Emotional Eye Movement Markup Language for Virtual Agents

(Extended Abstract)

Zheng Li
Beihang University
P.O.Box 206, 100191
buaa_david@yahoo.com.cn

Xia Mao
Beihang University
P.O.Box 206, 100191
moukyoucn@yahoo.com.cn

ABSTRACT

EEMML (Emotional Eye Movement Markup Language) is a scripting tool that enables authors to describe and generate emotional eye movement in virtual agents. The EEMML is capable of describing and generating both basic eye movement and emotional eye movement, including primary (joy, sadness, anger, fear, disgust and surprise) and intermediate (emotions that can be represented as the mixture of two primary emotions) emotions for virtual agents. The emotional eye movement generation framework is based upon the MPEG-4 FAP (facial animation parameters), and the animations are driven by parameters picked from the Cohn-Kanade AU-Coded facial expression database as well as real-time eye movement data(pupil size, blink rate and saccade).

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*input devices and strategies, interaction styles, theory and methods*

General Terms

Human factors

Keywords

Virtual Agent, Markup Language

1. INTRODUCTION AND MOTIVATION

The manner of a person's eye movement conveys much nonverbal information and emotional intent beyond speech. While the role and mechanics of emotional eye movement in virtual agents have been extensively studied [5][2], how the emotional eye movement in virtual agents can be described is not well explored. The goal of our paper is to gain a deeper understanding of how to script emotional eye movement in virtual agents easily. The Emotional Eye Movement Markup Language (EEMML) is such a scripting tool.

2. SYSTEM ARCHITECTURE

In this paper we propose a scripting language that enables authors to describe and generate emotional eye movement in virtual agents. The overall system of our research is illustrated in Fig. 1. The EEMML is based on emotional eye movement generation framework which picks the parameters from the AU-Coded facial expression database [1] and real-time eye movement data (pupil size, blink rate and saccade). We compute statistics about the involved FAP [4] which contains the values of eyebrow, eyelid and eyeball that associate with the eye movement. Finally, a rule based method is employed to generate emotional eye movement for primary and intermediate emotions in virtual agents. The realization of the framework is extensively discussed in paper[2]. The tags are defined to control the agents' emotional eye movement. Document Type Definition (DTD) for the tags is created to define the grammar. The validation module invokes the Document Type Definition (DTD) file to check the errors contained in the EEMML script file. The parser module employs the Simple APIs for XML (SAX) to parse the EEMML script. The converter module transforms the parsed EEMML script to FAP values according to the emotional eye movement generation framework.
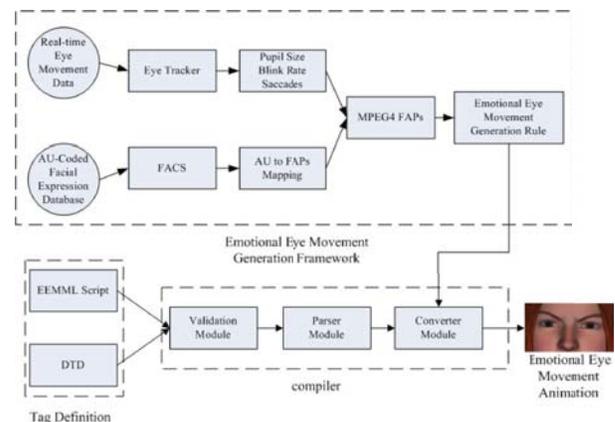

Figure 1: System of emotional eye movement markup language

EEMML is an XML-based language that composes a tag tree described in Fig. 2. All the tags are easy to learn and remember, as they follow the conventions of HTML.

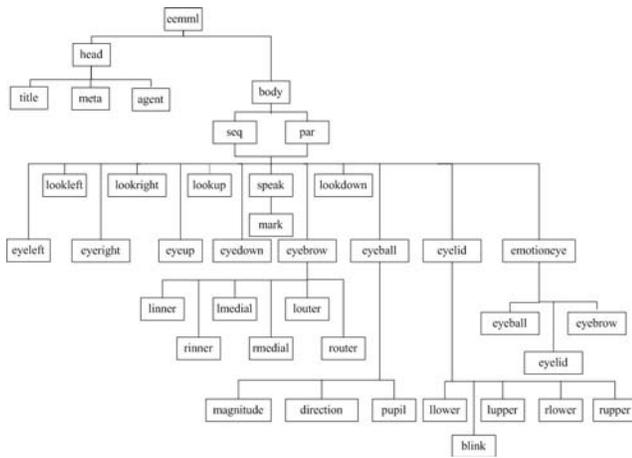The root element of an EEMML script is eemml. The root

Figure 2: Tag structures of EEMML

element's tagging structure <eemml> ... </eemml> contains all other tagging structures of the script. The head element specifies general information, the title, the meta (author's information) and the agent that will perform as a presenter. The tag body refers to the sequence of events comprising the actions of virtual agents.

The seq and par elements, which belong to the presentation control tags, are responsible for events that should be executed sequentially or in parallel, respectively. The eyeleft, eyeright, eyeup and eyedown elements are employed to control the agents' eyes rotation, while the lookleft, lookright, lookup and lookdown elements are employed to control the agents' both eyes and head rotation. The eyebrow element is defined to control the the agents' eyebrow, including inner part, medial part and outer parts for both eyes. The eyelid is employed to control the agents' eyelid, including lower part and upper parts for both eyes. The blink element is also a sub-element of the eyelid used to control the eye blink rate. The eyeball element is specified to control the agents' eyeball, including pupil size, magnitude and direction for both eyes. With the speak element, the agent speaks one or several sentences through a TTS (text-to-speech) systems. The mark element can be used to set an arbitrary mark at a given place in the text, so that an engine can report back to the calling application that it has reached the given location.

The emotioneye element is defined to control the emotional eye movement, twenty-two kinds of emotions (gratification, joy, pride, disappoint, anger, disgust, resentment, remorse, surprise, relief, satisfaction, sadness, pity, admiration, gratitude, hate, fear, shame, hope, reproach, love and gloating) can be generated. All the parameters related to the twenty-two kinds emotions are obtained through the emotional eye movement generation framework. If the scripting author wants to modify some parameters, they can use the basic eye movement control tags below the <emotioneye> tag, then the related parameters value will be replaced. As the example shows below, the blink rate value in joy emotion will be replaced by 3 per-10s.

<emotioneye type ="joy">
<blink rate = 3/ >
< /emotioneye>
A simple example of EEMML is shown as follow. The

agent will express "sad" emotion when she speaks "I'm sorry to hear that you have met an accident" then express "joy" emotion when she speaks "but you looks much better now". The "speak-begin" and "speak-end" in tag emotioneye represent when the emotional eye movement will begin and end respectively.

<eemml>
<head>
<title>EEMML sample< /title>
<meta authorname ="David"/ >
<agent id = "1" name= "Alice"/ >
< /head>
<body>
<par>
<speak id="1">
<mark id="1"/ >I'm sorry to hear that you had an accident <mark id="2"/ > but you look much better now<mark id="3"/ >
< /speak>
<seq>
<emotioneye type ="sad" speak-id = "1" speak-end="2"/ >
<emotioneye type ="joy" speak-id = "1" speak-begin="2"/ >
< /seq>
< /par>
< /body>
< /eemml>

3. DISCUSS

The EEMML focus on the emotional eye movement in virtual agents and can realize fine-grained emotional eye movement generation. Almost all the eye related parameters can be controlled by EEMML. The EEMML can also serve as a sub-module for some of the existent scripting languages for virtual agents, such as MIML[3] etc, to enhance the function of these languages.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1] J. Cohn, A. Zlochower, J. Lien, and T. Kanade. Automated face analysis by feature point tracking has high concurrent validity with manual facs coding. *Psychophysiology*, 36:35–43, 1999.

[2] Z. Li, X. Mao, and L. Liu. Providing expressive eye movement to virtual agents. In *Proceeding of ICMI-MLMI2009*, pages 241–244, 2009.

[3] X. Mao, Z. Li, and H. Bao. Extension of mpml with emotion recognition functions. In *Proceeding of IVA2008*, pages 289–295, 2008.

[4] MPEG4. Moving pictures expert group, iso/iec 14496, international standard on coding of audio-visual objects. Technical report.

[5] M. Thiebaux, B. Lance, and S. Marsella. Real-time expressive gaze animation for virtual humans. In *Proceeding of AAMAS2009*, pages 321–328, 2009.