Expression training for complex emotions using facial expressions and head movements

Andra Adams
Computer Laboratory
University of Cambridge
Cambridge, UK
Email: Andra.Adams@cl.cam.ac.uk

Peter Robinson
Computer Laboratory
University of Cambridge
Cambridge, UK
Email: Peter.Robinson@cl.cam.ac.uk

***Abstract*—Imitation is an important aspect of emotion recognition. We present an expression training interface which evaluates the imitation of facial expressions and head movements. The system provides feedback on complex emotion expression, via an integrated emotion classifier which can recognize 18 complex emotions. Feedback is also provided for exact-expression imitation via dynamic time warping. Discrepancies in intensity and frequency of action units are communicated via simple graphs. This work has applications as a training tool for customer-facing professionals and people with Autism Spectrum Conditions.**

Index Terms*—*affective computing*; *emotion recognition

I. Introduction

Facial expressions [1] and head motions [2] are important modalities that humans use to communicate their mental states to others. Interestingly, the human ability to correctly recognize emotions is believed to be dependent on our ability to imitate the expressions presented to us by others, a theory originally proposed by Lipps in 1907 [3] and supported by many others since then [4]–[7].

Thus we present an expression training interface which encourages and evaluates the imitation of facial expressions and head motions. The system provides users with feedback on the success of their face and head movements in (a) conveying complex emotions, and (b) precisely imitating expressions.

Expression training can be useful as an intervention for individuals with Autism Spectrum Conditions (ASC) [8]–[11], or as a training tool for neurotypical individuals to hone their emotion synthesis capabilities (such as actors or customer service representatives).

II. System Description

A screenshot of the system is given in Figure 1. The system uses a webcam to capture the facial expressions and head movements of the user. The user selects an emotion category and then selects a *target video* to imitate from the provided database of videos belonging to that emotion category [12].

Two animated bar graphs are displayed: one connected to the target video, and one connected to the webcam video. The values in these animated bar graphs change frame-by-frame to correspond with the changing action unit (AU) intensities in the videos. Thus the user is able to visually compare the AU intensities in her own expression to those in the target video.

When ready, the user presses "Record" and then uses her face and head to either (a) portray the selected emotion (in any way she wishes), or (b) precisely imitate the target video.

When she presses "Stop", the recorded frames are analyzed to gather aggregate intensity and frequency statistics for each AU. Based on the difference between the statistics of the imitation and those of the target emotion, an overall score for the imitation is given to the user, along with detailed feedback as outlined in the sections below. Recorded imitations and their corresponding feedback are saved by the system.

Proprioception alone is not enough to induce improvement on facial imitation tasks [13], and thus visual feedback is an important component for expression training. Our system provides the user with visual feedback of her expressions, via the real-time video stream from the webcam (Figure 1 (ii)) and via the ability to replay the videos of previous imitation attempts (Figure 1 (vii)).

Slowing down the presentation of stimuli has been found to be beneficial in prior facial expression imitation tasks used in an ASC intervention [9]. Thus our system provides the ability to slow down the playback of the target video. This change in speed does not negatively affect the imitation comparison mechanisms, since one depends on aggregate statistics (emotion imitation) and the other uses dynamic time warping (exact-expression imitation).

A. Emotion Feedback

An emotion classifier which recognizes 18 complex emotions has been integrated into the system. The classifier runs in real-time on the live webcam feed. It uses a sliding window of previous frames to collect the AU data from which the real-time emotion classification is calculated. The emotion recognized in the current sliding window is displayed to the user and updated periodically. The emotions recognized are: *Afraid*, *Angry*, *Ashamed*, *Bored*, *Disappointed*, *Disgusted*, *Excited*, *Frustrated*, *Happy*, *Hurt*, *Interested*, *Joking*, *Proud*, *Sad*, *Sneaky*, *Surprised*, and *Worried*.

Detailed feedback for emotion expression is provided to the user when she records her imitation. Feedback is given via bar graphs that depict any large differences in particular AU intensities between the imitation and the target emotion, as can be seen in Figure 1 (vi).

978-1-4799-9953-8/15/$31.00 ©2015 IEEE

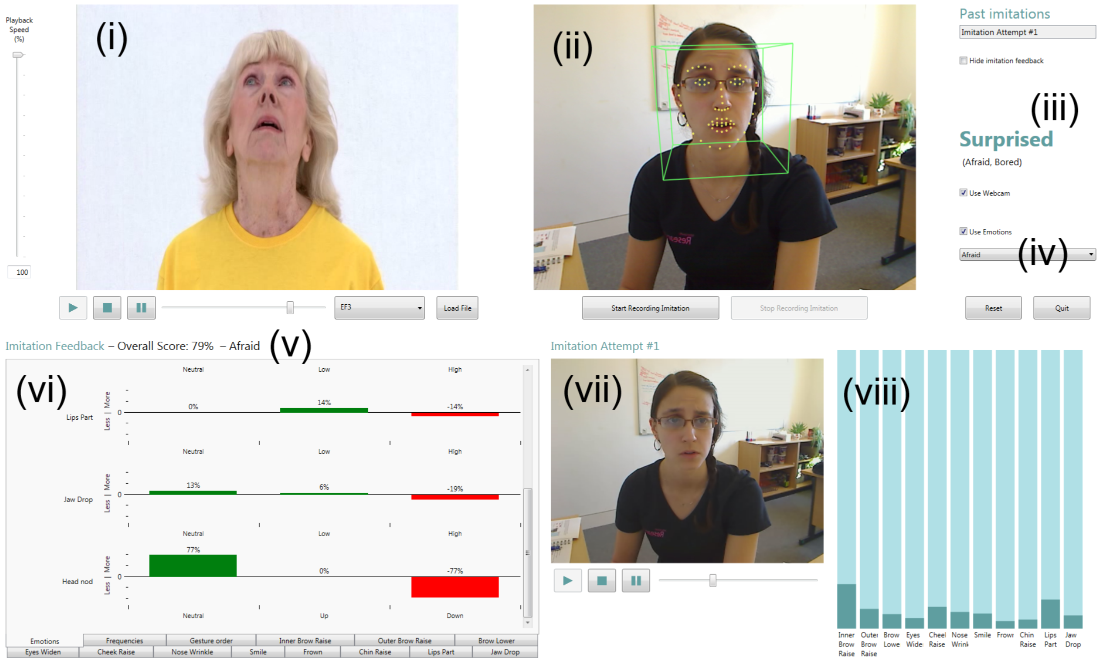

Fig. 1. A screenshot of the expression training interface, with detailed imitation feedback showing. **(i)** A target video is selected. **(ii)** The webcam feed is analyzed in real-time. When the subject is ready, she can record her imitation attempt. **(iii)** The real-time emotion classification label is displayed. **(iv)** The subject can select a target emotion from 18 emotion categories. **(v)** The overall emotion classification label for the most recent recorded attempt is displayed. **(vi)** Large differences in the intensities of particular action units are given as feedback. Green represents "Do more", red represents "Do less". **(vii)** Recorded imitations can be re-watched. The list of past imitations is in the upper right corner of the interface. **(viii)** The action unit intensities for each frame of the recorded imitation from **vii** are displayed in an animated bar graph as the imitation video plays.

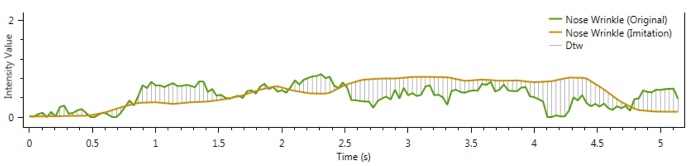

Fig. 2. An example of the visual feedback presented to the user to depict the dynamic time warping between the imitation and the target video.

B. Exact-Expression Feedback

The system also gives feedback on how precisely the user has imitated the exact expressions in the target video. Using the *NDTW* package[1], the system performs dynamic time warping (minimizing the total distance across all action units simultaneously) to determine the frame-by-frame differences in action unit intensities. These differences are presented to the user visually through simple line graphs (Figure 2).

III. Technical Contributions

This is the first real-time system that performs classification of facial expressions and head motions for such a wide range of categorical complex emotions. Previous systems have focused only on basic emotions [14], or have investigated only a small subset of complex emotions (such as el Kaliouby et al. [15] who classified six cognitive mental states, or Littlewort et al. [16] who looked only at pain).

[1]NDTW package available at: https://github.com/doblak/ndtw

Furthermore, this is the first system to give feedback to users for the purpose of improving emotional expression. Previous systems have focused only on static images [17], or have not been designed to coach emotional expression (despite analyzing features that carry emotional content, such as nodding and smiling) [18].

IV. Experiment Design

The system has not yet been studied to determine its effect on imitation abilities. At the conference we will be exploring the question "does practising with our system actually improve imitation ability?" by collecting the imitation scores of all participants, particularly targeting repeated imitation attempts.

V. Conclusion

We have presented an expression training interface which provides feedback on the expression of complex emotions through facial expressions and head movements. The interface also provides feedback on exact-expression imitation which compares the precise expressions in the target video to those of the imitation. The system has applications as an intervention for ASC or as a training tool for actors and other customer-facing professionals.

Acknowledgement

This work is supported by the Gates Cambridge Trust. The authors would like to thank David Dobias for many helpful discussions on this research.

REFERENCES

[1] J. Whitehill, M. S. Bartlett, and J. R. Movellan, "Automatic facial expression recognition," *Social Emotions in Nature and Artifact*, vol. 88, 2013.

[2] Z. Hammal and J. F. Cohn, "Intra-and interpersonal functions of head motion in emotion communication," in *Proceedings of the 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges*. ACM, 2014, pp. 19–22.

[3] T. Lipps, "Das wissen von fremden ichen," *Psychologische untersuchungen*, vol. 1, no. 4, pp. 694–722, 1907.

[4] L. M. Oberman, P. Winkielman, and V. S. Ramachandran, "Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions," *Social neuroscience*, vol. 2, no. 3-4, pp. 167–178, 2007.

[5] P. M. Niedenthal, M. Brauer, J. B. Halberstadt, and Å. H. Innes-Ker, "When did her smile drop? facial mimicry and the influences of emotional state on the detection of change in emotional expression," *Cognition & Emotion*, vol. 15, no. 6, pp. 853–864, 2001.

[6] H. G. Wallbott, "Recognition of emotion from facial expression via imitation? some indirect evidence for an old theory," *British Journal of Social Psychology*, vol. 30, no. 3, pp. 207–219, 1991.

[7] T. M. Field, R. Woodson, R. Greenberg, and D. Cohen, "Discrimination and imitation of facial expression by neonates," *Science*, vol. 218, no. 4568, pp. 179–181, 1982.

[8] S. Bölte, S. Feineis-Matthews, S. Leber, T. Dierks, D. Hubl, and F. Poustka, "The development and evaluation of a computer-based program to test and to teach the recognition of facial affect," *International Journal of Circumpolar Health*, vol. 61, 2002.

[9] C. Tardif, F. Lainé, M. Rodriguez, and B. Gepner, "Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial–vocal imitation in children with autism," *Journal of Autism and Developmental Disorders*, vol. 37, no. 8, pp. 1469–1484, 2007.

[10] K. A. Loveland, B. Tunali-Kotoski, D. A. Pearson, K. A. Brelsford, J. Ortegon, and R. Chen, "Imitation and expression of facial affect in autism," *Development and Psychopathology*, vol. 6, no. 03, pp. 433–444, 1994.

[11] J. A. DeQuinzio, D. B. Townsend, P. Sturmey, and C. L. Poulson, "Generalized imitation of facial models by children with autism," *Journal of applied behavior analysis*, vol. 40, no. 4, pp. 755–759, 2007.

[12] H. O'Reilly, D. Pigat, S. Fridenson, S. Berggren, S. Tal, O. Golan, S. Blte, S. Baron-Cohen, and D. Lundqvist, "The EU-Emotion Stimulus Set: A validation study," *In press*, 2015.

[13] R. Cook, A. Johnston, and C. Heyes, "Facial self-imitation objective measurement reveals no improvement without visual feedback," *Psychological science*, vol. 24, no. 1, pp. 93–98, 2013.

[14] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. Scherer, "The first Facial Expression Recognition and Analysis challenge," in *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011)*. IEEE, 2011, pp. 921–926.

[15] R. El Kaliouby and P. Robinson, "Real-time inference of complex mental states from facial expressions and head gestures," in *Real-time vision for human-computer interaction*. Springer, 2005, pp. 181–200.

[16] G. C. Littlewort, M. S. Bartlett, and K. Lee, "Faces of pain: automated measurement of spontaneousallfacial expressions of genuine and posed pain," in *Proceedings of the 9th international conference on Multimodal interfaces*. ACM, 2007, pp. 15–21.

[17] K. Ito, H. Kurose, A. Takami, and S. Nishida, "iface: A facial expression training system," in *Analysis, Design, and Evaluation of Human-Machine Systems*, vol. 10, no. 1, 2007, pp. 466–470.

[18] M. E. Hoque, M. Courgeon, J.-C. Martin, B. Mutlu, and R. W. Picard, "Mach: My automated conversation coach," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 2013, pp. 697–706.

978-1-4799-9953-8/15/$31.00 ©2015 IEEE