Adaptive Head Modeling System and Its Application

Yoshimitsu AOKI  Shuji HASHIMOTO

Dept. of Applied Physics, School of Science & Engineering
Waseda University
Okubo 3-4-1, Shinjuku-ku, Tokyo 169-8555, Japan
{aoki, shuji} @shalab.phys.waseda.ac.jp

Abstract

We propose an adaptive head modeling method which can be selective according to the purposes and available input resources. In each case, we utilize the generic head model for the head reconstruction, which is constructed based on the facial anatomy and the hierarchical structure. This model is capable of simulating a facial dynamics by the physics-based computation.

For the head reconstruction, we account for the input resources as facial images, X-ray images, 3D-CT, and 3D Digitizer. The proposed methods are the photogrammetric technique using two directional facial images and X-ray images, the combination of the 3D digitizer and X-ray images. Using the extracted shape information from those inputs, the generic head model is flexibly deformed to match the target head so that it is useful for wide aims. A facial dynamics simulation and a facial surgery simulation are introduced as the concrete applications.

1. Introduction

1.1 Related works

The facial modeling is very essential issue for the natural facial representation and the technique is utilized in several fields and applications, such as teleconferencing, man-machine interface, and realistic human representation in the virtual reality space. 3D reconstruction and visualization of the body shape information including a human head are also attached importance for the assistance of examinations in the medical fields.

Regarding a face modeling, a variety of methods have been proposed corresponding to each purpose. It is assumed that required accuracy of the model depends on the aims, and available input resources are different for each system. In addition to these points, the performance of computers should be considered as well. The typical processes of the modeling are composed by two stages; the feature extraction stage and the fitting stage. In the first stage, facial shape information is extracted from input resources. Then the prepared geometric face model can be modified to the specific person’s one using the extracted features in the second stage. Hereafter, we review some related works dealing with a facial modeling.

- Image-based Technique

The simple and effective method to reconstruct a facial model is the image-based technique using the shape features extracted from facial images. The most frequent method proposed is to use two-directional facial images captured from frontal and side views. [1] and [2] introduced the 3D face reconstruction for the visual communication. These kinds of techniques can provide a reconstructed facial model easily, although those 3D shape is not completely accurate. Therefore, reconstructed models are generally suitable for the visual communication system which requires realtime and interactive performances.

- 3D shape from Scanner, CT, MRI

For the accurate modeling, a non-contact 3D digitizer, CT, and MRI are useful. The 3D digitizer can obtain facial surface position data and corresponding facial texture at the same time. [3] proposed the hierarchical physics-based facial model which is constructed from a high resolution 3D digitizer for generating realistic facial animation. [4] developed a prototype system for the surgical planning and the prediction of facial...
shape after the craniofacial and maxillofacial surgery for patients with facial deformities, using finite element models from MRI data set [5]. H.Mita et al. proposed a elaborate model including underlying facial tissue from CT scanned data for a facial paralysis simulator [6]. The obtained 3D data in these cases is highly accurate, although amount of data and modeling costs increase. Especially respecting the medical applications, these kinds of modeling methods should be required for the reliable visualization and simulation.

As reviewed above, a variety of the facial modeling methods have been proposed in order to achieve each aim, with the consideration of the required accuracy and the amount of data.

1.2. Motivation

In this paper, we propose an adaptive head modeling method which can be selective according to the purposes and available input resources. In Table 1, our modeling techniques are listed. The first column indicates our available input resources; facial images, X-ray images, a 3D digitizer, and a 3D-CT. The second and third columns are the obtainable shape information from those inputs. From facial images, we can extract the facial feature points and contours. From X-ray images, we can obtain not only a facial contour but also some feature points of a skull (generally called anatomical measurement points). Using a 3D digitizer and 3D-CT, the full surface data of a face and a skull can be available for the head modeling. In proportion to downward of this table, the costs and accuracy of modeling increase.

In each case, we utilize the geometric head model for the head reconstruction as the initial head, which is constructed based on the facial anatomy and the hierarchical structure. This model is capable of simulating a facial dynamics by the physics-based computation (Detailed description of this model is in [7][8][9]).

For the head reconstruction, we account for the input resources as facial images, X-ray images, a 3D-CT, and a 3D Digitizer. The proposed methods are the photogrammetric technique using the two directional facial images and X-ray images, the combination of the 3D digitizer and X-ray images. Using the extracted shape information from those inputs, the generic head model is flexibly deformed to match the target head so that it is widely useful for a variety of aims. As the concrete applications, a facial dynamics simulation and a facial surgery simulation are introduced to show the performance of this method.

We organize this paper as follows; In the section 2, we describe the proposed head model and its dynamical mechanism. In the section 3 and 4, the photogrammetric technique using two directional facial images and X-ray image is introduced. In the section 5, we describe the head modeling method with highest accuracy which utilize both of a 3D digitizer (shape and texture information) and X-ray images. In the section 6, we introduce some applications using our head model constructed from the proposed methods. In the section 7, we finally conclude this paper with some future works.

2. Initial head model

2.1 Modeling method

We constructed the hierarchical head model that consists of a skin layer, a muscle layer, and a skull layer. The skin and skull layers are segmented from 3D CT data by thresholding, using the segmentation software called “Mimics” as shown in the Fig.1.

The skin layer is the wire-frame model constructed from CT scanned data, which is regarded as an elastic body. All frames composed of the skin are simulated by non-linear springs that can represent the skin elasticity by their elastic parameters.

<table>
<thead>
<tr>
<th>Input Resources</th>
<th>Obtainable Information</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial images (frontal &amp; side)</td>
<td>feature shape</td>
<td>Low</td>
</tr>
<tr>
<td>X-ray image</td>
<td>feature points contours</td>
<td></td>
</tr>
<tr>
<td>3D Digitizer</td>
<td>surface</td>
<td>High</td>
</tr>
<tr>
<td>3D CT</td>
<td>surface</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. A variety of modeling methods
The facial muscles are also modeled by non-linear springs which start from the skull layer, and are attached to facial surface just like the actual facial structure. The muscles are grouped by their position, and act with connected facial tissues in harmony. In the present, 14 kinds of facial muscles are simulated which are mainly concerned with facial expressions. Each muscle has the contraction parameter to generate facial expressions.

For the skull layer, the polygon model from ViewPoint Corp. is used. The jaw part of this model can simulate realistic jaw movement with six degrees of freedom and it is located under facial tissues referring the relationship between the face and skull.

Three layers composed of the hierarchical head model and integrated head model are illustrated in Fig.2, and detailed information about the model is summarized in Table 2.

This generic head model is used as the initial head model for the head reconstruction process.

Table 2. Detailed information of the model

<table>
<thead>
<tr>
<th>Layer</th>
<th>Data</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>751 points, 1288 polygons</td>
<td>elastic</td>
</tr>
<tr>
<td>Muscles</td>
<td>14 facial muscles</td>
<td>elastic</td>
</tr>
<tr>
<td>Skull</td>
<td>5072 points, 7757 polygons</td>
<td>rigid</td>
</tr>
</tbody>
</table>

2.2 Facial Dynamics

In generating facial expressions, points on the skin model are moved in order to obtain the modified face. The energy of the spring system can be changed by both of the muscular contraction and the jaw action, and the new position of each point on the facial surface is obtained by calculating the energy equilibrium point of the entire spring system. The kinetic equation for feature point $i$ on the facial surface is given as following eq.(1).

$$m \frac{d^2r_i}{dt^2} = - \sum_{j} k_{ij} (r_i - r_j) + m g - R \frac{dr_i}{dt}$$

Where $r_i$ represents the 3D coordinates of the feature point $i$. The first and second terms represent the total elastic force affected to the point $i$ and the attraction of gravity, respectively. The value of $k_{ij}$ is changed gradually by the contraction parameter $C_i$ of each spring so that it can represent the skin and muscles’ non-linear property. The third term is the viscous term of the facial model.

Using this model, we developed the facial action synthesizer which enable the user to create facial expressions derived from facial actions (See Fig.3). Required facial parameters in facial modification are the contraction rate of muscles and jaw action parameters, which can be controlled by the GUI (Fig.3.(b)). Facial actions are parameterized relative to typical facial expressions based on anatomical analysis.

![Figure 1. Segmentation software “Mimics”](image)

![Figure 2. The hierarchical head model](image)

![Figure 3. Facial action synthesizer](image)
Fig. 4 shows the result of generating the facial expression “surprise” which is caused by the muscular contraction around a forehead and jaw rotation. Thus, our head model can create flexible facial expressions derived from the giving parameters, such as contraction rates of muscles and action parameters for the jaw. Hereafter, we use the described model as the initial model for arbitrary persons. From the following section, three reconstruction methods are introduced respectively.

3. Reconstruction from Two Facial Images

As the simplest method for the facial reconstruction, two directional facial images are used to extract facial features. For this method, only two directional images from frontal and side views are needed and facial parts (eyes, mouth, nose, etc.) are dealt with in detail features of the facial appearance.

3.1 Feature Extraction

3.1.1 Frontal Facial Image

From a frontal facial image, the contours of eyes, a nose, a mouth and the facial outline are automatically detected to locate facial parts of the model. The automatic recognition is realized by the combination of popular techniques; the skin color extraction, the template matching, and the active contour model.

At first, the facial area is detected from the result of skin color extraction based on hue information. From the horizontal and vertical projections of blightness, the positions of pupils are detected roughly and their detailed positions are decided by the template matching. Then the approximate areas which enclose facial parts can be estimated from the distance between pupils, and the initial positions of active contour model are determined. Finally, active contour models search facial contours based on their energy functions. This software is basically developed by IPA, and we refine this to be useful for our method. More detailed information about this method is described in [9].

Throughout these processes, the contours of eyes, a nose, a mouth, and the facial outline are detected and some feature points on those contours are extracted at regular intervals. Fig. 5 shows the example of detected feature points on a frontal facial image.

3.1.2 Profile Facial Image

The facial profile which can be detected from a facial image from a side view is also important to improve the quality of the facial reconstruction result. The facial silhouette in the image can be easily separated from other regions by thresholding, if the input image has a simple color. The input image and the detected facial silhouette are shown in Fig. 6.

We detect the points on the facial contour which has a large curvature, as the feature points. The value of curvature $C_s$ along a facial profile can be computed using following eq.(2).

$$C_s = \frac{dD_s}{ds} \quad D_s = P_{i+1} - P_i$$

Where $P_i$ represents the coordinates of the point on the facial profile, and $D_s$ represents the difference between two adjacent points, respectively. $s$ is the line element along a facial profile. Fig. 7(b) indicates computed values of curvature along the vertical axis, and from this figure, we can detect the feature points by checking zero-cross points (+ to -: concave, - to +: convex). From the general facial structure, the sequence of the convex points and the concave points are determined.
and used for the feature points detection. Fig.7 (a) shows the detected feature points on the facial profile using this technique.

3.2 Results

Using the coordinates of the detected feature points from two directional facial images as shown in Fig.6(a) and Fig.7 (a), the corresponding points of the facial model can be deformed and the other points are moved by the interpolation based on their locations relative to feature points. By the method as mentioned earlier, we reconstructed the 3D face model from full automatic processing. Fig.8 is the reconstructed face model from several views. This result shows the effectiveness of proposed simple method.

4. Reconstruction from X-ray Images

In this section, we describe more precise head modeling method using x-ray images. To obtain three-dimensional shape information of a face and a skull, two-directional X-ray images and a frontal facial image are used in this method.

4.1 Measurement Points Extraction

4.1.1 Measurement Points

Anatomical measurement points on a face and a skull are usually utilized in order to analyze their shape quantitatively in orthodontics. Based on this analysis, orthodontists plan how to treat in each case of a patient by reconstructing 3D shape from images. Fig.9 shows the plotted measurement points on the sketches of X-ray images (cephalograms) from both of frontal and side views. Each position of measurement points is defined in detail, especially the point on the contour which has large curvature. We manually plot on the two-directional cephalograms by mouse operations, and use them as the typical control points of the head. The number of measurement points is 22 for a skull model, and 21 for a face, respectively.

4.1.2 3D Cephalogram Measurement

From the 2D coordinates of each point on the two-directional cephalograms, the 3D coordinates can be calculated by integrating them. However, a target object is usually projected onto the image planes as expanded and rotated images (See Fig.10, 11). So we have to adjust them for higher head modeling.
In capturing X-ray images, a head is fixed by the ear rods. Two x-ray sources and cameras are also bound, so we can obtain both of normalized facial images and x-ray images from frontal and side views. The head is projected onto the picture planes as expanded images of actual measurement. Therefore, the expansion adjustment should be considered to obtain an accurate 3D shape of the head. As shown in Fig.10, the point on the subject \( A_n (x_o, y_o, z_o) \) is enlarged and projected as the point \( A_{LAT} (y_{LAT}, z_{LAT}) \) on the picture plane (YZ plane). The coordinates of the measurement point \( A_n \) can be computed by following adjustment eq.(3). The value of \( k \) is the distance from Y axis, so that it depends on each measurement point.

\[
y_o = y_{LAT} \times \frac{1500 + k}{1650} \quad z_o = z_{LAT} \times \frac{1500 + k}{1650}
\]

Although the head is fixed by the ear rods, slight head rotation around horizontal axis (Fig.11) is unavoidably in capturing images so that the measurement point on the subject is projected to the incorrect position \((x_{PA}, y_{PA})\) on the XY plane. The rotation adjustment can be calculated by eq.(4). Then, the line \((y=y_{PA})\) is drawn on the frontal cephalogram, and the exact value of \( x_{PA} \) is obtained by the cross section with the skull and facial contour using eq.(5). Thus, the accurate 3D coordinates of the measurement point \((x_o, y_o, z_o)\) can be ready for the face reconstruction process.

\[
y_{PA} = \frac{1650(1500 + i)(y_{LAT} \sin \theta + y_{LAT} \cos \theta)}{1500 \times 1650 + (1500 + i)(z_{LAT} \cos \theta - y_{LAT} \sin \theta)}
\]

\[
x_o = x_{PA} \times \frac{1500 + z_o}{1650}
\]

4.2 Facial Feature Points Extraction

From two directional X-ray images, rough profiles of the head can be extracted. For more detailed fitting of the face shape, we use facial feature points on the contours of facial parts (eyes, nose, mouth). The detection algorithm is same as described in section 3. Fig.12 shows the results of facial feature points extraction from normalized frontal images.

4.3 Deformation of the Head Model

4.3.1 Corresponding Points

In order to deform the generic head model according to obtained 3D coordinates of measurement points, we should also set the corresponding points on the head model. We selected the corresponding points as shown in Fig.13.

![Figure 13. Measurement Points](image)

(a) Skull                                    (b) Face

4.3.2 Deformation Method

Using the width measured from frontal image \((x_o, x_j)\) and the corresponding width of the initial head model \((x_{1(初)}, x_{2(初)})\), rates of the width are calculated so that the x coordinate of the point \(i\) is computed by the linear interpolation eq.(6) (See Fig.14). \( S_i \) represents the number of steps between point \(I\) and \(2\). After width and height fitting, the coordinates of \((y_j, z_j)\) is used for the fitting of depth and height directions. The whole head is divided into 3 parts, back part and frontal parts (upper and lower) as shown in Fig.15, considering anatomical structure of a head. The movement vector of each measurement point is calculated, and rest of the points between two measurement points are computed by the linear interpolation eq.(7).

\[
x_i = x_{PA} + \frac{i}{S_j} (x_{PA} - x_{PA})
\]

\[
\Delta r_i = \frac{i \cdot p_i + (S_j - i) \cdot p_j}{S_j}
\]

\( \Delta r_i \): The movement vector of the point \(i\)

\( p_i, p_j \): The movement vectors of measurement points
5. Reconstruction from 3D Digitizer

The highest accurate modeling introduced in this paper is using a non-contact 3D Digitizer for a facial surface and X-ray images for a skull modeling. Using the 3D Digitizer (VIVID700, Minolta, Japan), the 3D points data and their corresponding texture image can be captured simultaneously. Therefore, 3D facial models with textures are easily obtained. The whole head model is reconstructed by the integration with the skull modeled from X-ray images. Fig. shows the modeling result. This result indicate extremely high quality compared to the results shown in former sections.

6. Applications

6.1 Muscular Action Analysis from Texture

Making use of the ability of our model that can analyze the relationship between the muscular actions and the facial expressions, we attempt to estimate the muscular actions relative to facial expressions. For this analysis, the method using two directional facial images is used to fit the generic model to the target faces. In this experiment, we manually adjusted the contraction parameters of muscles and elasticity of the skin, by comparing the synthesized facial animations with the actual sequences of facial expressions. In Fig.18, two different “smiling” faces and their internal state are displayed in raws. In this case, the muscular motion of the person A is more dynamic so that generated facial expression is more active as well, compared to the person B. Thus, this model realizes the mutual investigations, “muscular motion to facial expression (synthesis)” and “facial expression to muscular motion (analysis)”.

6.2 Facial Surgery Simulation

For the medical application of our head model, we developed a facial surgery system for orthognathic surgery (in cooperation from Kyushu Univ.). The predic-
section of facial morphology after surgery is a critical issue for both of patients and doctors. The developed system can predict the surgical result by calculating facial modification caused by the jaw movement, and display it using 3D computer graphics.

For this simulation, we reconstructed both of the face and the skull from X-ray images of an actual patient with mandibular prognathism who was required to set back the jaw, by the method described in section 4. We input surgical parameters of the jaw movement by referring the result of surgical planning which dentists decided. The surgical result and the simulated result are shown in Fig.19, Fig.20, respectively.

![Before surgery](image1.png) ![After surgery](image2.png)

**Figure 19. Result of orthognathic surgery**

![Before surgery](image3.png) ![After surgery](image4.png)

**Figure 20. Simulation result**

### 7. Conclusion and Future Work

We proposed an adaptive head modeling method which can be selective according to the purposes and available input resources. In each case, we utilize the generic head model for the head reconstruction, which is constructed based on the facial anatomy and the hierarchical structure. Proposed methods are the photogrammetric technique using two directional facial images and X-ray images, the combination of the 3D digitizer and X-ray images. Using the extracted shape information from those inputs, the generic head model is flexibly deformed to match the target head so that it is widely useful. A facial dynamics simulation and a facial surgery simulation are introduced as the concrete applications. For future works, we are considering to develop the more precise head modeling system by introducing the 3D skull database and detailed muscular and skeletal dynamics.

### References


S.Morishima, K.Aizawa, H.Harashima (1989) An intelligent fa-