closer to

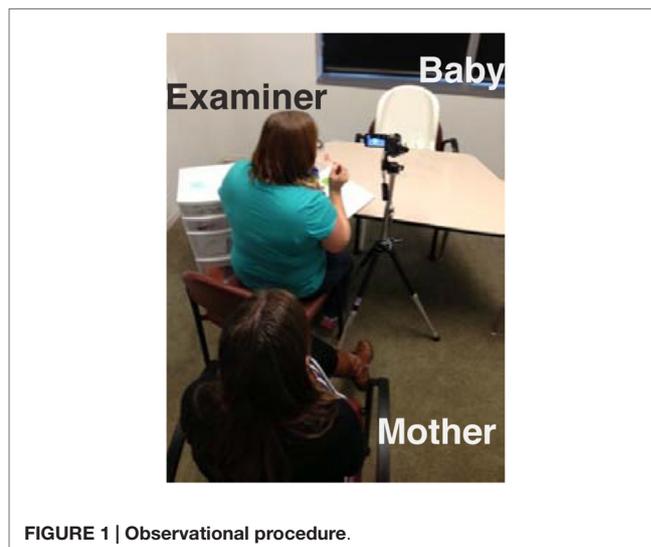


FIGURE 1 | Observational procedure.

the table as shown in **Figure 1**. Two age-appropriate emotion inductions were administered with each consisting of multiple trials.

Emotion Induction

The "Bubble Task" was intended to elicit positive affect (surprise, interest, and amusement); the "Toy-Removal Task" was intended to elicit negative affect (frustration, anger, and distress).

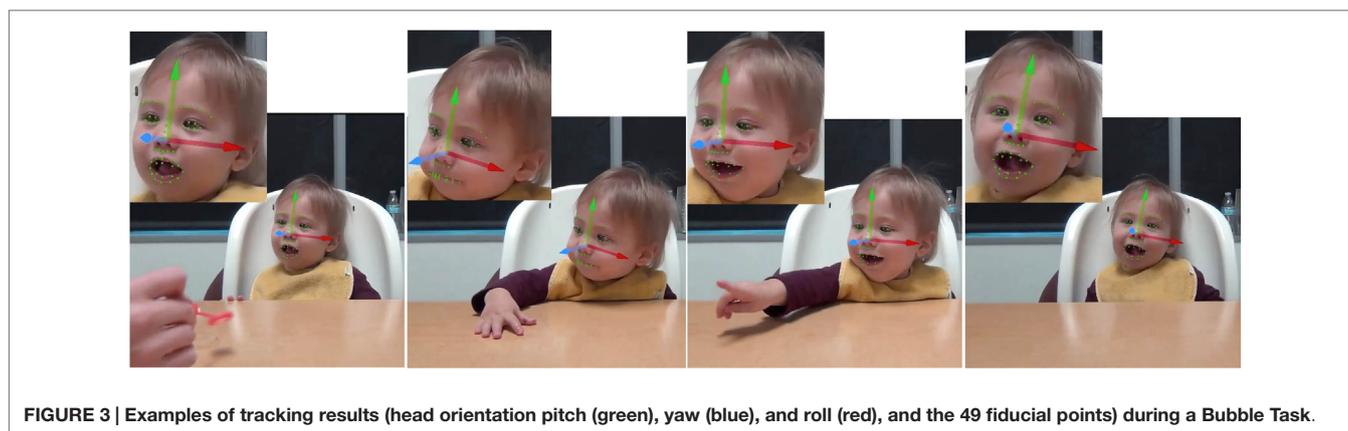
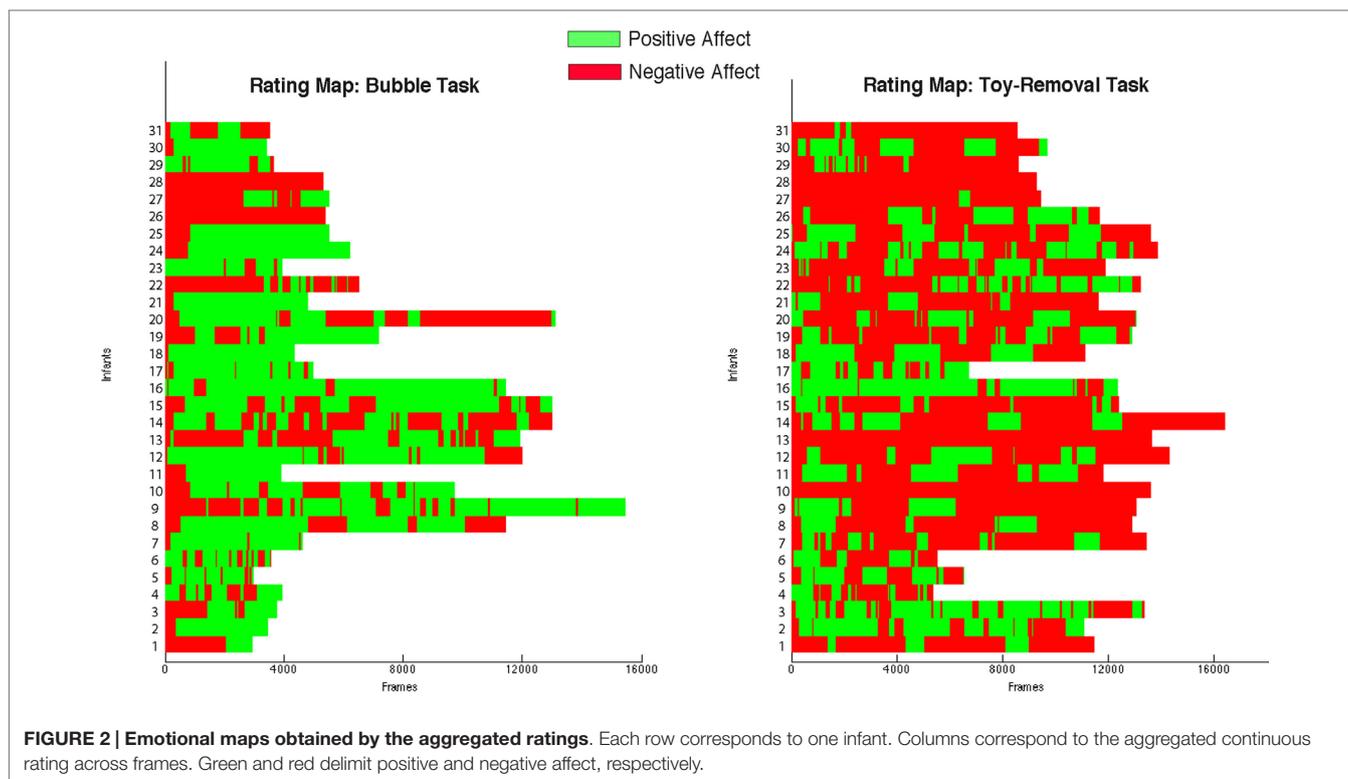
Bubble Task

Soap bubbles were blown toward the center of the table and below camera view (see **Figure 3**, first row). Before blowing bubbles, the examiner attempted to build suspense (e.g., counting 1–2–3 and then saying "Ooh, look at the pretty bubbles), Can you catch one? Where's that bubble going?" and scaffolding infant engagement (e.g., "allow all bubbles to pop before continuing again"). This procedure was repeated multiple times.

Toy-Removal Task

The examiner showed a car to the infant and demonstrated how it works by "running" the car on the table between her own hands to generate interest (e.g., car is "wound up" by pulling it back a few inches with all four wheels on the table surface) (Goldsmith and Rothbart, 1999). The toy car then was placed within the infants' reach, allowing them to play with it for few seconds (see **Figure 4**, first row). The examiner then gently took back the toy car and placed it inside a clear plastic bin just out of the infant's reach for 30 s (see **Figure 4**, rows 2–4). After 30 s, the toy was returned to the infant. This procedure was repeated one to three times.

Source videos were manually reviewed to identify the beginning and ending of the entire segments of Bubble and Toy-Removal tasks, respectively. For all infants, the Bubble Task segments started when the interviewer first said, "look what I have," "ready?" or count down before they first blow the bubbles and ended when interviewers said, "all done." For all infants, the Toy-Removal Task segments started when the interviewer said,



“look what I have” while introducing the car to the infants for the first time or the off-screen sound of the car on the table was heard. The Toy-Removal Task segments ended when the interviewer said, “all done” or “good job!” after the last repetition.

Neither task required an active response from the infant. Qualitative observations suggest that head pose and hand position appeared to be independent of task. For instance, as shown in **Figures 3** and **4**, frontal and non-frontal pose and hand placements were observed in both tasks. The infant reached toward the bubbles in **Figure 3** and withdrew his hand from the toy in **Figure 4**.

The trials were recorded using a Sony DXC190 compact camera at 60 frames per second. Infants' face orientation to the

cameras was approximately 15° from frontal, with considerable head movement.

Manual Continuous Ratings

Despite the standardized emotion inductions, positive and negative affect could occur during both tasks. To control for this possibility and to provide actual ground truth for positive and negative affect, independent manual continuous ratings of valence were obtained from four naïve raters. Observers rated intensity of both positive and negative affect. This approach to measuring emotion is informed by a 45° rotation of Russell's circumplex as proposed by Watson and Tellegen (1985). Baker et al. (2010) and Messinger et al. (2009) used a similar approach.

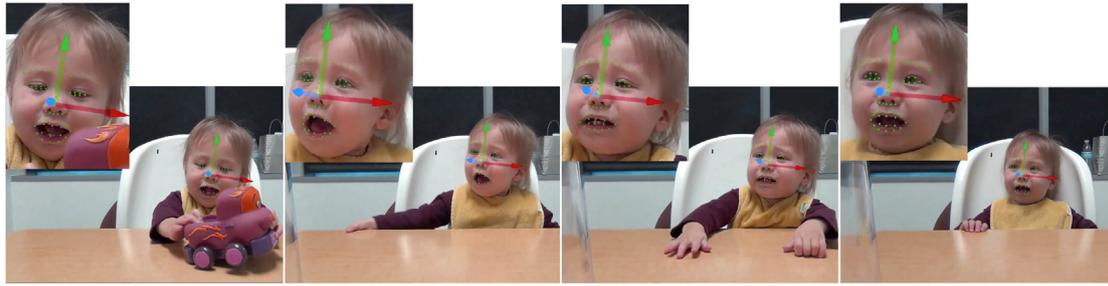


FIGURE 4 | Examples of tracking results (head orientation pitch (green), yaw (blue), and roll (red), and the 49 fiducial points) during a Toy-Removal Task.

More precisely, each of 62 video recordings (31 infants \times 2 conditions, Bubble and Toy-Removal Tasks) was independently rated from the video on a continuous scale from “-100”: maximum of intensity negative affect to “+100”: maximum of intensity positive affect. Continuous ratings were made using custom software (Girard, 2014). This produced a time series rating for each rater and each video. To achieve high effective reliability, following Rosenthal (2005), the time series were averaged across raters. Effective reliability was evaluated using intraclass correlation (McGraw and Wong, 1996; Rosenthal, 2005). The ICC was 0.91, which indicated high reliability.

Figure 2 shows the emotional maps obtained from the aggregated ratings. Continuous ratings of Bubble Tasks are shown on the right, and continuous ratings of Toy-Removal Tasks are shown on the left. In most cases, segments of positive and negative affect (delimited in green and red, respectively in **Figure 2**) alternated throughout the interaction regardless of the experimental tasks. In a few cases, infants never progressed beyond negative affect even in the nominally positive Bubble Task (see rows 4 and 6 from the top in **Figure 2**). This pattern highlights the importance of using independent criteria to identify episodes of positive and negative affect instead of relying upon the intended goal of each condition.

Infants were more positive in the Bubbles Task and more negative in the Toy-Removal Task. The ratio of positive to negative affect was significantly higher in the Bubbles Task compared to the Toy-Removal Task ($t = 7.60$, $df = 30$, $p \leq 0.001$, see green segments in **Figure 2** right), and the ratio of negative to positive affect was significantly higher in the Toy-Removal Task compared to the Bubbles Task ($t = 7.60$, $df = 30$, $p \leq 0.001$, see red segments in **Figure 2** left). To be conservative, for analysis we compared segments rated as positive in the Bubbles Task (see green segments in **Figure 2** right) and segments rated as negative in Toy-Removal Task (see red segments in **Figure 2** left).

The differences between the selected affective segments were specific to valence (i.e., positive and negative affect) rather than the intensity. The absolute intensity of behavioral ratings did not differ between segments of positive and negative affect ($t = -1.41$, $df = 28$, $p = 0.168$). The difference was of valence.

Automatic Tracking of Head Orientation and Facial Landmarks

A recently developed generic 3D face tracker (ZFace) was used to track the registered 3D coordinates from 2D video of 49 facial landmarks, or fiducial points, and the 3 degrees of rigid head movements (i.e., pitch, yaw, and roll, see **Figures 3** and **4**). The tracker uses a combined 3D supervised descent method (Xiong and de la Torre, 2013), where the shape model is defined by a 3D mesh and the 3D vertex locations of the mesh. The tracker registers a dense parameterized shape model to an image such that its landmarks correspond to consistent locations on the face. The robustness of the method for 3D registration and reconstruction from 2D video was validated in a series of experiments [for details, please see Jeni et al. (2015)].

To further evaluate the concurrent validity of the tracker for head pose, it was compared with the CISRO cylinder-based 3D head tracker (Cox et al., 2013) on 16 randomly selected videos of Bubble and Toy-Removal Tasks from eight infants. Concurrent correlations between the two methods for pitch, yaw, and roll were 0.74, 0.93, and 0.92, respectively. Examples of the tracker performance for head orientation and facial landmarks localization are shown in **Figures 3** and **4**.

AUTOMATIC MEASUREMENT OF HEAD AND FACIAL MOVEMENT

In the following, we first describe the results of the automatic tracking and the manual validation of the tracked data. We then describe how the extracted head orientations and fiducial points were reduced for analysis. Measures of displacement, velocity, and acceleration were computed for both head and facial landmark movements.

Evaluation and Manual Validation of the Tracking Results

For each video frame, the tracker outputs the 3D coordinates of the 49 fiducial points and 3 degrees of freedom of rigid head movement (pitch, yaw, and roll) or a failure message when a frame could not be tracked. To guard against possible tracking error,

tracking results were overlaid on the source videos and manually reviewed. Examples are shown in **Figures 3** and **4**. Frames that could not be tracked or failed visual review were not analyzed. Failures were due primarily to self-occlusions (e.g., hand in the mouth) and extreme head turn (out of frame). The percentage of well-tracked frames was lower for Toy-Removal Task than for Bubble Task (see **Table 1**).

Measurement of Head Movement

Head angles in the horizontal, vertical, and lateral directions were selected to measure head movement. These directions correspond to head nods (i.e., pitch), head turns (i.e., yaw), and lateral head inclinations (i.e., roll), respectively (see blue, green, and red rows in **Figures 3** and **4**). Head angles were converted into angular displacement, angular velocity, and angular acceleration. For pitch, yaw, and roll, angular displacement was computed by subtracting the overall mean head angle from each observed head angle within each valid segment (i.e., consecutive valid frames). Similarly, angular velocity and angular acceleration for pitch, yaw, and roll were computed as the derivative of angular displacement and angular velocity, respectively. Angular velocity and angular acceleration allow measurements of changes in head movement from one frame to the next.

The root mean square (RMS) was then used to measure the magnitude of variation of the angular displacement, the angular velocity, and angular acceleration. The RMS value of the angular displacement, the angular velocity, and the angular acceleration was computed as the square root of the mean value of the squared values of the quantity taken over each consecutive valid segment for each condition. The RMS of the horizontal, vertical, and lateral angular displacement, angular velocity, and angular acceleration were then calculated for each consecutive valid segment, for each condition, and for each infant separately. RMSs of head movement as used in the analyses refer to the mean of the RMSs of consecutive valid segments.

Measurement of Facial Movement

The movement of the 49 detected 3D fiducial points (see **Figures 3** and **4**) was selected to measure facial movement. The movement of the 49 fiducial points corresponds to the movement (without rigid head movements) of the corresponding facial features such the eyes, eyebrows, and mouth. First, the mean position in the face of each detected fiducial point within all valid segments (i.e., consecutive valid frames) was computed. The displacement of each one of the 49 fiducial points was computed by measuring the Euclidian distance between its current position in the face and its mean position in the face. The velocity of movement of the 49 detected fiducial points was computed as the derivative of the corresponding

displacement, measuring the speed of change of the fiducial points position on the face from one frame to the next. Similarly, acceleration of each one of the 49 detected fiducial points was computed as the derivative of the corresponding velocities.

The RMS was then used to measure the magnitude of variation of the fiducial points' displacement, velocity, and acceleration, respectively. Similarly to head movement, the RMS values of the fiducial points' displacement, the fiducial points' velocity, and the fiducial points' acceleration were computed for each consecutive valid segment for each condition. RMSs of facial movement as used for analyses refer to the mean of the RMSs of consecutive valid segments. For simplicity, in the following sections the fiducial points' displacement, fiducial points' velocity, and fiducial points' acceleration are referred to as facial displacement, facial velocity, and facial acceleration, respectively.

Because the movements of individual points were highly correlated, principal components analysis (PCA) was used to reduce the number of parameters (a compressed representation of the 49 RMSs of facial displacement, velocity, and acceleration, respectively). The first principal components of displacement, velocity, and acceleration accounted for 60, 71, and 73% of the respective variance and were used for analysis.

DATA ANALYSIS AND RESULTS

Because of the repeated-measures nature of the data, mixed analyses of variance (ANOVA) were used to evaluate mean differences in head and facial movements between positive and negative affect. The ANOVAs included sex (between-subjects factor), condition (within-subjects factor) and the interaction between sex and condition.¹ Separate ANOVAs were used for each measure (i.e., head and facial RMSs of displacement, velocity, and acceleration). Paired *t*-tests were used for *post hoc* analyses following significant ANOVAs. In a follow-up analysis, discriminant analysis was used to assess to what extent the significant measures alone allow to detect positive and negative affect.

Infant Head Movement During Positive and Negative Affect

For RMS angular displacement of yaw and roll but not pitch, there was a main effect for sex ($F_{2,56} = 8.9$, $p = 0.004$, $F_{2,56} = 7.35$, $p = 0.008$, and $F_{2,56} = 0.54$, $p = 0.46$, respectively).² Yaw and roll angular amplitudes were higher for girls than for boys. The sex-by-condition interaction was not significant. The sex effect should, however, be interpreted with caution given the unbalanced distribution of male and female infants.

TABLE 1 | Percentage of valid tracking during Bubble and Toy-Removal Tasks.

	Bubble Task	Toy-Removal Task
Infants	85	70

$t = 4.76$, $p \leq 0.01$, $df = 30$.

¹An interaction effect is a change in the simple main effect of one variable over levels of another variable. A sex-by-episode interaction is a change in the simple main effect of sex over levels of episode or the change in the simple main effect of episode over levels of sex.

² $F_{2,56}$ corresponds to the F statistic and its associated degrees of freedom. The first number refers to the degrees of freedom for the specific effect and the second number to the degrees of freedom of error.

For RMS angular velocity and RMS angular acceleration, there were no sex- or interaction effects for pitch, yaw, or roll. Consequently, for angular velocity and angular acceleration, only main effects of condition were tested. For RMS angular velocity of pitch, yaw, and roll, there were significant effects for condition ($F_{1,58} = 5.51, p = 0.02$; $F_{1,58} = 11.21, p = 0.001$; and $F_{1,58} = 6.03, p = 0.01$, respectively). Pitch, yaw, and roll RMS angular velocities increased in negative affect compared with positive affect (see **Table 2**). Similarly, for RMS angular acceleration of pitch, yaw, and roll, there were significant effects for condition ($F_{1,58} = 3.99, p = 0.05$; $F_{1,58} = 5.29, p = 0.02$; and $F_{1,58} = 4.04, p = 0.04$, respectively). Pitch, yaw, and roll RMS angular accelerations increased in negative compared with positive affect (see **Table 2**).

Infants Facial Movement During Positive and Negative Affect

Similar to head movement, preliminary mixed ANOVAs included sex and sex-by-condition interaction. For facial movement, neither effect was significant; they were thus dropped from the model. For RMS facial displacement, RMS facial velocity, and RMS facial acceleration, there were significant effects for condition ($F_{1,58} = 3.91, p = 0.05$; $F_{1,58} = 4.16, p = 0.04$; and $F_{1,58} = 4.65, p = 0.03$, respectively). RMS facial displacement, RMS facial velocity, and RMS facial acceleration were significantly higher in negative affect compared with positive affect (**Table 3**).

Automatic Detection of Positive and Negative Affect from Head and Facial Movement

To further evaluate the predictive value of head and facial movement for positive and negative affect, all of the significant measures identified in the previous sections were entered together into a linear discriminant analysis. These consisted of nine parameters: pitch, yaw and roll angular velocities and angular accelerations, facial amplitude, facial velocity, and facial acceleration. For a compact representation, principal component analysis was first used to reduce the nine parameters to two components accounting for 79.36% of the variance. The resulting discriminant function was highly significant (Wilks' lambda = 0.86, $p = 0.015$). Using leave one-case out cross-validation, the mean classification rate for positive and negative affect was 65%. Kappa, a measure of agreement, which adjusts for chance, was 0.30, which represents a fair agreement.

Correlation Between Head and Facial Movements and Behavioral Ratings

Because both head and facial movements for most measures were faster during negative than positive affect, we wished to investigate whether head and facial movements were themselves correlated and to investigate the correlation between head and facial movements and the behavioral ratings. To reduce the number of facial movement parameters to a compact

TABLE 2 | Post hoc paired t-tests (following significant ANOVAs) for pitch, yaw, and roll RMS angular velocities, and RMS angular accelerations.

	Positive	Negative	Differences by paired t-test			
	M (SD)	M (SD)	t	df	Observed p	Holm p
Angular velocity						
Pitch	0.66 (0.14)	0.77 (0.20)	-3.56	28	0.001	0.005
Yaw	0.59 (0.12)	0.72 (0.17)	-4.59	28	<0.001	<0.004
Roll	0.32 (0.08)	0.40 (0.14)	-3.06	28	0.004	0.03
Angular acceleration						
Pitch	0.90 (0.19)	1.02 (0.25)	-3.01	28	0.005	0.025
Yaw	0.77 (0.14)	0.89 (0.23)	-3.07	28	0.004	0.016
Roll	0.40 (0.11)	0.48 (0.19)	-2.19	28	0.03	0.05

To control for experiment-wise error, a Holm correction was used for each set of multiple comparisons.

M, mean of RMSs; t, t-ratio; df, degrees of freedom; p, probability.

Critical values for family-wise error per Holm correction (Holm, 1979; Aickin and Gensler, 1996) are 0.025, 0.016, and 0.05 for Pitch, Yaw, and Roll, respectively.

TABLE 3 | Post hoc paired t-tests (following significant ANOVAs) for RMS facial displacements, RMS facial velocities, and RMS facial accelerations.

	Positive	Negative	Differences by paired t-test			
	M (SD)	M (SD)	t	df	Observed p	Holm p
Facial displacement	-0.96 (3.38)	0.90 (3.90)	-2.30	28	0.02	0.05
Facial velocity	-0.12 (0.37)	0.11 (0.51)	-2.85	28	0.008	0.016
Facial acceleration	-0.16 (0.48)	0.15 (0.67)	-2.80	28	0.01	0.025

To control for experiment-wise error, a Holm correction was used for each set of set of multiple comparisons.

M, mean of RMSs; t, t-ratio; df, degrees of freedom; p, probability.

Critical values for family-wise error per Holm correction (Holm, 1979; Aickin and Gensler, 1996) are 0.05, 0.016, and 0.025 for facial displacement, velocity, and acceleration, respectively.

TABLE 4 | Mean correlations between infants' time series of angular displacement of pitch, yaw, and roll, and facial displacement.

	Bubble Task			Toy-Removal Task		
	Comp-1	Comp-2	Comp-3	Comp-1	Comp-2	Comp-3
Pitch	0.25	0.22	0.15	0.20	0.28	0.18
Yaw	0.20	0.17	0.11	0.18	0.16	0.13
Roll	0.19	0.14	0.14	0.11	0.11	0.14

All correlations differed significantly from 0, $p \leq 0.01$.

representation, PCA was used to reduce the 49 fiducial point time series to three time series components. These three components accounted for a total of 70% of the variance (see **Table 4**). We then computed the correlation between infants' time series of angular displacement of pitch, yaw, and roll, and facial displacement (measured using the compressed three time series: Comp-1, Comp-2, and Comp-3) during Bubble and Toy-Removal Tasks (see **Table 4**). Overall, the obtained mean correlations suggest that head and facial movements were only weakly correlated.

We next examined the relative importance of infants' head angular displacement (i.e., pitch, yaw, and roll), and facial displacement (Comp-1, Comp-2, and Comp-3) to the observers' ratings. To do this, we performed multiple linear regressions to learn more about the relationship between the behavioral ratings and infants' head angular displacement of pitch, yaw, and roll, and facial displacement measured using the compressed three time series components (Comp-1, Comp-2, and Comp-3). Three regression models were fitted. First, a regression model between the time series of angular displacement of pitch, yaw, and roll, and continuous ratings. Second, a regression model between the time series of facial displacement (measured using the compressed three times series Comp-1, Comp-2, and Comp-3) and continuous ratings. Finally, a regression model was fitted between both the infants' time series of angular displacement of pitch, yaw, and roll, and facial displacement, and the continuous ratings. The adjusted *R*-squared values for each model are reported in **Figures 5** and **6**. *R*-squared is a measure of variance accounted for in the dependent variable by the set of predictors.

The obtained results show that both head angular displacement and facial displacement accounted for significant variation of the continuous ratings. Head angular displacement accounted, respectively, for 16 and 15% of the variance in behavioral ratings of infant affect during Bubble and Toy-Removal tasks. Compared to head movement, facial movement accounted, respectively, for 26 and 24% of the variance in behavioral ratings of infant affect during Bubble and Toy-Removal tasks. The two modalities together accounted, respectively, for 36 and 32% of the variance in behavioral ratings of infant affect during Bubble and Toy-Removal tasks. For both tasks, facial movement accounted for more of the variance of the continuous rating than did head movement ($p \leq 0.01$, $df = 1$, $p \leq 0.01$, $df = 1$, for Bubble and Toy-Removal tasks, respectively) and head and facial movement together accounted for more of the variance of the continuous ratings than head and facial movement separately ($p \leq 0.01$, $df = 1$, $p \leq 0.05$, $df = 1$, for Bubble and Toy-Removal tasks, respectively).

DISCUSSION AND CONCLUSION

We tested the hypothesis that head and facial movements in infants as measured by AAA vary between positive and negative affect. We found striking differences in head and facial dynamics between positive and negative affect. Velocity and acceleration for both head and facial movements were greater during negative than positive affect.

Were these effects specific to the affective differences between conditions? Alternative explanations might be considered. One is whether the differences were specific to intensity rather than valence. From a dimensional perspective, affective phenomena exist in a two-dimensional space defined by intensity and valence [i.e., a "circumplex" model as elaborated by Russell and Bullock (1985)]. If the intensity of observer ratings were greater during one or the other condition that would suggest that the valence differences we observed were confounded by intensity. To evaluate this possibility, we compared observer ratings of absolute intensity between conditions. Absolute intensity failed to vary between conditions.

This negative finding suggests that intensity of response was not responsible for the systematic variation we observed between conditions. They were specific to valence of affect. Another possible alternative explanation could be demand characteristics of the tasks. Toy-removal elicits reaching for the toy; bubble task elicits head motion toward the bubbles. To consider whether these characteristics may have been a factor, we compared the amount of head movement (i.e., displacement) between conditions. Displacement of the head failed to vary between tasks. The lack of differences in head displacement suggests that demand characteristics did not account for the differences in acceleration and velocity we observed. Lastly, it is possible that missing data from tracking error played a role. Further improvements in tracking infants' head and facial movements especially during distress would be needed to rule this alternative.

The obtained results are consistent with prior work in adult psychiatric populations that found that head movement systematically varies between affective states. For instance, our findings are consistent with a recent study that found greater velocity of head motion in intimate couples during episodes of high vs. low conflict (Hammal et al., 2014). They are consistent as well with the finding that velocity of head movement was inversely correlated with depression severity. Girard and colleagues (Girard et al., 2014) found that velocity was lower when depressive symptoms were clinically significant and increased as patients improved. Reduced reactivity is consistent with many evolutionary theories of depression (Klinger, 1975; Nesse, 1983; Fowles, 1994) and highlights the symptoms of apathy and psychomotor retardation.

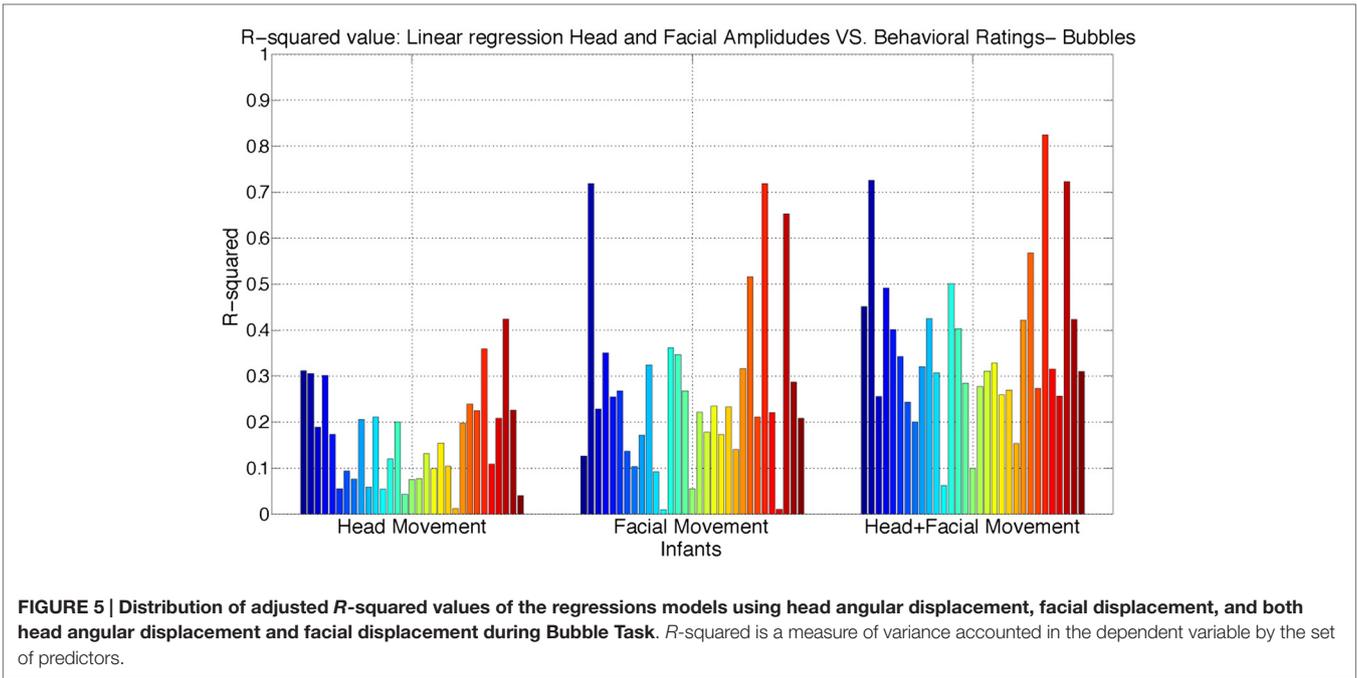


FIGURE 5 | Distribution of adjusted *R*-squared values of the regressions models using head angular displacement, facial displacement, and both head angular displacement and facial displacement during Bubble Task. *R*-squared is a measure of variance accounted in the dependent variable by the set of predictors.

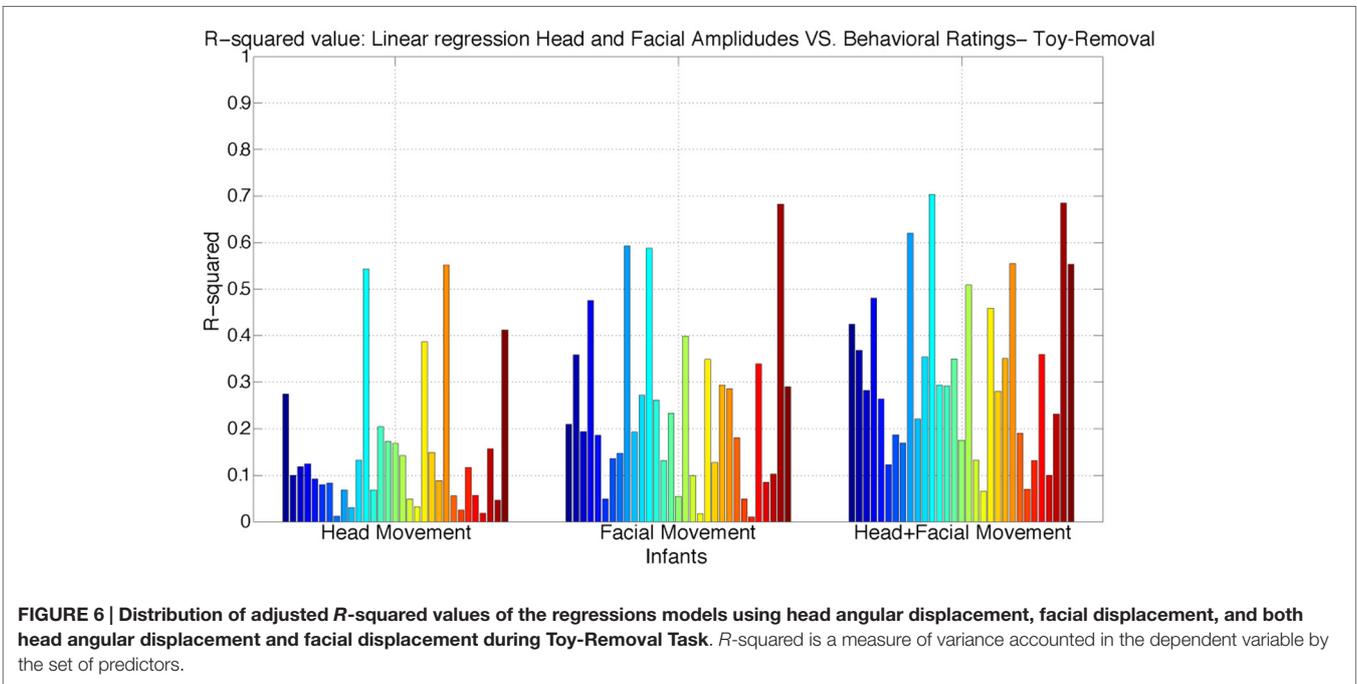


FIGURE 6 | Distribution of adjusted *R*-squared values of the regressions models using head angular displacement, facial displacement, and both head angular displacement and facial displacement during Toy-Removal Task. *R*-squared is a measure of variance accounted in the dependent variable by the set of predictors.

Previous studies have found that depression is marked by reductions in general facial expressiveness (Renneberg et al., 2005) and head movement (Fisch et al., 1983; Joshi et al., 2013). Together, these findings suggest that head and facial movements are lowest during depressed affect, increase during neutral to positive affect, and increase yet again during conflict and negative affect. Future research is needed to investigate more fully the relation between affect and the dynamics of head and facial movement.

Finally, some infants had mild CFM. While CFM could potentially have moderated the effects we observed, too few infants with CFM have yet been enrolled to evaluate this hypothesis. Case/control status has not yet been unblinded to investigators. For these reasons, we focus on the relation between head and facial movement and affect independent of CFM status. Once study enrollment is further advanced, the possible role of CFM as a moderator can be examined.

We also examined the relative importance of infants' head facial displacement and the behavioral ratings independently of specific facial expressions and AUs. The dynamics of head and facial movement together accounted for about a third of the variance in behavioral ratings of infants' affect. Furthermore, preliminary results using significant measures of velocity and acceleration of head and facial movement achieved a mean classification rate of positive and negative affect equal to 65% using discriminant linear analysis. The next step will be to evaluate how do head and facial movement temporally covary with specific facial expression and AU and to compare the relative contribution of movement vs. static features in automatically detecting positive and negative affect.

The results of this study will enable further research on health care outcomes such as in infants with CFM. Facial asymmetries and impairments to the facial nerve are common in infants severely affected with CFM. From a developmental perspective, one of the most important consequences of limitations in facial muscle movement is its potentially negative impact on affect communication. Automatic measurement of expressiveness as a predictor of CFM outcomes would be of great support for clinical applications in which the cost of manual coding [e.g., FACS (Ekman and Rosenberg, 2005)] would be prohibitive. Automated measurement of expressiveness has potential in future studies of children with CFM and other conditions that affect facial expressiveness. Individual clinical assessments of infant facial movements could be used to target children for specialized interventions, both medical (e.g., surgical interventions to improve

facial symmetry and/or animation) and behavioral (e.g., behavioral studies of children's compensatory strategies for limitations in facial expressiveness, parent coaching to reduce the impact of infants' muted facial expressions on parent-child interactions). Automatic analyses of non-verbal expressiveness, such as facial expressiveness, could also be used in research to assess pre-/post intervention changes in facial movements.

The next steps will be first to understand how the expressiveness, measured by head and facial movement dynamics, is manifested in children with CFM compared to healthy matched controls. Second, compare the static (or structural) facial asymmetries to facial movement asymmetries in children with CFM, a distinction that may lead to novel diagnostic and treatment methods.

In summary, we found strong evidence that head and facial movements vary markedly with affect. The dynamics of head and facial movement together communicate affect independently of static facial expressions.

ACKNOWLEDGMENTS

Research reported in this publication was supported in part by the National Institute of Dental and Craniofacial Research of the National Institutes of Health under Award Number DE022438. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- Aickin, M., and Gensler, H. (1996). Adjusting for multiple testing when reporting research results: the bonferroni vs. Holm methods. *Am. J. Public Health* 86, 726–728. doi:10.2105/AJPH.86.5.726
- Baker, J. K., Haltigana, J. D., Brewstera, R., Jaccardb, J., and Messinger, D. S. (2010). Non-expert ratings of infant and parent emotion: concordance with expert coding and relevance to early autism risk. *Int. J. Behav. Dev.* 34, 88–95. doi:10.1177/0165025409350365
- Beebe, B., and Gerstman, L. (1980). The "packaging" of maternal stimulation in relation to infant facial-visual engagement. A case study at four months. *Merrill Palmer Q.* 26, 321–339.
- Bell, C. (1844). *The Anatomy and Philosophy of Expression as Connected with the Fine Arts*, 3 Edn. London: John Murray.
- Busso, C., Deng, Z., Grimm, M., Neumann, U., and Narayanan, S. S. (2007). Rigid head motion in expressive speech animation: analysis and synthesis. *IEEE Trans. Audio Speech Lang. Process.* 15, 1075–1086. doi:10.1109/TASL.2006.885910
- Campos, J. J., Caplowitz, K., Lamb, M. E., Goldsmith, H. H., and Stenberg, C. (1983). "Socio-emotional development," in *Handbook of Child Psychology*, 4th Edn, Vol. II, eds Haith M. M. and Campos J. J. (New York, NY: Wiley), 1373–1396.
- Chang, K. Y., Liu, T. L., and Lai, S. H. (2009). "Learning partial ly-observed hidden conditional random fields for facial expression recognition," in *IEEE Conference on Computer Vision and Pattern Recognition CVPR* (Miami), 533–540.
- Cohen, I., Sebe, N., Chen, L., Garg, A., and Huang, T. S. (2003). Facial expression recognition from video sequences: temporal and static modeling. *Comput. Vis. Image Underst.* 91, 160–187. doi:10.1016/S1077-3142(03)00081-X
- Cohn, J. F., Campbell, S. B., and Ross, S. (1991). Infant response in the still-face paradigm at 6 months predicts avoidant and secure attachment at 12 months. *Dev. Psychopathol.* 3, 367–376. doi:10.1017/S0954579400007574
- Cohn, J. F., and De la Torre, F. (2015). "Automated face analysis for affective computing," in *Handbook of Affective Computing*, eds Calvo R. A., D'Mello S. K., Gratch J., and Kappas A. (New York, NY: Oxford), 131–150.
- Cohn, J. F., and Schmidt, K. L. (2004). The timing of facial motion in posed and spontaneous smiles. *Int. J. Wavelets Multiresolut. Inf. Process.* 2, 1–12. doi:10.1142/S021969130400041X
- Cox, M., Nuevo-Chiquero, J., Saragih, J. M., and Lucey, S. (2013). "CSIRO face analysis SDK," in *10th IEEE Int. Conf. Autom. Face Gesture Recog. Shanghai, China*, Brisbane, QLD.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: Murray.
- Dibeklioglu, H., Hammal, Z., Yang, Y., and Cohn, J. F. (2015). "Multimodal detection of depression in clinical interviews," in Proceedings of the ACM International Conference on Multimodal Interaction (ICMI), Seattle, WA.
- Duchenne, G. (1990). *Mecanisme de la Physionomie Humaine in 1862. [The Mechanism of Human Facial Expression]*. (New York, NY: Cambridge University Press). Translated by R. Andrew Cuthbertson.
- Ekman, P. (1972). "Universals and cultural differences in facial expressions of emotions," in *Nebraska Symposium on Motivation*, ed. Cole J. (Lincoln, NB: University of Nebraska Press), 207–282.
- Ekman, P., and Friesen, W. V. (1978). *Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., and Hager, J. C. (2002). *Facial Action Coding System [E-book]*. Salt Lake City, UT: Research Nexus.
- Ekman, P., and Rosenberg, E. L. (2005). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System*, 2nd Edn. New York: Oxford University Press.
- Fisch, H. U., Frey, S., and Hirsbrunner, H. P. (1983). Analyzing nonverbal behavior in depression. *J. Abnorm. Psychol.* 92, 307–318. doi:10.1037/0021-843X.92.3.307
- Fowles, D. C. (1994). "A motivational view of psychopathology," in *Integrative Views of Motivation, Cognition, and Emotion*, eds Spaulding W. D. and Simon H. A. (University of Nebraska Press), 181–238.
- Girard, J. (2014). CARMA: software for continuous affect rating and media annotation. *J. Open Res. Softw.* 2(1, e5), 1–6. doi:10.5334/jors.ar
- Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., and Rosenwald, D. P. (2014). Nonverbal social withdrawal in depression: evidence

- from manual and automatic analyses. *Image Vis. Comput. J.* 32(10), 641–647. doi:10.1016/j.imavis.2013.12.007
- Goldsmith, H. H., and Rothbart, M. K. (1999). *The Laboratory Temperament Assessment Battery*. Eugene, OR: University of Oregon.
- Hammal, Z., Cohn, J. F., and George, D. T. (2014). Interpersonal coordination of head motion in distressed couples. *IEEE Trans. Affect. Comput.* 5, 155–167. doi:10.1109/TAFFC.2014.2326408
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.
- Izard, C., Woodburn, E., Finlon, K., Krauthamer-Ewing, S., Grossman, S., and Seidenfeld, A. (2011). Emotion knowledge, emotion utilization, and emotion regulation. *Emot. Rev.* 3, 44–52. doi:10.1177/1754073910380972
- Izard, C. E. (1971). *The Face of Emotion*. New York, NY: Appleton-Century-Crofts.
- Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychol. Bull.* 115, 288–299. doi:10.1037/0033-2909.115.2.288
- Jeni, L. A., Cohn, J. F., and Kanade, T. (2015). *Dense 3D Face Alignment from 2D Videos in Real-Time 11th IEEE International Conference on Automatic Face and Gesture Recognition*, Ljubljana, Slovenia.
- Joshi, J., Goecke, R., Parker, G., and Breakspear, M. (2013). “Can body expressions contribute to automatic depression analysis?” in *International Conference on Automatic Face and Gesture Recognition*, Shanghai, China.
- Keltner, D. (1995). The signs of appeasement: evidence for the distinct displays of embarrassment, amusement, and shame. *J. Pers. Soc. Psychol.* 68, 441–454. doi:10.1037/0022-3514.68.3.441
- Klinger, E. (1975). Consequences of commitment to and disengagement from incentives. *Psychol. Rev.* 82(1), 1–25. doi:10.1037/h0076171
- Knapp, M. L., and Hall, J. A. (2010). *Nonverbal Behavior in Human Communication*, 7th Edn. Boston, MA: Cengage.
- Koelstra, S., Pantic, M., and Patras, I. (2010). A dynamic texture based approach to recognition of facial actions and their temporal models. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 1940–1954. doi:10.1109/TPAMI.2010.50
- Li, Y., and Ji, Q. (2013). Simultaneous facial feature tracking and facial expression recognition. *IEEE Trans. Image Process.* 22, 2559–2573. doi:10.1109/TIP.2013.2253477
- McGraw, K. O., and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* 1, 30–46. doi:10.1037/1082-989X.1.1.30
- Mergl, R., Mavrogiorgou, P., Hegerl, U., and Juckel, G. (2005). Kinematical analysis of emotionally induced facial expressions: a novel tool to investigate hypomimia in patients suffering from depression. *J. Neurol. Neurosurg. Psychiatr.* 76, 138–140. doi:10.1136/jnnp.2004.037127
- Messinger, D. S., Mahoor, M. H., Chow, S. M., and Cohn, J. F. (2009). Automated measurement of facial expression in infant-mother interaction: a pilot study. *Infancy* 14, 285–305. doi:10.1080/15250000902839963
- Messinger, D. S., Mattson, W., Mohammad, M. H., and Cohn, J. F. (2012). The eyes have it: making positive expressions more positive and negative expressions more negative. *Emotion* 12, 430–436. doi:10.1037/a0026498
- Michel, G. F., Camras, L., and Sullivan, J. (1992). Infant interest expressions as coordinative motor structures. *Infant Behav. Dev.* 15, 347–358. doi:10.1016/0163-6383(92)80004-E
- Moore, G. A., Cohn, J. F., and Campbell, S. B. (2001). Infant affective responses to mother's still-face at 6 months differentially predict externalizing and internalizing behaviors at 18 months. *Dev. Psychol.* 37, 706–714. doi:10.1037/0012-1649.37.5.706
- Nesse, R. M. (1983). Is depression an adaptation? *Arch. Gen. Psychiatry* 57, 2000.
- Renneberg, B., Heyn, K., Gebhard, R., and Bachmann, S. (2005). Facial expression of emotions in borderline personality disorder and depression. *J. Behav. Ther. Exp. Psychiatry* 36, 183–196. doi:10.1016/j.jbtep.2005.05.002
- Rosenthal, R. (2005). “Conducting judgment studies: some methodological issues,” in *The new handbook of methods in nonverbal behavior research. Series in Affective Science*, eds Harrigan J. A., Rosenthal R., Scherer K. R. (New York, NY: Oxford University Press) 199–234.
- Russell, J. A., and Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *J. Pers. Soc. Psychol.* 48, 1290–1298. doi:10.1037/0022-3514.48.5.1290
- Sariyanidi, E., Gunes, H., and Cavallaro, A. (2015). Automatic analysis of facial affect: a survey of registration, representation and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 113–133. doi:10.1109/TPAMI.2014.2366127
- Shang, L., and Chan, K.-P. (2009). “Nonparametric discriminant HMM and application to facial expression recognition,” in *IEEE Conference on Computer Vision and Pattern Recognition CVPR* (Miami), 2090–2096.
- Tong, Y., Chen, J., and Ji, Q. (2010). A unified probabilistic framework for spontaneous facial action modeling and understanding. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 258–274. doi:10.1109/TPAMI.2008.293
- Tong, Y., Liao, W., and Ji, Q. (2007). Facial action unit recognition by exploiting their dynamics and semantic relationships. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 1683–1699. doi:10.1109/TPAMI.2007.1094
- Tronick, E. Z. (1989). Emotions and emotional communication in infants. *Am. Psychol.* 44, 112–119. doi:10.1037/0003-066X.44.2.112
- Valstar, M., Girard, J., Almaev, T., McKeown, G., Mehu, M., Yin, L., et al. (2015). “FERA 2015 – second facial expression recognition and analysis challenge,” in *Proc. IEEE Int'l Conf. Face and Gesture Recognition*, Ljubljana, Slovenia.
- Valstar, M. F., and Pantic, M. (2007). “Combined support vector machines and hidden markov models for modeling facial action temporal dynamics,” in *Proceedings of IEEE Workshop on Human Computer Interaction*, Vol. 4796 (Rio de Janeiro: LNCS), 118–127.
- Valstar, M. F., Pantic, M., Ambadar, Z., and Cohn, J. F. (2006). “Spontaneous vs. posed facial behavior: automatic analysis of brow actions,” in *Proc. ACM Intl. Conf. on Multimodal Interfaces (ICMI'06)* (Banff, AB), 162–170.
- Valstar, M. F., Schuller, B., Smith, K., Eyben, F., Jiang, B., Bilakhia, S., et al. (2013). “AVEC 2013 – the continuous audio/visual emotion and depression recognition challenge,” in *Proc. 3rd ACM International Workshop on Audio/Visual Emotion Challenge* (New York, NY: ACM), 3–10.
- Wang, Z., Wang, S., and Ji, Q. (2013). “Capturing complex spatio-temporal relations among facial muscles for facial expression recognition,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 13)* (Portland: Oregon), 3422–3429.
- Watson, D., and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychol. Bull.* 98, 219–235. doi:10.1037/0033-2909.98.2.219
- Xiong, X., and de la Torre, F. (2013). *Supervised Descent Method and its Applications to Face Alignment. Computer Vision and Pattern Recognition (CVPR)* (Portland: Oregon), 532–539.
- Zeng, Z., Pantic, M., Roisman, G. I., and Huang, T. S. (2012). A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 39–58. doi:10.1109/TPAMI.2008.52

Conflict of Interest Statement: 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.

Copyright © 2015 Hammal, Cohn, Heike and Speltz. 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) or licensor 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.