**From:** Jonathan Gratch   
**Sent:** Wednesday, June 12, 2019 2:28 PM  
**To:** Kalin Stefanov <[kstefanov@ict.usc.edu](mailto:kstefanov@ict.usc.edu)>; Mohammad Soleymani <[soleymani@ict.usc.edu](mailto:soleymani@ict.usc.edu)>  
**Subject:** head gestures

You guys may know this stuff but here is some work related to head gestures (annotation schemes and automatic techniques) it case it is of some interest

**Annotation Schemes**

* (Kousidis, Malisz et al. 2013): distinguishes several gestures

·                    

* (Wlodarczak, Buschmeier et al. 2012): annotated nod, jerk, tilt, turn, protrusion and retraction, plus number of times each repeated. Tried to infer communicative function
* (Paggio and Navarretta 2011)Tried to infer “feedback function” of different head movements (Nod, Jerk, HeadForward, HeadBackward,Tilt, SideTurn, Shake, Waggle, HeadOther)

**Work on annotating or automatically recognizing or interpreting head gestures**

* (Gunes and Pantic 2010): maps the amount and direction of head motion, and occurrences of head nods and shakes into arousal, expectation, intensity, power and valence level.  Use SAL dataset
* (Anis, Zakia et al. 2018): think looking at raw head kinematics to detect depression
* (Hammal, Cohn et al. 2014): looks like examined raw kinematics in distressed couples
* (Hammal, Cohn et al. 2015): head motion in mother-infant dyads
* (Morency, Sidner et al. 2005): nods and shake recognition
* (Adams, Mahmoud et al. 2015): this paper tries to show that head motions contribute to the interpretation of facial expressions

**Corpora**

* (Buschmeier, Malisz et al. 2014): a corpus of “active listening behaviors”
* (Bilakhia, Petridis et al. 2015): Dyadic data. contains “head movements”
* ICT rapport dataset (annotated for nods and shakes)

Adams, A., et al. (2015). Decoupling facial expressions and head motions in complex emotions. 2015 International Conference on Affective Computing and Intelligent Interaction (ACII).

Perception of emotion through facial expressions and head motion is of interest to both psychology and affective computing researchers. However, very little is known about the importance of each modality individually, as they are often treated together rather than separately. We present a study which isolates the effect of head motion from facial expression in the perception of complex emotions in videos. We demonstrate that head motions carry emotional information that is complementary rather than redundant to the emotion content in facial expressions. Finally, we show that emotional expressivity in head motion is not limited to nods and shakes and that additional gestures (such as head tilts, raises and general amount of motion) could be beneficial to automated recognition systems.

Anis, K., et al. (2018). Detecting Depression Severity by Interpretable Representations of Motion Dynamics. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018).

Recent breakthroughs in deep learning using automated measurement of face and head motion have made possible the first objective measurement of depression severity. While powerful, deep learning approaches lack interpretability. We developed an interpretable method of automatically measuring depression severity that uses barycentric coordinates of facial landmarks and a Lie-algebra based rotation matrix of 3D head motion. Using these representations, kinematic features are extracted, preprocessed, and encoded using Gaussian Mixture Models (GMM) and Fisher vector encoding. A multi-class SVM is used to classify the encoded facial and head movement dynamics into three levels of depression severity. The proposed approach was evaluated in adults with history of chronic depression. The method approached the classification accuracy of state-of-the-art deep learning while enabling clinically and theoretically relevant findings. The velocity and acceleration of facial movement strongly mapped onto depression severity symptoms consistent with clinical data and theory.

Bilakhia, S., et al. (2015). "The MAHNOB Mimicry Database: A database of naturalistic human interactions." Pattern Recognition Letters **66**: 52-61.

People mimic verbal and nonverbal expressions and behaviour of their counterparts in various social interactions. Research in psychology and social sciences has shown that mimicry has the power to influence social judgment and various social behaviours, including negotiation and debating, courtship, empathy and helping behaviour. Hence, automatic recognition of mimicry behaviour would be a valuable tool in various domains, and especially in negotiation skills enhancement and medical help provision training. In this work, we present the MAHNOB mimicry database, a set of fully synchronised, multi-sensory, audiovisual recordings of naturalistic dyadic interactions, suitable for investigation of mimicry and negotiation behaviour. The database contains 11 h of recordings, split over 54 sessions of dyadic interactions between 12 confederates and their 48 counterparts, being engaged either in a socio-political discussion or negotiating a tenancy agreement. To provide a benchmark for efforts in machine understanding of mimicry behaviour, we report a number of baseline experiments based on visual data only. Specifically, we consider face and head movements, and report on binary classification of video sequences into mimicry and non-mimicry categories based on the following widely-used methodologies: two similarity-based methods (cross correlation and time warping), and a state-of-the-art temporal classifier (Long Short Term Memory Recurrent Neural Network). The best reported results are session-dependent, and affected by the sparsity of positive examples in the data. This suggests that there is much room for improvement upon the reported baseline experiments.

Buschmeier, H., et al. (2014). ALICO: A multimodal corpus for the study of active listening. LREC 2014, Ninth International Conference on Language Resources and Evaluation, 26-31 May, Reykjavik, Iceland.

Gunes, H. and M. Pantic (2010). Dimensional Emotion Prediction from Spontaneous Head Gestures for Interaction with Sensitive Artificial Listeners, Berlin, Heidelberg, Springer Berlin Heidelberg.

This paper focuses on dimensional prediction of emotions from spontaneous conversational head gestures. It maps the amount and direction of head motion, and occurrences of head nods and shakes into arousal, expectation, intensity, power and valence level of the observed subject as there has been virtually no research bearing on this topic. Preliminary experiments show that it is possible to automatically predict emotions in terms of these five dimensions (arousal, expectation, intensity, power and valence) from conversational head gestures. Dimensional and continuous emotion prediction from spontaneous head gestures has been integrated in the SEMAINE project [1] that aims to achieve sustained emotionally-colored interaction between a human user and Sensitive Artificial Listeners.

Hammal, Z., et al. (2014). "Interpersonal Coordination of Head Motion in Distressed Couples." IEEE Transactions on Affective Computing **5**(2): 155-167.

In automatic emotional expression analysis, head motion has been considered mostly a nuisance variable, something to control when extracting features for action unit or expression detection. As an initial step toward understanding the contribution of head motion to emotion communication, we investigated the interpersonal coordination of rigid head motion in intimate couples with a history of interpersonal violence. Episodes of conflict and non-conflict were elicited in dyadic interaction tasks and validated using linguistic criteria. Head motion parameters were analyzed using Student's paired t-tests; actor-partner analyses to model mutual influence within couples; and windowed cross-correlation to reveal dynamics of change in direction of influence over time. Partners' RMS angular displacement for yaw and RMS angular velocity for pitch and yaw each demonstrated strong mutual influence between partners. Partners' RMS angular displacement for pitch was higher during conflict. In both conflict and non-conflict, head angular displacement and angular velocity for pitch and yaw were strongly correlated, with frequent shifts in lead-lag relationships. The overall amount of coordination between partners' head movement was more highly correlated during non-conflict compared with conflict interaction. While conflict increased head motion, it served to attenuate interpersonal coordination.

Hammal, Z., et al. (2015). "Head Movement Dynamics during Play and Perturbed Mother-Infant Interaction." IEEE Transactions on Affective Computing **6**(4): 361-370.

We investigated the dynamics of head movement in mothers and infants during an age-appropriate, well-validated emotion induction, the Still Face paradigm. In this paradigm, mothers and infants play normally for 2 minutes (Play) followed by 2 minutes in which the mothers remain unresponsive (Still Face), and then two minutes in which they resume normal behavior (Reunion). Participants were 42 ethnically diverse 4-month-old infants and their mothers. Mother and infant angular displacement and angular velocity were measured using the CSIRO head tracker. In male but not female infants, angular displacement increased from Play to Still-Face and decreased from Still Face to Reunion. Infant angular velocity was higher during Still-Face than Reunion with no differences between male and female infants. Windowed cross-correlation suggested changes in how infant and mother head movements are associated, revealing dramatic changes in direction of association. Coordination between mother and infant head movement velocity was greater during Play compared with Reunion. Together, these findings suggest that angular displacement, angular velocity and their coordination between mothers and infants are strongly related to age-appropriate emotion challenge. Attention to head movement can deepen our understanding of emotion communication.

Kousidis, S., et al. (2013). Exploring annotation of head gesture forms in spontaneous human interaction. Proceedings of the Tilburg Gesture Meeting (TiGeR 2013).

Face-to-face interaction is characterised by head gestures that vary greatly in form and function. We present on-going exploratory work in characterising the form of these gestures. In particular, we define a kinematic annotation scheme and compute various agreement measures among two trained annotators. Gesture type mismatches among annotators are compared against kinematic characteristics of head gesture classes derived from motion capture data.

Morency, L.-P., et al. (2005). Contextual Recognition of Head Gestures. 7th International Conference on Multimodal Interactions, Torento, Italy

Paggio, P. and C. Navarretta (2011). Learning to classify the feedback function of head movements in a Danish Corpus of first encounters. Proceedings of ICMI2011 Workshop Multimodal Corpora for Machine Learning.

This paper deals with the automatic classification of feedback by head movement in the Danish NOMCO corpus, a collection of dyadic interactions in which speakers get to know each other for the first time. The results show that by using a combination of features related to the shape of head movements and facial expressions, together with features of the words these gestures are related with, good results (an F-score of 0.72) are achieved in distinguishing head movements used to express feedback from those that serve a different communicative function. Moreover, the distinction between feedback give and feedback elicit, can also be learn  with very good accuracy (an F-score of 0.913), although this result should be taken with caution due to the fact that one of the behaviours is much more dominant than the other in the corpus

Wlodarczak, M., et al. (2012). Listener head gestures and verbal feedback expressions in a distraction task. Proceedings of the Interdisciplinary Workshop on Feedback Behaviors in Dialog, INTERSPEECH2012 Satellite Workshop.

We report on the functional and timing relations between head movements and the overlapping verbal-vocal feedback expressions. We investigate the effect of a distraction task on head gesture behaviour and the co-occurring verbal feedback. The results show that head movements overlapping with verbal expressions in a distraction task differ in terms of several features from a default, non-perturbed conversational situations, e.g.: frequency and type of movement and verbal to nonverbal display ratios.