Journal of Visual Languages and Computing 23 (2012) 299–310

Contents lists available at SciVerse ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

Emotional eye movement generation based on Geneva Emotion Wheel for virtual agents ☆

Zheng Li, Xia Mao *

School of Electronic and Information Engineering, Beihang University, Beijing, People's Republic of China

ARTICLE INFO

Article history:
Received 24 June 2010
Received in revised form
28 March 2012
Accepted 17 June 2012
Available online 26 June 2012

Keywords:
Virtual agents
Human–computer interaction
Eye movement synthesis
Computer animation

ABSTRACT

Eye movement plays an important role in face to face communication in that it conveys nonverbal information and emotional intent beyond speech. Our research proposes a computational framework that enables virtual agents to convey different emotional expressions to users through eye movement. We build this framework based on eye movement parameters derived from the Kohn-Canade AU(action unit)-Coded Facial Expression Database as well as real-time eye movement data. We also describe a rule-based approach to generate emotional eye movement based on the Geneva Emotion Wheel. We then present one experiment in which subjects evaluated the emotional eye movement generated by this framework. When our proposed model was employed, the results showed a higher rate of recognition of the agent intended emotion, proving the validity of our approach.

1. Introduction

Virtual agents are an emerging technology in multimodal interfaces. They have been created for the personalisation of user interfaces and for the presentation of a wide variety of tasks, including tutors in e-learning systems, recommenders in e-commerce systems, actors in entertainment systems and partners in chatting systems etc [22,34,30]. They allow for a natural way of presenting information and interacting with users through their multimodal expressions (speech, gesture, facial expression and so on). In order to realize harmonious and lively human–agent interaction, virtual agents should possess the capability for displaying emotions similar to humans. Although many researchers have devoted theirs work to this direction, they were mainly concentrated on facial expression generation, emotional speech synthesis and emotional gesture generation. Nevertheless, eye movement is also a crucial part of human communication. To interact with humans in a rich and natural way, virtual agents need to use this communicative channel effectively. While the role and mechanics of human eye movement are extensively studied, how it might be used effectively by virtual agents to convey different emotional expression to users is not well explored. The goal of our paper is to gain a deeper understanding of how to generate emotional eye movement for virtual agents. We also consider related aspects of behavior that contribute to expressiveness as a whole and add realism to the animation, such as eyebrow, eyelid and nose movement.

The emotions might be categorized into primary emotions and intermediate (secondary) emotions [8]. Typical primary emotions refer to emotions which are supposed to be innate. These include joy, sadness, anger, fear, disgust and surprise. Intermediate emotions are those emotions that are assumed to arise from higher cognitive process, based on an ability to evaluate preferences over outcomes and expectations. They can be represented primary emotions mixture. The Geneva Emotion Wheel (GEW) is a theoretically derived and empirically tested

☆ This paper has been recommended for acceptance by S.-k. Chang.

* Corresponding author. Tel.: +86 10 82316739.
E-mail address: moukyoucn@yahoo.com.cn (X. Mao).

1045-926X/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jvlc.2012.06.001

instrument to measure emotional reactions to objects, events, and situations. The emotions in GEW, including primary emotions (joy, sadness, anger, fear, disgust and surprise) and 14 intermediate emotions, are symmetrically arranged in a wheel shape with the axes being defined by two major appraisal dimensions: high/low appraisal and positive/negative appraisal. In this paper, we propose an approach to synthesize emotional eye movement for primary and intermediate emotions based on GEW. We present our emotional eye movement generation framework based on the MPEG-4 (Moving Picture Experts Group) facial animation parameters (FAPs) utilized for facial animation purpose [23,26]. We derive the FAPs values from the Cohn-Kanade AU(action unit)-Coded facial expression database, as well as real-time eye movement data we have collected. By analyzing the users' opinions in our evaluation procedure, we conclude that our framework significantly enhanced the emotional expression of a virtual agent.

The remaining paper is organized as follows. The next section reviews some of the related research on the virtual agents eye movement. Section 3 presents an overview of our system framework and identifies a subset of FAPs for use in generating eye movement and shape. Then we describes the derivation of eye movement parameters from the AU-Coded facial expression database, as well as real-time eye movement data we have collected. We also describe our approach to generating emotional eye movement for primary and intermediate emotions based on these eye movement and FAPS parameters In Section 4, experimental results, which illustrate the performance of the presented approach, are given. Finally, we propose future research directions.

2. Related work

Being "a window to the mind", the eye and its behavior are tightly coupled with human cognitive processes. Several studies have been carried out, mostly in the Psychology and Physiology fields, to find how eye movement regulates interaction between individuals. Meanwhile, there have been many implementations of eye movement in virtual agent applications. Fukayama et al. proposed an eye movement model that enables virtual agents to convey different expressions to users [12]. Cassell et al. have explored eye engagement during social interactions or discourse between virtual agents. They discussed limited rules of eye engagement between animated participants in conversation [6,7]. Mutlu's research gained a deeper understanding of how eye movement affects people's interactions with virtual agents and how people can design eye movement for effective communication [24]. The research of Poggi and Pelachaud has been concentrated on the study and generation of coordinated linguistic and eye communicative acts, namely the ones performed by blink, eyebrow and eyelid (upper and lower) movement. They proposed a formalism where a communicative eye movement is represented by two elements: a meaning (that corresponds to a set of goals and beliefs that the agent has the purpose to transmit to the interlocutor) and a signal (that is the nonverbal expression of that meaning) [29]. Breton et al. brought forward an eye movement model applied in a real time multimodal dialogue interaction with users. The agents' eye movement control results from the fusion of a rational dialogue engine based on natural language interaction and a multi-user face tracker [5]. Dziemianko et al. presented a technique to automatically synthesis eye blinking from a speech signal. Animating the eyes of a talking head is important as they are a major focus of attention during interaction. The developed system predicts eye blinks from the speech signal and generates animation trajectories automatically employing a "Trajectory Hidden Markov Model" [9]. Wissner et al. described the implementation of a parameterizable gaze behavior system based on psychological notions of human gaze. The resulting gaze patterns and agent behaviors cover a wide range of possible uses due to this parametrization [35]. Bee et al. implemented an eye-gaze based model of interaction to investigate whether flirting tactics help improve first encounters between a human and an agent. Unlike earlier work, they concentrated on a very early phase of human-agent conversation (the initiation of contact) and investigated which nonverbal signals an agent should convey in order to create a favourable atmosphere for subsequent interactions and increase the user's willingness to engage in an interaction with the agent [2]. Lee et al. proposed an approach for synthesizing the trajectory kinematics and statistical distribution of saccadic eye movements in speaking mode and listening mode. They presented an eye movement model which is based on both empirical studies of saccade and statistical models of eye tracking data [19] and we employed this model in our research. However, these works mainly concentrated on communicative or task-related eye movement, not on how eye movement reveals emotion.

To date, only few works have explicitly considered virtual agents capable of revealing theirs emotional states through the manner of their eye movement. Among these, Faigin illustrated that downcast eyes, upraised eyes, eyes looking sideways, and even out-of-focus eyes are all suggestive of states of mind [11]. Gu and Badler stared with the main goal to make virtual agents convey emotional information by eye movement [13]. Another interesting work is done by Lance and Marsella, who realized a model of emotionally expressive head and body movement during eye movement shifts based on Gaze Warping Transformation (GWT), which is a combination of temporal scaling and spatial transformation parameters that describe the manner of an emotionally expressive eye movement shift [16]. After that they proposed a model of realistic, emotionally expressive eye movement that builds upon the GWT by improving the transformation implementation, and by adding a model of eye movement inspired from the visual neuroscience literature [17]. Then they provided the results of a reverse engineering study which presents a preliminary mapping between gaze behaviors and emotional states that could be used with a variety of gaze or emotion models. However, most of the aforementioned research only focused on analysis and synthesis of primary emotions, whereas intermediate