Generation of Comic Facial Expressions on Variant 3D Mesh Models for Intelligent Avatar Communication Systems

Sang-Woon KIM, Yong-Who LEE and Yoshino AOKI*

Dept. of Computer Science & Engineering, Myongji University, Korea
*Graduate School of Engineering, Hokkaido University, Japan

1 Introduction

As a means of overcoming the linguistic barrier in the Internet cyberspace, recently a couple of studies on intelligent avatar communications between avatars of different languages such as Japanese-Korean have been performed [1], [2]. From these studies, it was revealed that facial expression qualities of more realistic avatar models are not as good when compared with simple models, even though the polygon number of the former is much larger than that of the latter. As an avatar model to be employed in cyberspace, therefore, a comic-style simple model may be better than the realistic complex models. However, the avatars who are participating in the communication, are different from each other in their mesh models and polygon numbers. So, in this paper we propose a method of generating comic facial expressions with variant 3D mesh models. To simulate that, we employ some kinds of polygonal mesh models, from which Action Units parameters are extracted and prototypic emotional expressions are regenerated.

2 Intelligent Avatar Communication

In the intelligent avatar communication systems, a set of animation parameters such as Joint Angles (JAs) of the arm gestures and the hand silhouette and Action Units (AUs) for facial expressions are transmitted instead of sending the entire-real motion pictures. As an example, consider an avatar communication between Korean and Japanese, where two avatars, named ‘A’ and ‘B’, communicate with each other in their own sign-languages. Korean Sign-Language (KSL) messages of ‘A’ are re-constructed (animated) with the corresponding Japanese Sign-Language (JSL) on ‘B’ after being transmitted as parameters of the joints angles and the action units [1]. In this case, the Korean avatar ‘A’ sends a KSL message, “NaNeun HakKyoEa KamNiDa (I go to school)”, to ‘B’, then the Japanese avatar ‘B’ receives the corresponding message, which has been translated into the JSL like “WatashiHa GakKouNi Ikimasu (I go to school)”.

Table 1: A combination set of AUs and PAUs to generate comic facial expressions, where the 1+2+12, etc. represent the combinations of \{ AU1, AU2, AU12 \}, etc.

<table>
<thead>
<tr>
<th>Expr.</th>
<th>Combination of AUs</th>
<th>PAUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joy</td>
<td>1+2+12, 1+2+6, 1+2+6+12</td>
<td>12</td>
</tr>
<tr>
<td>Sadness</td>
<td>1+4+7+15</td>
<td>15</td>
</tr>
<tr>
<td>Disgust</td>
<td>2+4+10, 2+4+25, 1+4+20+25</td>
<td>20</td>
</tr>
<tr>
<td>Anger</td>
<td>2+4+5, 2+4+15, 2+4+5+15</td>
<td>5+15</td>
</tr>
<tr>
<td>Surprise</td>
<td>1+2+5, 1+2+26, 1+2+5+26</td>
<td>26</td>
</tr>
<tr>
<td>Fear</td>
<td>1+4+20, 1+4+25, 1+4+20+25</td>
<td>48</td>
</tr>
</tbody>
</table>

3 Comic Facial Expressions

A widely used scheme for describing the facial expression is the Facial Action Coding System (FACS) [3], which describes the set of all possible basic facial muscle action units (AUs) performable by the human face. With the AUs, we can generate numerous expressions. For example, a “joy” expression can be edited with the combinations of AU’s such as \{ AU1, AU2, AU12 \}, \{ AU1, AU2, AU25 \}, \{ AU1, AU2, AU12, AU25 \}, etc. Among them, a few of the action units such as AU 1, AU 2 and AU 4, are used in common. On the other hand, each universal expression has one or two special AU’s, which have the greatest influence on generating the expression. We define here the special action units as the Principal Action Units (PAUs) [2]. Also we define the Expression Intensity to exaggerate each expression comically. The intensity of a PAU, which is defined as a value within [0.0 ~ 1.0], corresponds with the amount of its activation (that is, the impressiveness). Higher PAU intensity can produce more joyful (or angry) expression, which will be discriminated easily. Table 1 shows a combination set of AUs and PAUs to generate comic facial expressions.

4 Movements of the Cheeks

Movements of the cheeks as well as those of the eye-brow, the eye and the mouth, play significant roles in generating the facial expression. So, we include the movements in generating the facial expression. To implement the movements, we utilize the parabolic partial differential equation defined as:

$$u_i(t+1) = u_i(t) + \alpha \frac{\Delta t}{\Delta x^2} (u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)),$$  \(1\)

where $\Delta x$ and $\Delta t$ are the mesh sizes of the $x$-direction and the $t$-direction, respectively, the $i$ and $j$ are the mesh point indices in the $x$-plane and the $\alpha$ is a parameter which represents the height of the parabolic curve. The $\alpha$, $\Delta x$ and $\Delta t$ are determined experimentally as values of $0.1 \leq \alpha \leq 10.0$, $\Delta x = D/(N-1)$ and $\Delta t < \Delta x^2/(2\alpha)$, respectively, where $D$ is the diameter of the cheek’s area and $N$ is the mesh number of the $x$-direction.
5 Generation of Expressions

The process of generating an expression on a given mesh model, \( M \), can be formalized as follows:

1. Classify the polygons of \( M \) by using their color and vertex information. From the classification, facial components of \( M \) and their corresponding AUs are extracted.

2. Search the Neighborhood area of each AU (NAU):
   (a) Extract NAUs of the left and right eyebrows, the left and right eyelids, and the upper and lower lips, respectively.
   (b) Separate NAUs of the eyebrows from those of eyelids. Also, separate NAUs of the upper lip from those of the lower lip.

3. Extract NAUs of the cheeks using the coordinate information of the facial components of \( M \).

4. Generate the facial expression on \( M \) using the corresponding AUs and their Expression Intensities.

In 2., to extract the neighborhood area of AUs, NAUs, we utilize two rectangles: a small one and a large one. The former is for near neighborhood area, the latter is for a little far neighborhood area. These two regions are different from their Intensities.

6 Experiments

The proposed method was experimented with 3D different mesh models using the Visual C++ 5.0 and the Open Inventor on the Windows’ platforms. Fig. 1 shows the four different mesh models: (a) ‘Nancy1’, which is composed of the 472 polygons and the 296 vertex, (b) ‘Nancy2’, which is constructed with the 1,888 polygons and the 1,065 vertex, (c) ‘Girl1’, which is built with the 4,288 polygons and the 2,334 vertex, and (d) ‘Girl2’, which is produced with the 17,152 polygons and the 8,940 vertex.

The six kinds of prototypic comic expression such as the joy, the sadness, the disgust, the anger, the surprise, and the fear, were experimentally reproduced with different Expression Intensities. Figs. 2 and 3 show the “joy” expressions generated without the cheeks movements and with the cheeks movements, respectively.

![Fig. 1: Experimental 3D mesh models. (a) Nancy1 (472 : 296), (b) Nancy2 (1,888 : 1,065), (c) Girl1 (4,288 : 2,334), and (d) Girl2 (17,152 : 8,940). Here (· · ·) represents the number of polygons and vertex.](image)

![Fig. 2: The “joy” expression generated without the cheeks movements on the four experimental 3D mesh models of Fig. 1.](image)

![Fig. 3: The “joy” expression generated with the cheeks movements on the four experimental 3D mesh models of Fig. 1.](image)

Table 2: Generation time of the “joy” expression on the four experimental 3D mesh models of Fig. 1 (sec.).

<table>
<thead>
<tr>
<th>Model</th>
<th>Time (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nancy1</td>
<td>0.02</td>
</tr>
<tr>
<td>Nancy2</td>
<td>0.24</td>
</tr>
<tr>
<td>Girl1</td>
<td>1.29</td>
</tr>
<tr>
<td>Girl2</td>
<td>67.10</td>
</tr>
</tbody>
</table>

7 Conclusion

From the experiments, we confirmed a possibility that the proposed facial expression method could be used for inter-communication between avatars of different languages on the Internet cyberspace. However the number of polygons is greater than 10,000, we could not generate expressions in real-time. Texture mapping of the 3D mesh models is the future work. This work was supported by the Ministry of Monbu-Kagaku under the Grant 13305026.

References