Expressing Complex Mental States Through Facial Expressions

Xueni Pan[1], Marco Gillies[1], Tevfik Metin Sezgin[2], and Celine Loscos[1,3]

[1] Department of Computer Science, University College London, London, UK
s.pan@cs.ucl.ac.uk
[2] Department of Computer Science, Cambridge University, Cambridge, UK
[3] Departament d'Informtica i Matemtica Aplicada, Universitat de Girona, Spain

A face is capable of producing about twenty thousand different facial expressions [2]. Many researchers on Virtual Characters have selected a limited set of emotional facial expressions and defined them as basic emotions, which are universally recognized facial expressions. These basic emotions have been well studied since 1969 and employed in many applications [3]. However, real life communication usually entails more complicated emotions. For instance, communicative emotions like "convinced", "persuaded" and "bored" are difficult to describe adequately with basic emotions. Our daily face-to-face interaction is already accompanied by more complex mental states, so an empathic animation system should support them. Compared to basic emotions, complex mental states are harder to model because they require knowledge of temporal changes in facial displays and head movements as opposed to a static snapshot of the facial expression. We address this by building animation models for complex emotions based on video clips of professional actors displaying these emotions.

The first step of our work is to extract the facial and head movements from video clips. We have adopted the recognition framework proposed by [4] to recognize head and facial displays, which allows facial displays to be based on not just the current facial action, but also on a pre-determined number of previous facial actions. Using the framework in [4], a commercial face tracker is used to locate and track 24 landmarks on the face. These are then mapped into facial actions and displays using Hidden Markov Models (HMMs) trained on video clips from the Mind Reading DVD compiled by researchers at the University of Cambridge Psychology Department [1].

The likelihoods of the facial displays, computed by the HMMs from the input video sequences, are used to drive the animation of our virtual characters. Six head action displays(*head tilt, head turn, head forward, head backward, head shake* and *head nod*) and three three facial displays (*lip pull, lip pucker*, and *brow raised*) are extracted for every frame and applied on the Virtual Character. The first four of the HMM outputs for head movement were translated directly into rotation about a single axis whereas *head shake* and *head nod* were translated into periodic animations. The facial displays are applied on the Virtual Characters as morph targets.

As described above complex mental states differ from basic emotions in that they cannot be effectively recognized from static facial expressions, only from

A. Paiva, R. Prada, and R.W. Picard (Eds.): ACII 2007, LNCS 4738, pp. 745–746, 2007.

Fig. 1. Animation result of the emotion *interest*

facial movements over time. This implies that we must use different animation techniques, ones based on time series rather than static methods such as morph targets or principal component analysis. The approach we take in this paper is to base our animations on a corpus of capture facial motion data using a data structure called a motion graph[5]. A motion graph enables new animations to be generated from a corpus of animation data by resequencing animation clips in different orders. It is a directed graph structure in which edges are motion clips and nodes are possible transition points between clips. These transition points are chosen so that they ensure smooth transitions between different clips. New animations are created by walking the graph. We use a dynamic programming method to calculate a value function over edges which is used to generate graph walks that give a good representation of the desired mental state.

Six communicative complex mental states were chosen based on their importance in giving feedback during conversations: *agreement, disagreement, concentration, interest, thinking* and *uncertainty*. We used eighteen video clips, three for each emotion, from the Mind Reading database. We first extracted the facial movements for each video clip, then applied those on the Virtual Character as separated animations. From the eighteen animations we then created a motion graph labeled with the six emotions. Finally, from the motion graph, we constructed animations for each of the six complex emotions by walking the graph. Figure 1 is an example of our final results of the emotion *interest*.

References

1. Baron-Cohen, S., Golan, O., Wheelwright, S., Hill, J.: Mind-reading: The interactive guide to emotions (January 2004)
2. Birdwhistell, R.L.: Kinesics and context: essays on body motion communication. Allen Lane The Penguin Press (1971)
3. The Duy Bui. Creating Emotions And Facial Expressions For Embodied Agents. PhD thesis (2004)
4. el Kaliouby, R., Robinson, P.: Real-time inference of complex mental states from facial expressions and head gestures. Real-time vision for HCI, 181–200 (2005)
5. Kovar, L., Gleicher, M., Pighin, F.: Motion graphs. ACM Transactions on Graphics 21(3), 473–482 (2002)