Do Users Anticipate Emotion Dynamics in Facial Expressions of a Virtual Character?

[1] Matthieu Courgeon *[2]Michel-Ange Amorim [2]Caroline Giroux [1] Jean-Claude Martin

[1] LIMSI-CNRS
Université Paris-Sud
courgeon@limsi.fr
martin@limsi.fr
www.limsi.fr/Individu/[courgeon;martin]

[2] UR CIAMS - Motor Control and Perception
Univ. Paris-Sud * Institut Universitaire de France
michel-ange.amorim@u-psud.fr
caroline.giroux@orange.fr
www.staps.u-psud.fr/fr/recherche/ciams/cmp.html

Abstract

Although expressive and social virtual agents are spreading around, little is known about user's cognitive processing of the dynamic aspects of facial expressions while watching agents' emotional reactions. The present psychophysical study evaluates how users anticipate emotional dynamics in facial expressions of a virtual character. Users memorized the emotional intensity expressed by the virtual face that reappeared statically after a brief mask period, in order to decide if this test intensity was greater or lesser than the actual expressed intensity. Results indicate that subjects anticipate a return to a more neutral expression after an intense expression. However, the size of this emotional momentum varies depending on emotion. These findings may help to improve the design of dynamic facial expression of emotions in virtual faces, especially for camera cutting effects on scenes with multiple social agents' interactions.

Keywords: expressive social agents, automatic camera cutting, facial expressions, emotional momentum

1. Introduction

Expressive and social virtual agents are nowadays used in several domains, e.g. perceptive tests, video games, e-learning, virtual presenters, and so on. These agents are expressing emotions, reacting to events, and speaking. When several agents are displayed, camera goes from one agent to another. Although automatic camera cutting (e.g., shot/reserve-shot pattern [1]) is an important issue for conversational scenes [2], little is known about user's cognitive processing of the dynamic aspects of facial expressions while watching agents' emotional reactions (to dialogue). Yet, the issue of affective chronometry [3] must be addressed when modeling emotional facial expressions on virtual agents (e.g. [4] for setting the parameters of the temporal subcomponents such as latency, rise time, magnitude, duration, and offset). In the present study, we provide psychophysical data about the effect of emotion type and expressed emotional intensity on the perception and memory for facial expressions, which may turn useful for setting the temporal parameters of virtual agents dynamics (e.g., to trigger camera cut).

Social neuroscience has well-established that during communicative behavior our brain obtains knowledge about other people's emotional states from covert imitation [5]. This emotion mirroring is part of an emulation process [6] that is crucial to understand and anticipate the behavior of our conspecifics.

Representational momentum [7] reflects our natural tendency to mentally extrapolate the motion of a target stimulus into the future, which causes to misremember an event as having continued beyond the point at which it ceases to be visible. Representational momentum is a by-product of the emulation process that assists our perceptual system in real-time[8] in order to preserve the continuity of an event despite dramatic visual changes such as when cutting on motion [9] or during event occlusion [10]. The evidence of an emotional momentum for facial expression is scarce and contradictory in the literature. One study reported an emotional momentum towards perceiving more intensified expressions than what actually appear [11], whereas another indicated that observers anticipate a return to a more neutral expression [12], which is a backwards memory displacement. Surprisingly, although emotion (anger, disgust, fear, happiness, sadness, and surprise) was manipulated in the former study, it did not affect the results.

Our study presents an experiment aiming at evaluating this emotional. Contrary to previous studies we used facial expression of emotion from real-time animated virtual faces rather than sequences of face photographs morphed [11]. Animation of virtual faces was shown to allow for identification of basic emotions similarly to natural faces [13].

2. MARC's facial animation

Our interactive facial animation platform, MARC, relies on GPU technology to perform fast realistic animation. It renders high quality virtual face (20.000 polygons) using light simulation techniques to enhance visual realism. Facial expressions are performed at 60 frames per second using extended MPEG4 specifications, i.e. blending faces in 3D. Skin is highly detailed and expressive wrinkles are computed automatically [14]. Figure 1 provides an example of MARC's face.

Figure 1. MARC expressing joy

MARC was already used in several experimental setups [15]. In this study, the user interface displayed during the experiment was integrated within the animation module to ensure maximum time precision. Indeed, some pictures were displayed during a very short period of 250ms. Expression dynamics was the important subject of our experiment, time precision was crucial.

3. Experiment

3.1 Procedure

17 subjects participated in this experiment, 8 men and 9 women, between 20 and 53 years old. They were instructed to memorize the Emotional Intensity (EI) expressed by our virtual face who reappeared as a static face after a 250 ms (retention interval) mask period (pixelated face), in order to decide if this Test EI was located *« after »* (greater than) or *« before »* (lesser than) the actual expressed EI. A typical experimental trial is illustrated in Figure 2 starting with a 2 sec warning signal during which participants could press the space bar to make a pause. Participants responded to the Test EI using an AZERTY keyboard ("before" = A key; "after" = P key). They were instructed to respond as quickly and accurately as possible. Next trial started automatically after response or if reaction time to the Test EI onset exceeded 3 sec. The Test EI corresponded to n-30%, n-15%, n, n+15%, n+30% given that *n* was the Expressed EI to be memorized (50%, 70%, or 90% depending on trial).

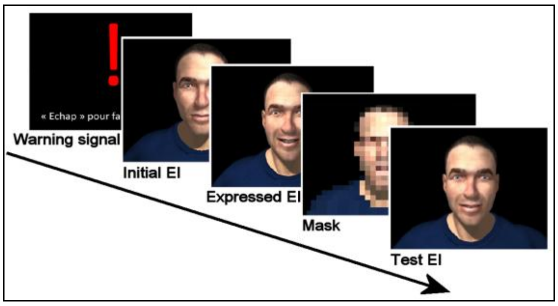

Figure 2. Flow chart of a typical trial. 2s warning signal. Smooth transition between a neutral expression and a target expression. 250ms mask. Then the test expression.

The experimental session began with practice trials, not included in the data set. Both practice and experimental trials were presented in a different random order for each participant, and without feedback. In summary, each participant performed 360 experimental trials: 3 Expressed EI × 6 Emotions (anger, disgust, fear, happiness, sadness, and surprise) × 5 Test EI × 4 Repetitions. For each emotion displayed, we defined a facial expression using Ekman's descriptions [16]. There were 30 trials blocks separated by automatic pauses.

3.2 Dependent variables

Points of subjective equality (PSEs) were computed by fitting sigmoid curves to the responses of each participant at the various Test EI. We used a cumulative logistic function in order to infer the memorized EI. The Test EI was either the same as the Expressed EI, or that 15% or 30% lesser or greater. PSE values corresponded to the memory bias reflecting the accuracy of the responses (see Figure 2): a value of PSE equal to 0% indicates that the subject's performance is accurate. A memory bias equal to 10% indicates that the memorized EI is shifted in the direction of the 100% EI, about 10% from the actual Expressed EI. Such a positive "memory bias" reflects a representational momentum effect.

Just noticeable difference (JND) values were computed, in order to evaluate the precision of the remembered EI (see Figure 3). JND is the minimum amount by which a stimulus (EI) must be changed to produce a noticeable perceptual difference. It is the difference between two stimuli (expressed vs. remembered EI) that is detected as often as it is undetected.

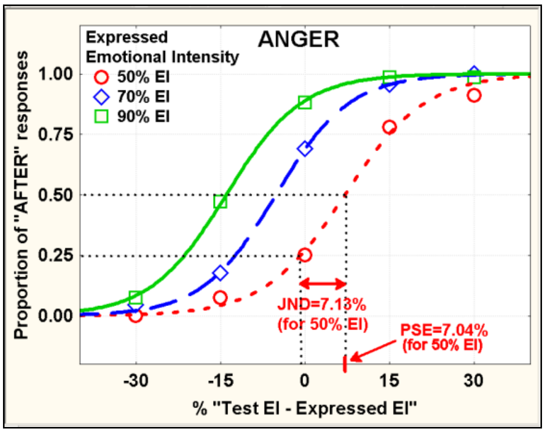

Figure 3. Example of PSE and JND computation**.**

In other words, the greater the JND values, the shallower the sigmoid slope (see Figure 3), and the less precise the remembered EI. PSE and JND provide two independent and complementary measures of performance.

3.3 Results

As a first step, we performed separate ANOVAs on mean proportion of *« after »* responses (Test EI is *greater than* memorized EI) for each Emotion (6) with Expressed EI (3) and Test EI (5) treated as within-subject factors. Then, we computed PSEs and JNDs for each Emotion and Test EI as illustrated in Figure 2 for Anger. The ANOVAs showed that the effect of Test EI on response varied significantly with Expressed EI for each Emotion but Hapiness ($F(8,128)=1.28$, $p=.26$) and Sadness ($F(8,128)=1.86$, $p=.07$).

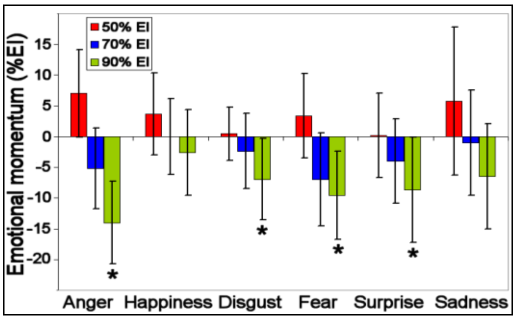

Figure 4. Mean emotional momentum (PSE) for each Expressed EI as a function of Emotion. Error bars represent ± 1 JND. Stars indicate when PSE>JND.

Figure 4 illustrates the general trend of forward memory displacements for 50% EI and increasing backwards memory shifts with greater Expressed EI. However, one might consider that proper emotional momentum takes place when greater than response precision (PSE>JND). Then, results indicate that emotional momentum is observed mainly for 90% EI and especially for Anger, Disgust, Fear and Surprise (see Figure 4).

4. Discussion

Our result show comparable effects of expressed intensity on every emotions category. These results are consistent with Thornton study [12]: with higher intensities, subjects anticipate a return to a neutral expression. We can assume that physical limits to facial deformations might explain that subject perceive the high intensity stimulus as the maximum deformation and anticipate a return to the neutral expression.

We observed significant perceptual differences between emotions. For example, this effect is greater with Anger expression than Joy expression. We can postulate that this is most likely related to the facial expression we used: in Joy expression, the mouth width provides a perceptual cue for intensity that is not available in Anger expression. Thus, in our next study we will use several expressions for each emotion, and each expression will be quantified in terms of facial deformation [17]. Moreover, if we postulate that subjects focus on facial features, it might imply that we did not only measure emotional momentum, but a standard representational momentum of facial movements. To test this hypothesis, we intend to conduct a similar experiment but displaying the face upside down (as literature suggests faces are less recognized when displayed reversed [18]) and using gaze tracking technologies to identify which facial features are observed and to what extent.

5. Conclusion

In this paper we presented a study on perception of emotional momentum in facial expression of virtual characters. Our results tend to validate previous results [12]. Human perception seems to anticipate a return to a more neutral expression after an intense expression. However, we observe inter emotions differences. For example, Joy expression seems to be less misremembered than Anger expression. Our future work will focus on differentiating which effect is related to emotional perception, and which one is related to a visual representational momentum of facial features.

The results presented in this paper help to design facial expressions dynamics of emotions, especially for camera cutting effects on multiple social agents' interactions.

References

[1] S.D. Katz. Film directing shot by shot: Visualizing from concept to screen.

[2] L. He, M.F. Cohen and D.H. Salesin. The virtual cinematographer: a paradigm for automatic real-time camera control and directing. *SIGGRAPH*, 1996

[3] R.J. Davidson. Affective style and affective disorders: perspectives from affective neuroscience. *Cognition and Emotion*, 307-320, 1998

[4] N. Mana and F. Pianesi. Modelling of emotional facial expressions during speech in synthetic talking heads using a hybrid approach. *AVSP2007*, 2007

[5] J.A.C.J. Bastiaansen, M. Thioux and C. Keysers. Evidence for mirror systems in emotions. *Phil. Trans. Royal Soc.*, 2391-2404, 2008

[6] R. Grush. The emulation theory of representation. *Behavioral & Brain Science*, 377-442, 2004

[7] I.-M. Thornton and T.-L. Hubbard. Representational Momentum: New Findings, New Directions. *Psych. Press Hove*, 2002

[8] M. Wilson and G. Knoblich. The case for motor involvement in perceiving conspecifics. *Psych. Bull.* 460-473:2005

[9] M. Jarraya, M.-A. Amorim and B. Bardy. Optical flow and viewpoint change modulate the perception and memorization of complex motion. *Perception & Psychophysics*, 951-961:2005

[10] M. Graf, B. Reitzner, C. Corves, A. Casile, M. Giese and W. Prinz. Predicting point-light actions in real-time. . *NeuroImage*, 2:22-32, 2007

[11] S. Yoshikawa and W. Sato. Dynamic facial expressions of emotion induce representational momentum. *Cognitive, Affective, & Behav. Neurosc.*, 25-31 2008

[12] I.M. Thornton, *The perception of dynamic human faces.* PhD Thesis of University of Oregon, Eugene, 1998

[13] J. Kätsyri and M. Sams. The effect of dynamics on identifying basic emotions from synthetic and natural faces. *Int. J. of Human-Computer Studies*, 233-242, 2008

[14] M. Courgeon, S. Buisine and J.-C. Martin. Impact of Expressive Wrinkles on Perception of a Virtual Character's Facial Expressions of Emotions. *Proceedings of the 9th Int. Conf. on Intelligent Virtual Agents*, Amsterdam, The Netherlands, 2009

[15] M. Courgeon, C. Clavel and J.-C. Martin. Appraising Emotional Events during a Real-time Interactive Game. *ICMI09 International workshop on Affective-aware Virtual Agents and Social Robots*, Boston, MA, USA 2009

[16] P. Ekman and W.V. Friesen. Unmasking the face. A guide to recognizing emotions from facial clues. *Englewood Cliffs*, New Jersey: P-H. , 1975

[17] C. Alvino, C. Kohler, F. Barrett, R.E. Gur, R.C. Gur and R. Verma. Computerized measurement of facial expression of emotions in schizophrenia. *J of Neurosc. Methods*, 350-361, 2007

[18] C. Jacques, O. D'arripe and B. Rossion. The time course of the inversion effect during individual face discrimination. *Journal of Vision*, 8:1-9, 2007