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 geometoric 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 avaliable 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.

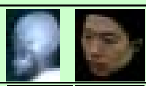
2. Initial head model

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 1. A variety of modeling methods

Input Resources		Obtainable Information: Facial shape	Obtainable Information: Skull shape	Accuracy
Facial images (frontal & side)		feature points contours	-	Low
X-ray images Facial image		feature points contours	feature points contours	↓
3D Digitizer X-ray images		surface	feature points contours	↓
3D CT		surface	surface	High

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 surmmarized in Table 2.

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

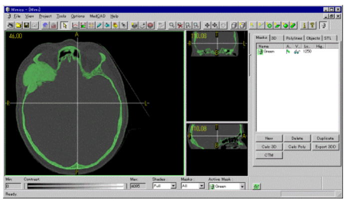
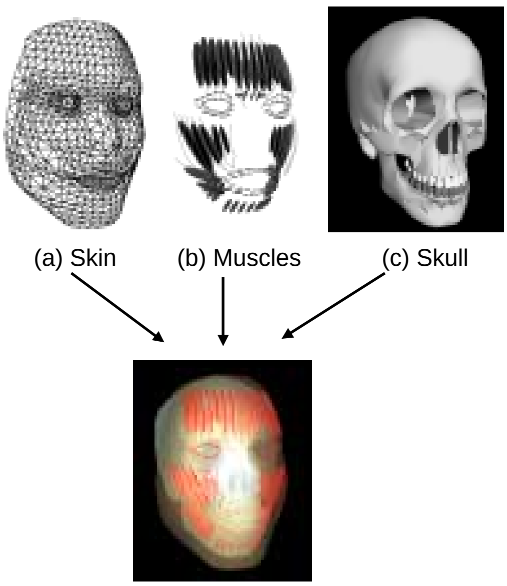


Figure 1. Segmentation software “Mimics”



(a) Skin (b) Muscles (c) Skull

(d) Integrated head model

Figure 2. The hierarchical head model

Table 2. Detailed information of the model

Layer	Data	Property
Skin	751 points, 1288 polygons	elastic
Muscles	14 facial muscles	elastic
Skull	5072 points, 7757polygons	rigid

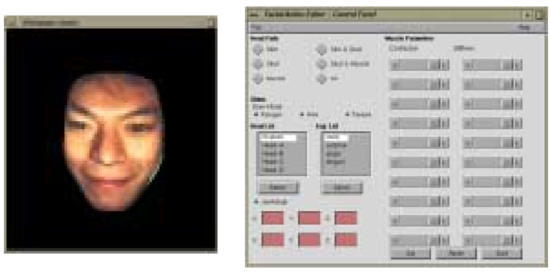
2.2 Facial Dynamics

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).

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = -\sum_{i,j} k_{ij}(c_r)\left(\mathbf{r}_i - \mathbf{r}_j\right) + m_i\mathbf{g} - R\frac{d\mathbf{r}_i}{dt} \qquad (1)$$

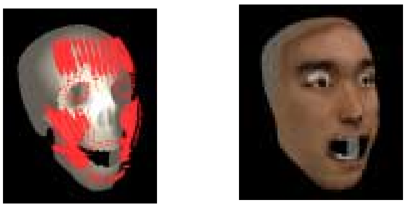
Where r_i represents the 3D coordinates of the feature point i. The first and second terms represent the total elastic force affected to the point i and the attraction of gravity, respectively. The value of k_{ij} is changed gradually by the contraction parameter C_r of each spring so that it can represent 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 (Fig3.(b)). Facial actions are parameterized relative to typical facial expressions based on anatomical analythis.



(a) Generated image (b) Facial action editor

Figure 3. Facial action synthesizer



(a) Muscle and jaw action (b) Facial surface

Figure 4. Facial Expression

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 apperance.

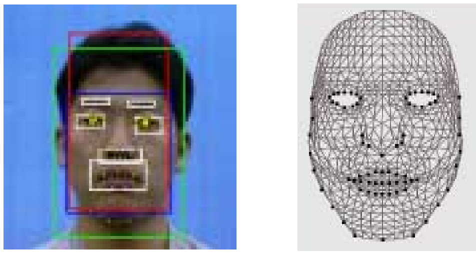
3.1 Feature Extraction

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 appoximate 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 thier 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.

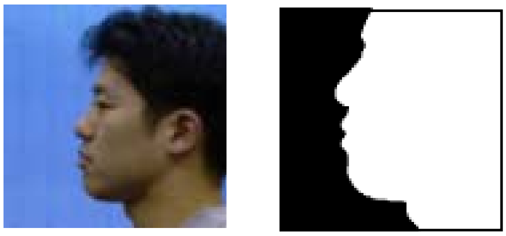


(a) Feature points (b) Corresponding points

Figure 5. Feature extraction from a frontal 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 thoresholding, if the input image has a simple color. The input image and the detected facial silhouette are shown in Fig.6.



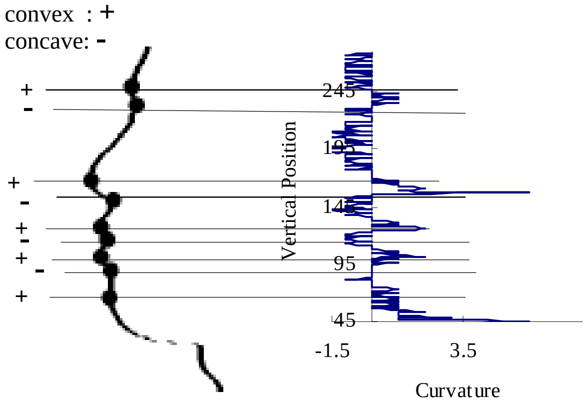
(a) Input image (b) Segmented facial profile

Figure 6. Facial profile detection

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).

$$C_s = \frac{dD_s}{ds} \qquad D_s = P_{i+1} - P_i \tag{2}$$

Where P_i represents the coordinates of the point on the facial profile, and D_s represents the difference between two adjusent 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



(a) Feature points (b) Curvature

Figure 7. Feature points detection

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.

Figure 8. Reconstructed face model

In this case, shape data of a skull is unavailable to reconstruct the skull model. Therefore, we deform the skull model roughly by automatic size adjustment which derived from the facial size. Then the slight position adjustment between a face and a skull is needed by manual operation by the user.

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.

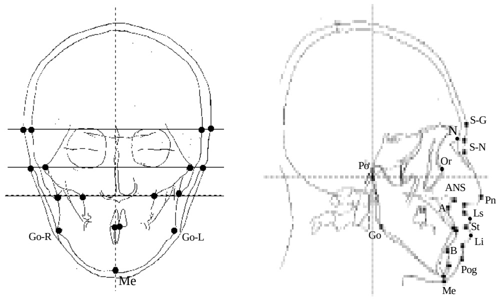
4.1 Measurement Points Extraction

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 skeches 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 Mesurement

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.



(a) Frontal view (b) Side view

Figure 9. Measurement points

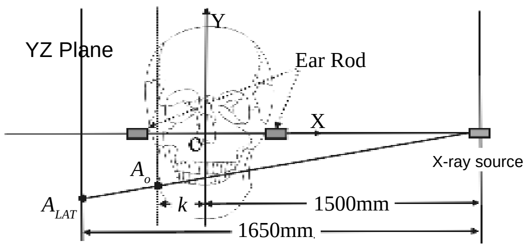


Figure10. X-ray projection

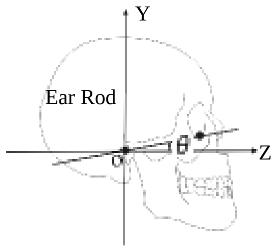


Figure 11. Head rotation

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_o (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_o 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.

$$y_o = y_{LAT} \times \frac{1500+k}{1650} \qquad z_o = z_{LAT} \times \frac{1500+k}{1650} \tag{3}$$

Although the head is fixed by the ear rods, slight head rotation around horizontal axis (Fig.11) is unavoidable 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.

$$y_{PA} = \frac{1650\,(1500+k)(z_{LAT}\sin\theta + y_{LAT}\cos\theta)}{1500\times 1650 + (1500+k)(z_{LAT}\cos\theta - y_{LAT}\sin\theta)} \tag{4}$$

$$x_o = x_{PA} \times \frac{1500+z_o}{1650} \tag{5}$$

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.

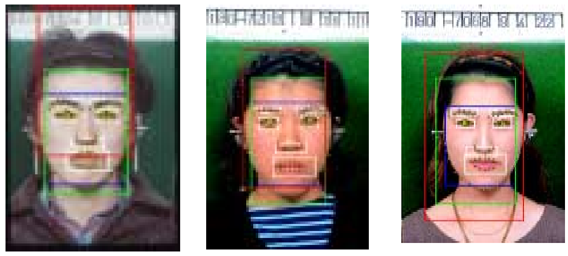
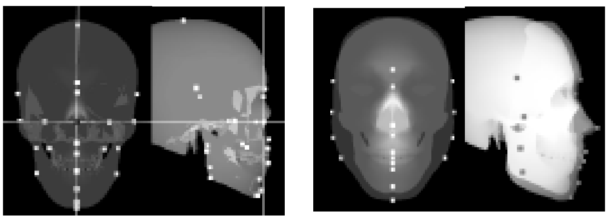


Figure 12. Facail feature points extraction from normalized facial images

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.



(a) Skull (b) Face

Figure 13. Measurement Points

4.3.2 Deformaion Method

Using the width measured from frontal image (x_1, x_2) and the corresponding width of the initial head model ($x_{1(std)}$, $x_{2(std)}$), 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_n represents the number of steps between point 1 and 2. After width and height fitting, the coordinates of (y_o, z_o) 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).

$$x_{ri} = \frac{i\, x_{r2} + (s_n - i)\, x_{r1}}{s_n} \tag{6}$$

$$\Delta \boldsymbol{r}_i = \frac{i\, \boldsymbol{p}_2 + (s_n - i)\, \boldsymbol{p}_1}{s_n} \tag{7}$$

$\Delta\boldsymbol{r}_i$: The movement vector of the point i

$\boldsymbol{p}_1$, $\boldsymbol{p}_2$: The movement vectors of measurement points

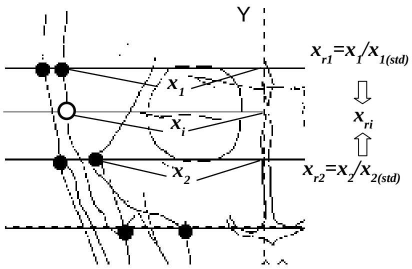
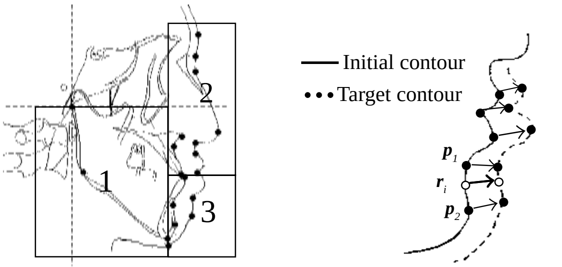


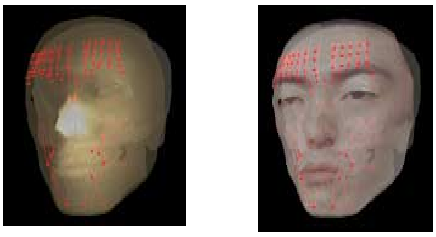
Figure 14. Fitting of the width



(a) Area division (b) Vector interpolation

Figure 15. Fitting of height and depth

Through these steps, the target head model can be reconstructed using shape features extracted from two X-ray images and a frontal facial image. Fig.16 shows the example of the fitting result for an actual patient, (a) is the result deformed only by information from X-ray images. (b) is the last result of detailed fitting using both of X-ray images and a facial image. The improvement of the quality can be seen from this result.



(a) Rough fitting (b) Detailed Fitting

Figure 16. Reconstruction result

5. Reconstruction from 3D Digitizer

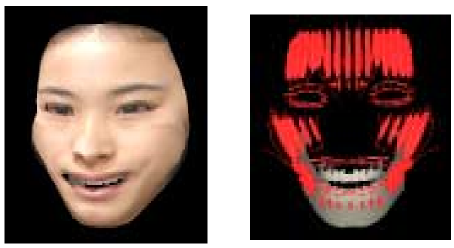
The highest accurate modeling introduced in this paper is using a non-contact 3D Digitizer for a facial surface and X-ray images for a skull modeling. Using the 3D Digitizer (VIVID700, Minolta, Japan), the 3D points data and their corresponding texture image can be captured simultanously. Therefore, 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.

Figure 17. 3D model from 3D Digitizer

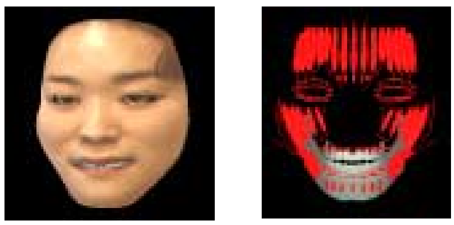
6. Applications

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)”.



(a) Person A (Left: “Smiling” face, Right: Internal state)



(b) Person B (Left: “Smiling” face, Right: Internal state)

Figure 18. Muscular action analysis

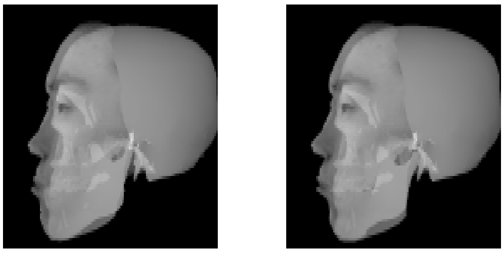
6.2 Facial Surgery Simulation

For the medical application of our head model, we developed a facial surgery system for orthgnathic surgery (in cooperation from Kyushu Univ.). The prediction 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.

(a) Before surgery (b) After surgery

Figure 19. Result of orthognathic surgery



(a) Before surgery (b) After surgery

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

[1] M.Gotla, Z.Huang, "A Minimalist Approach to Facial Reconstruction," Proceedings of the 5th Int'l Conference on Multi Media Modeling (MMM'99), pp.377-387, Ottawa, 1999.

[2] B.Dariush, S.B.Kang, K.Waters, "Spatiotemporal analysis of faceprofiles: Detection, segmentation, and registration," Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition (FG'98), pp.248-253, Nara, 1998.

Y.Lee, D.Terzopoulos et al (1995) Realistic Modeling for Facial Animation. SIGGRAPH 95 Proceedings, pp.55-62.

Y.Aoki, S.Hashimoto (1998) Physical Facial Model based on 3D-CT Data for Facial Image Analysis and Synthesis. Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98), pp.448-453.

Y.Aoki, S.Hashimoto (1999) Physical Modeling of Face Using Spring Frame based on Anatomical data. Trans. on IEICE (A), 82-A (4):573-582.

Y.Aoki, S.Hashimoto, M.Terajima, A.Nakasima (1999) Simulation of Postoperative 3D Facial Morphology using Physics-base Head Model. Proceedings of the International Conference on Multi Media Modeling (MMM'99), pp.377-387.

T.Yokoyama et al. (1999) Extracting Contours and Features from Frontal Face Images. The Journal of the Institute of Image Information and Television Engineers, 53(11): 1605-1614.

[11] D.Terzopoulos and K.Waters, "Analysis And Synthesis of Facial Image Sequences Using Physical and Anatomical Models," IEEE Trans. on PAMI, Vol.15, No.6, pp.569-579, June 1993.

[12] K.Waters, "A Physical Model of Facial Tissue and Muscle Articulation Derived from Computer Tomography Data," Proceeding of the Second Conference Visualization in Biomedical Computing, Chapel Hill, N.Carolina, SPIE, Vol.1808, pp.574-583, 1992.

[13] W.Lee, P.Kalra, N.M.Thalmann, "Model Based Face Reconstruction for Animation," Proceeding of the 4th International Conference on Multi-Media Modeling (MMM1.

S.Hashimoto, Y.Sato and H.Oda, "Modeification of Facial Expression Using Spring Frame Model", Proceedings of IASTED, pp.9-15, 1989.

'97), pp.323-338, 1997.

K.Aizawa, H.Harashima, T.Saito (1989) Model-based analysis synthesis image coding system for a person's face. Signal Processing: Image Commun., 1(2):139-152.

S.Morishima, T.Ishikawa, D.Terzopoulos (1998) Facial Muscle Parameter Decision from 2D Frontal Image", Proceeding of 14th International Conference on Pattern Recognition (ICPR'98), Vol.I, pp.160-162, 1998.

F.I.Parke (1974) A Parametric model for Human Faces. Master's thesis, University of Utah, Salt Lake City, UT, UTEC-CSs-75-047.

P.Ekman, W.V.Friesen (1978) Manual for Facial Action Coding System. Consulting Phycologists Press, Inc., Palo Alto, CA.

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