Rigid Bones Grouping Scheme for Facial Expressions Synthesis Utilizing Three-Dimensional Head Model

Maja Kocoń, West Pomeranian University of Technology

Abstract
In the recent years the demand for intuitive communication between user and machine is growing. Therefore, the synthesis and analysis of emotions play significant role in computer systems [1]. Facial expressions can improve man-machine interactions, because the human face is natural medium of communication. We have created three-dimensional model according to face topology and anatomy. In our approach, the control signals are exploited to change model properties based on emotional state using FACS [2] system. The structure of proposed model has been prepared taking movement aspects into account. The problem of grouping bones in emotion mapping process is presented in this work. Defined groups of bones can be applied to simplify the number of control parameters to changing emotional state in the human face model.

1. Introduction
In the face analysis or synthesis fields, the biggest challenge is the exploration how people react to the world and interact with other people [3]. Thus, not only face but also facial movements play a major role in emotions and interpreting situation.

Today, the facial expression analysis and synthesis techniques are becoming popular, because the demand for the intuitive communication between people and machines (i.e. user-friendly interfaces, interface agents, human-robot interaction etc.) is growing. Studies in the field of facial expression may be helpful in such different areas as virtual agents, video teleconferencing [4] and man-machine interactions. Facial expression can provide a means of communication - the basic information about the user's needs and requirements.

The synthesis process of 3-D shape model of the human head and face with defined emotion or pose exploits several modeling techniques [8]. In order to expression-based shape and movements deformations, the Facial Action Coding System (FACS) is frequently used [2]. Accurate measurements of 3-D face structure with various emotions are still a difficult tasks [4].

The following sections describe the human head model based on anatomical properties along with definition of rigid bones required to express emotions. The modification of bones' controls play main role in animation quality [8]. Therefore a simple approach to simplify control process is presented.

2. Three-Dimensional Head Model
The three-dimensional model of the human head has to be designed taking many aspects into account. For example, high quality facial expression mapping needs a lot computations for control points trajectories estimation to achieve realistic movements. Therefore it is important to respect the human face anatomy including muscles mechanics [6-8]. Knowledge of the structure of the face, the movements of muscle and their relationship is important when the model is supposed to be used to improving man-computer interactions (i.e. three-dimensional agent) [6].

Model presented in this paper has been designed manually using extrusion modeling technique in order to easiness in placement of vertices at the beginning of modeling process [5]. The mesh of 3-D model is shown in figure 1 with different views.

The most important part of the model from emotions point of view is the face. Face carry most information describing human emotional state [3]. Existing Facial Action Coding System [2] explains how to classify facial behaviors based on the face muscles and how activity of every muscle is related to face appearance. FACS describes all facial activity with 44 action units (UA's) in several categories of head and face areas positions. Each action unit has a code, for example: AU 1 - Inner Brow Raise, AU 17 - Chin Raiser or AU 26 - Jaw Drop, etc. [2]. Based on FACS system rules we defined the face topology presented in figure 2. Exploiting proposed model in our experiments, we have concentrated on the facial movements, ignoring head and eye movements properties.
3. Proposed Approach

Facial expressions have been implemented with the help of bones applied to vertices of the mesh in 3-D model [5, 8]. Expressions were obtained by bones translation (bones are assigned to face’s areas like of the real muscles [7]).

Mapping the expressions on the 3-D model is associated with the problem of the large number of points, which require many of calculations and memory resources. For this reason, we tried to apply techniques of grouping elements responsible for facial expression changes. In the grouping process, elements are selected in order to direction similarity of movement vectors. It is possible to reduce the effort needed to animation control. As grouping criterion we adopted the Euclidean distance calculated between bones positions in different emotional expressions (in our study these distances were calculated between neutral and joy states). Bones with highest distance are marked as master in the group and the rest group members are bones with the same movement vectors as master. The idea is to store information about the source model and master bones translation vector while changing emotional states of the model. All positions of members in each group are dependent on translation of master and are modified accordingly to their relative angles (defined in polar system) as depicted in figure 3. Regions marked in gray define boundaries for $\varphi$ and $\theta$ angles.

The size of the group and the number of groups are determine efficiency and applicability of proposed approach. Let’s define the group of bones having the same direction of motion vectors and being in accordance with anatomical aspects:

\[
[M^{(0)}, M^{(1)}, \ldots, M^{(N-1)}],
\]

where:

- $N$ - number of bones,
- $M^{(i)} = [x_n(i), y_n(i), z_n(i)]$,
- $M^{(0)}$ - the bone determining the location of the group $M^{(0)} = [x_g, y_g, z_g]$.
In order to modify the model for each bone in the group, except master (\(M^{(0)}\)), the following transformation needs to be calculated:

\[
\begin{align*}
    x_n(i) &= x_g + t_x \cdot f(\varphi, \theta) \cdot e_x, \\
    y_n(i) &= y_g + t_y \cdot f(\varphi, \theta) \cdot e_y, \\
    z_n(i) &= z_g + t_z \cdot f(\varphi, \theta) \cdot e_z
\end{align*}
\] (2)

where:

- \(T = [t_x, t_y, t_z]\) is a translation vector of \(M^{(0)}\) bone,
- \(f(\varphi, \theta)\) - mapping function which define scaling value depending on \(\varphi\) and \(\theta\) angles for each bone in the group,
- \(\varphi, \theta\) - angles of the bones in the polar system,
- \(e_x, e_y, e_z\) - scaling constant selected for a specific 3-D model, \(1 \leq i < N - 1\).

We used the angles determining the location of the bone due to facilitate the mapping of geometrical characteristics of the head. As the result, spherical coordinates and radially placed bones can be used for controlling members in groups (similarly as for the growth simulation [8]).

Thanks to this approach we can store only information about the location of bone in a neutral state, while the change to a specific state will be require only information about the \(T\) vector for each master bone in the group. In the result, it is sufficient to store changes to one bone in all groups, instead of storing the changes to all bones in model. Such approach requires precise mapping of anatomy aspects of the final 3-D model. The grouping scheme is presented in figure 4. It’s important to specify the number of groups carefully in order to making bones’ movements dependent on \(\varphi\) and \(\theta\) angles.

1) Select number of groups using anatomical properties
2) Group bones accordingly to their movement vectors
3) Calculate Euclidean distances \(d_j(s, s')\) between bones in neutral \((s)\) and selected emotional state \((s')\)
4) In each group select master bone exploiting maximum distance: \(M^{(0)} = \max[d_j(s, s')]\)

Fig.4. Grouping scheme

Fig.5. Bones configuration (front, side and top view) .
Table 1 shows an example of bones categorization into groups. The main criterion was the compatibility of motion vectors (master bones in each group is marked in bold).

<table>
<thead>
<tr>
<th>Group</th>
<th>Bones</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Nose2.R, Cheek2.R</td>
</tr>
<tr>
<td>5</td>
<td>Cheek3.R, Cheek.R</td>
</tr>
</tbody>
</table>

4. Conclusion

Presented approach can be used for simplifying the process of modification 3-D human head model from facial expression mapping point of view. The selection of number of groups and bones remain an open question. Fewer control parameters simplifies the process of mapping emotions to the model, but simultaneously it influences on the final representation accuracy. Modification of defined groups needs to assign functions to selected regions of face model. To obtain these functions, authors currently working on analysis of relationships between emotional states accordingly to specified groups.

Bibliography


Author:
MSc. Maja Kocoń
West Pomeranian University of Technology
ul. Sikorskiego 37
70-313 Szczecin
tel. (091) 449 41 50
e-mail: maja.kocon@zut.edu.pl