In Proceedings of 1^st Workshop on Emotion and Computing at KI 2006, 29th Annual Conference on Artificial Intelligence, June, 14-19, 2006, Bremen, Germany

Psychologically Grounded Avatars Expressions

Marco Paleari and Christine Lisetti

Affective Social Computing Group,
Eurecom Institute, 2229 route des crêtes,
Sophia Antipolis

Abstract. Progress in computer graaphics over the last decade has rendered the creation of believable anthropomorphic graphical avatars possible. Issues in rendering these animated graphical avatars believable and engaging during Human-Computer Interaction still remain. In this article, we focus on the animation of avatar's facial expressions. We explain how we created our animation on Scherer's theory of emotion generation associated with facial expressions to create five facial expression animations (happiness, disgust, sadness, fear and anger) that are congruent with Scherer's theory. We discuss the specific steps and issues we followed as well as the evaluation results of a user study we conducted.

1 Introduction

Progress in computer graphics over the last decade has rendered the creation of believable anthropomorphic graphical avatars possible. Issues in rendering these animated graphical avatars believable and engaging during Human-Computer Interaction still remain.

In this article, we focus on the animation of avatar's facial expressions from a commercially a available generator software (Haptek, [1] which we evaluated as providing the best graphical rendering for believable anthropomorphic avatars. We created our animation on Scherer's theory of emotion generation associated with facial expressions [2–4] because we are concurrently developing an emotion-based architecture for intelligent social agents [5, 6] and we want to be able to link dynamically internal emotion-like states to external facial expressions in a manner that will be as close as possible to Scherer's theory of emotion generation.

We have currently created five facial expression animations (happiness, disgust, sadness, fear and anger) that are congruent with Scherer's theory using Haptek generator and we discuss the steps and issues involved in the process. We also show the results of an initial evaluation of the avatar's expression in terms of recognition and believability compared to that of a human.

2 Haptek avatars

Haptek avatars have been developed to represent believable human faces, and Haptek tools are commercially available software generators to enable the insertion of avatars in applications or web pages. Haptek animation is based on dedicated technology, similar to MPEG-4 FAP (Facial Action Parameters). There are different levels of control over Haptek avatars: from the control of global facial expressions, morph and position of the avatar to the control of basic facial movements. Basic control of the avatar is possible by Haptek hypertext technology. Through hypertext one can control text to speech, avatar position and launch Haptek switches. Switches are collections of states which represent the still expression of the avatar as well as its morphs in term of combinations of facial parameters defined by Haptek. Through switches one can, therefore, control the evolution of states over time as well as the softness of the transitions from one state to another, i.e. the evolution of the avatar expression.

3 Scherer psychological theory

We based our work on Scherer's multi-level process theory of emotions. According to Scherer [2–4], emotions are experienced depending upon the result of the individual's evaluation or appraisal of the events surrounding him in terms of their significance for one's desires and aversions. Scherer describes this process of appraisal as a process of sequential evaluation of parameters that he calls Sequential Evaluation Checks (SECs). The number and the nature of these checks have changed over time but they always refer to four categories: 1) *relevance*, or how relevant the event is for someone, 2) *implications*, or what the implications of this event are, 3) *coping potentials*, or how one can cope with these consequences and 4) *normative significance*, or what the significance of this event is with respect of one's self-concept and to social norms and values.

Sequential Evaluation Checks (SECs) are evaluated one after the other in temporal order from those representing *relevance* to the *significance* ones. For some of them Scherer also gives predictions about the corresponding facial expressions in terms of Ekman's [7] facial Action Units (AUs), the smallest independent facial muscle mouvement/action possible in the human face. Combining properly these predictions, it is possible to, on the one hand, recognize displayed emotions, and on the other hand display facial expressions (and the corresponding emotion), not in terms of a label (happiness, sadness etc.) as many computational animations do, but rather in terms of the underlying SECs structure of the emotion.

The use of SECs will also play a fundamental role in our cognitive-affective archicture.

4 Avatar Animation based on Scherer's SECs

It is not straightforward to automatically convert SECs to Haptek parameters, and we have identified four steps for the process: 1) convert each SEC to AUs; 2) convert AUs to Haptek parameters; 3) find appropriate intensities for the AUs; 4) exploit the temporal and intra-SEC correlation adapting AUs intensities.

For *Step 1 - Converting SECs to AUs*, we used tables directly extrapolated from Scherer's [2–4]. The other three points are fairly more complicated.

For *Step 2 - Converting AUs to Haptek parameters*, we created a conversion table and software designed to create switches representing single AUs at different intensities. A problem is that some movements are not really describable in term of Haptek parameters (we would need sub-parameters). To resolve this, we designed the AUs representing movements that are at least similar to the original ones. The average quality of the conversion has been evaluated as very good by an informal internal questionnaire.

For *Step 3 - Finding the right intensities for the AUs*, we based our decisions on the Scherer theory, [2–4] and converted SECs intensities to the Ekman scale from 'a' (min. intensity) to 'e' (max. intensity).

For *Step 4 - Exploiting temporal information*, the problem is that Scherer's theory does not exploit any intra-SEC or temporal correlation. It does not explain how the SEC sub-expressions are linked to each other over time, nor how long one single sub-expression should be displayed on the face before it vanishes. We referred to standard videos contained in a database of different actors expressing six universal expressions [7]. As it became evident that some AUs must be activated in time over different SECs and have their intensities reduced in a fluid way, we relaxed the constraints over AU intensities and adapted the intensity to reach fluid expressions. The transitions for *fear* are shown in Figure 1 and all of the five expressions will also be demoed at the KI Workshop Demonstration Session applied to an e-tutoring context [8].

Fig. 1. Possible evolution of the expression for Fear according to Scherer theory

5 Recognition and Believability Evaluation User Study

We conducted a user study to compare the resulting expressions against the ones of human actors and generic expressions developed by Haptek. The results of the study of five different expressions (happiness, disgust, sadness, fear and anger), experimented on a group of 16 people, are shown in figure 2. The subjects were first asked to recognize the shown emotion using a closed questionnaire and then to express their opinion, by chosing a mark between 1 and 5, about the expressions shown by one actor, an avatar tuned by standard Haptek expressions and another tuned with the designed parameters.

The scores are good in that the shown emotions were recognized in the 94% (see table in Fig. 3) of the cases and the believability of the developed expression

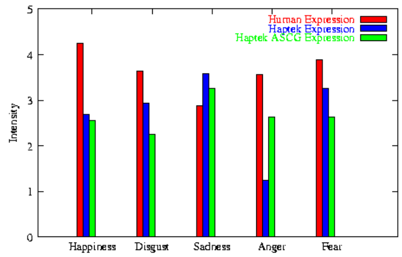

Fig. 2. Expressions believability

shown / recogn.	Happin.	Disgust	Sadness	Fear	Anger
Happiness	**100%**	0%	0%	0%	0%
Disgust	0%	**63%**	0%	0%	38%
Sadness	0%	0%	**100%**	0%	0%
Fear	0%	6%	0%	**94%**	0%
Anger	0%	13%	0%	0%	**88%**

Fig. 3. Recognition scores

has been judged in a similar way to the one of human expression and of the Haptek default expressions (see [9] for further details).

6 Conclusions and future work

This study shows the feasibility of facial emotion expression based on Scherer's component process theory. Expresions created through this paradigm have shown to be believable and understandable. Nevertheless this study has been developed over only five expressions and will be expanded to other combinations of SECs, as well as linked with an affective-cognitive computational architecture.

References

1. Unknown: Haptek website: www.haptek.com (2006)
2. Scherer, K.: Emotions can be rational. Social Science Information **24**(2) (1985) 331–335
3. Scherer, K.R.: Emotion as a process: Function, origin and regulation. Social Science Information **21** (1982) 555–570
4. Scherer, K.R.: Toward a dynamic theory of emotion: The component process model of affective states. Geneva Studies in Emotion and Communication, 1(1), 1-98 (1987)
5. Lisetti, C.L., LeRouge, C.: Affective computing in tele-home health: design science possibilities in recognition of adoption and diffusion issues. In: In Proceedings 37th IEEE Hawaii International Conference on System Sciences, Hawaii, USA (2004)
6. Lisetti, C.L., Gmytrasiewicz, P.J.: Emotions and personality in agent design. In: In Proceedings of AAMAS 2002. (2002)
7. Ekman, P., Friesen, W.V., Hager, J.C.: Facial Action Coding System Invistagator's Guide. A Human Face (2002)
8. Paleari, M., Lisetti, C.L., Lethonen, M.: VALERIE: a virtual agent for a learning environment, reacting and interacting emotionally. In: AIED 2005, 12th International Conference on Artificial Intelligence in Education, Amsterdam, The Netherlands (2005)
9. Grizard, A., Paleari, M., Lisetti., C.: Adaptating psychologically grounded facial emotional expressions to different platforms. (2006)