MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

# 3—22
# Automatic Reconstruction of 3D Human Face from CT and Color Photographs

Ali Md. Haider and Toyohisha Kaneko *
Department of Information and Computer Sciences
Toyohashi University of Technology

## Abstract

This paper proposes an automatic method for reconstructing a realistic 3D facial image from CT (computer tomography) and three color photographs: front, left and right views, which can be linked easily with the underlying bone and soft tissue models. This work is done as a part of surgery simulation. The 3D facial surface derived from CT by the *marching cubes algorithm* is obviously colorless. Our task is to add the color texture of the same patient actually taken with a digital camera to the colorless 3D surface. To do this it needs an accurate registration between the 3D facial image and the color photograph. Our approach is to set up a virtual camera around the 3D facial surface to register the virtual camera images with the corresponding color photographs by automatically adjusting seven parameters of the virtual camera. The camera parameters consists of three rotations, three translations and one scale factor. The registration algorithm has been developed based upon Besl and McKay's *iterative closest point* (ICP) algorithm. Successful experiments on three persons have been carried out.

## 1 Introduction

The task of accurately modeling the expressive human face by computer remains a major challenge. An important problem is the creation of a 3D facial model of a specific person. Current technology allows us to generate a precise 3D shape of an object easily by using a 3D shape acquisition system such as a Cyberware Color Digitizer$^{TM}$. However the model obtained from this system consists of only the outer shape, which is not sufficient for generating human animation. For example, the animation of a face, changing its expression requires a model that includes shape data for underlying bone and soft tissues in addition to the shape data for the face surface. The model must contain structure data for controlling eye direction, eyelid opening, jaw rotation, mouth shape and so on[3].

We propose an automatic method for creating a realistic 3D colored human face of a specific person that can be easily linked with the underlying bone shape data and soft tissues. The problem we address in the paper is how to map the color photographs onto the 3D face surface as accurately as possible. Our basic idea for mapping is to set up a virtual camera around the 3D face surface and to find three sets of camera parameters for the three photographs so that for each color photograph the corresponding perspective image on the virtual camera looks the same.

In our previous work[1] we carried out the same objective by performing the registration task based on some manually selected (semi-automatically) facial land mark points. The relative transformations between the point sets with known correspondence were calculated automatically by using *Simplex Method*.

Initially we tried to automate our previous work by making the landmark selection process automatic. However we found that some landmarks like eye-mid and mouth-top point in the color photograph were very difficult to identify automatically. Instead we came to a conclusion that the line features such as external contours and internal contours like eyes, nose, mouth and ears from both the 3D image and color photographs have better utilization for matching.

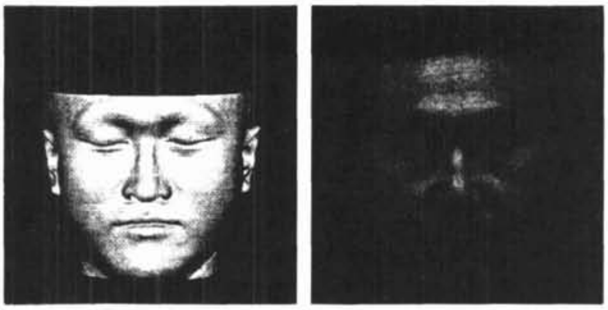
## 2 Registration Algorithm

The basic problem of our registration algorithm is to match two face pictures shown in Fig.1 by adjusting the direction, position and magnification of the virtual camera which is set in front of the 3D face in Fig.1(a). 3D face surface has been reconstructed from the CT slices using the *Volume-Preserving Marching Cubes Algorithm*[5]. In this paper we decided to utilize edges which can be detected automatically. Typical examples of edges are outlines of the face against the background and internal edges from the eyes, mouth, ears and nose.

We used color photographs with approximately known lighting conditions also we know the approximate orientation of the 3D face surface data. We made the following note-worthy features:

**(1)** The perspective image of the 3D CT face surface and the corresponding photographic image

*1-1, Hibarigaoka, Tempaku-cho, Toyohashi 441-8580, Japan. E-mail: h_ali or kaneko@mmip.tutics.tut.ac.jp



(a) 2D Perspective of 3D (b) Photograph

Figure 1: 2D and 3D front Face

should looks similar if the virtual camera parameters and lighting conditions are approximately equal to that of original photograph.

**(2)** The edge shapes from both the images should also looks similar if same edge detector is applied on both the images.

## 2.1 Edge Based Registration

Edge-based registration involves the determination of corresponding edges in different images and estimation of relative transformations using these corresponding edges. The main difference between previous point-based registration and this edge-based registration algorithm is in the availability of point correspondence information.

Let 3D edge shape consists of a points set $P$ and the 2D edge shape consists of $\chi$, where there is no correspondence between them. The general approach to search iteratively for the rigid-body transformation $\tau'$ that minimizes the cost function[6]

$$d(\tau') = \sqrt{\frac{1}{N_p}\sum_{j=1}^{N_p} ||x_j - \tau'(p_j)||^2} \quad (1)$$

where $N_p$ is the number of points in $P$, $\tau' = \varrho\tau$, $\tau$ is 3D transformation function, $\varrho$ is the perspective projection function and

$$x_j = \Re[\tau'(p_j), \chi] \quad (2)$$

where $x_j$ is a point in the 2D edge shape $\chi$ "corresponding" to the 3D point $p_j$ and $\Re$ is a "correspondence" function. The point set $P$ and the corresponding points set $X \in \chi$ have been called respectively the "hat" and "head" or "data" shape and "model" shape.

## 2.2 Besl and McKay's ICP Registration Algorithm

Best and McKay[2] presented a general purpose, representation independent, shape-based registration technique which they call the "iteration closest point" (ICP) algorithm. This method can be used with a variety of geometric primitives including point sets and line segment sets. Almost all cases are handled by first assigning one shape to be the "data" shape and the other shape to be the "model" shape. The data shape must decomposed into a point set (if it is not already in point set form). Then the data shape is registered to the model shape by iteratively finding model points closest to the data points. The technique has been noted by several workers in the medical image registration field[6], and recently implemented for image-to-physical space registration.

## 2.3 Point to Line-segment Distance

In this subsection, the method for computing the closest point to a given point on the point set or line-segment set is described. Additional information is found in[2].

The Euclidean distance $\delta(\vec{x}_1, \vec{x}_2)$ between the two 2D points $\vec{x}_1 = (u_1, v_1)$ and $\vec{x}_2 = (u_2, v_2)$ is $\delta(\vec{x}_1, \vec{x}_2) = ||\vec{x}_1 - \vec{x}_2|| = \sqrt{(u_2-u_1)^2 + (v_2-v_1)^2}$. Let $\chi$ be a 2D points set with $N_a$ points denoted by $\vec{x}_i$: $\chi = \{\vec{x}_i\}$ for $i = 1, ......N_a$. The distance between the projected 3D point $\vec{p}$ and the line segments set $\chi$ is

$$\delta(\vec{p}, \chi) = \min_{i \in \{1,\ldots,N_a\}} \delta(\vec{p}, \vec{x}_i) \quad (3)$$

The closest point $\vec{x}_j$ of $\chi$ satisfy the equality $\delta(\vec{p}, \vec{x}_j) = \delta(\vec{p}, \chi)$.

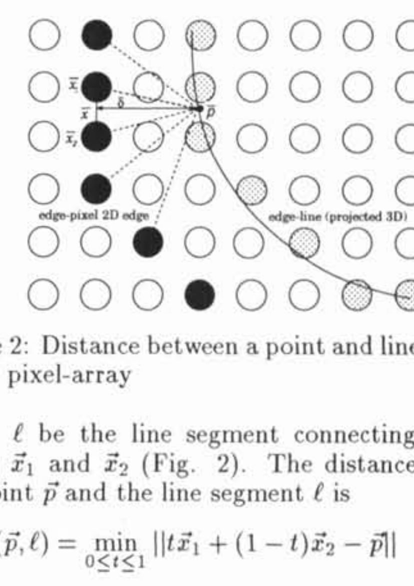



Figure 2: Distance between a point and line-segment on 2D pixel-array

Let $\ell$ be the line segment connecting the two points $\vec{x}_1$ and $\vec{x}_2$ (Fig. 2). The distance between the point $\vec{p}$ and the line segment $\ell$ is

$$\delta(\vec{p}, \ell) = \min_{0 \le t \le 1} ||t\vec{x}_1 + (1-t)\vec{x}_2 - \vec{p}|| \quad (4)$$

Let $L$ be the set of $N_\ell$ line-segments denoted by $\ell_i$, and let $L = \{\ell_i\}$ for $i = 1, ...., N_\ell$. The distance between the point $\vec{p}$ and the line-segment set $L$ is

$$\delta(\vec{p}, L) = \min_{i \in \{1,\ldots,N_\ell\}} \delta(\vec{p}, \ell_i) \quad (5)$$

The closest point $\vec{x}_j$ on the line segment $L$ satisfies the equality $\delta(\vec{p}, \vec{x}) = \delta(\vec{p}, L)$.

The method of computing the closest point from a given point to the edge shape consisting points set or line segments set have been outlined. The

3D transformation and projection function have already been outlined in our previous work[1], where the parameters are three rotations $R_x, R_y$ and $R_z$, three translations $T_x, T_y$ and $d$ (camera distance or $T_z$) and the last parameter is scaling factor $l$, which is the position of the projection plane.

In the description of the algorithm, a 3D edge points set $P$ is moved (registered, positioned called "data" shape) to be the best matched with a 2D line segments set $L$ (called "model"shape).

To guarantee the global minimum for our edge-based matching, the proposed algorithm will perform two stage registration. In the first stage registration, called coarse registration, we have assumed only four parameters as variables, they are: one rotation ($R_z$), two translations ($T_x$ and $T_y$) and scaling factor $l$. In the second stage, called fine registration, all the seven unknown parameters have been assumed as variables. In each stage at-least four initial states have been assumed. The selection of initial states or starting points are based on the general rule of thumb comes from the from the matching results of many faces. This technique may not applied for any kind of shape matching but almost accurate convergence is guaranteed for 3D to 2D edge matching in the projected space. The convergence towards global minimum also partially depends on the edge similarity between the two edges.

Each iteration loop of the proposed registration algorithm can be stated as:

(1) Extract edge from the perspective image of 3D face using initial camera parameters;
(2) back-project the edge on the 3D face to extract 3D edge;
(3) For each 3D edge-point in projected plane, compute the corresponding point (closest point) in the 2D photographic edge;
(4) Compute camera parameters by minimizing errors between two points sets in (3);
(5) If the error is less than a predefined value then exit;
(6) Else update camera parameters;
(7) If change in any of the rotation angle is more 5 deg. then goto (1), Else goto (3);

## 2.4 Edge Detection

The edge quality should be one-pixel thick with better localization accuracy for the proposed edge-based matching.

### 2.4.1 2D Edge Detection

Color photographic image has been separated into three primary color filtered images. Simple intensity gradient edge detector has been applied on each filtered image separately. Percentage histogram for edge thresholding has been selected assuming approximately same lighting conditions for all photographs. Edge-direction based edge thinning have been utilized. Finally edges obtained from three different primary color images have been OR'ed. Necessary image enhancement, background separation and smoothing operations were performed as preprocessing. Edges obtained in this way had a good localization accuracy.

### 2.4.2 3D Edge Detection

One of the requirements for the proposed edge based matching is edge similarity. The better edge similarity, the better matching accuracy. Keeping this in mind 3D edge detection algorithm has been developed. The steps for 3D edge detector are;

**(1)** The same edge detector have been applied on both the perspective image of 3D face and corresponding color photograph.

**(2)** Obtained perspective edge shape was back-projected on the 3D face surface to select the edge vertices from the visible surface.

Eye, nose and mouth edges from the front face were detected using the above technique. For facial outlines from both the front and side faces and ear edge from the side faces were extracted by using the technique shown in Fig.3.

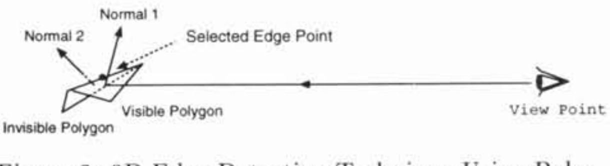



Figure 3: 3D Edge Detection Technique Using Polygon Position

Obtained edge-shapes are changeable in the projected space according to the new camera-parameters within a limited range. The 2D and 3D edge for front are shown in Fig. 4.

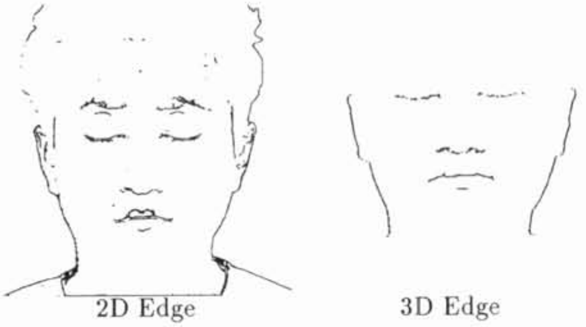



Figure 4: 2D and 3D front edge

In 3D front edge (Fig. 4) the upper-part of the side-edges from the ear-top have been deleted by automatically detecting the ear-top point from the 3D surface data. This is due to the dis-similarity with the 2D correspondence.

## 3 Mapping Process

The color mixing and mapping process has been elaborately described in our previous paper[1].

## 4 Experimental Results

Experiments were carried out on three persons and it was found that the process was operable totally automatically. The virtual camera parameters were identifiable separately with the initial rotation zero degree for the front-view and $\pm 90^0$ for the right and left view. The camera distance (Z translation) could be estimated even with a 50% estimation error.

Before Coarse Fine

Figure 5: Edge-based Registration Results for Front-Face

The coarse and final registration obtained from the proposed registration algorithm is given in Fig. 5. A list of relative transformations is also given in table 1.

| Parameters | Before | Coarse | Fine |
|---|---|---|---|
| Rotation $R_x$(deg) | 0 | 0* | 4.170 |
| Rotation $R_y$(deg) | 0 | 0* | 0.820 |
| Rotation $R_z$(deg) | 0 | 1.324 | 0.907 |
| Transla. $T_x$(mm) | 0 | -1.243 | -0.768 |
| Transla. $T_y$(mm) | 0 | -38.696 | -33.998 |
| Camera Dist. $d$(mm) | 2000 | 2000* | 1504.51 |
| Scale Factor $l$(mm) | 120 | 126.903 | 122.260 |
| Average Error (mm) | 4.631 | 0.722 | 0.480 |
| Exec. Time (min) | | 8 | 20 |

[* not used as variables in coarse registration]

Table 1: Parameters for Front Face Matching

The CT data we employed is a size of $512 \times 512 \times 175$ with a resolution of $0.46mm \times 0.46mm \times 1mm$. The $480 \times 480$ pixels color photographs were taken with a SONY digital video camera of $640 \times 480$ resolution. A blue sheet of approximately $1m \times 1m$ was used as the background of taking photographs.

The input 2D images and the resulting 3D images is shown in Figs.6 and 7.

## 5 Conclusion and Future Works

We have succeeded in creating 3D realistic faces by estimating the camera parameters automatically.

Left Front Right

Figure 6: Color photographs of resolution $480 \times 480$

$45^0$ Left $0^0$ $45^0$ Right

Figure 7: Resulting 3D images

No special trained person is needed to perform the operation.

Our ultimate goal is to develop a system for predicting realistic post-surgical appearance of a cancer patient. This requires modeling the internal structures such as soft tissue and bone. A change in the underlying structures will alter the facial skin appearance before surgery, which we have dealt with in this paper. Such information will benefit both a doctor and his patient for selecting a surgery alternative for the mutual understanding.

## References


[1] Ali Md. Haider, Eiji Takahashi and Toyohisha Kaneko, "A 3D Face Reconstruction Method from CT Image and Color Photographs", will be published in IEICE Trans. on Inf. and Systems Vol. E81-D, No.10, October 1998.

[2] Paul J. Besl and Neil D. McKay, "A Method for Registration of 3-D Shapes", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-14, No. 2, Feb. 1992, pp. 239-257.

[3] T. Akimoto and Y. Suenaga, "Automatic Creation of 3D Facial Model", IEEE Computer Graphics and Applications, Vol. 13, No. 4, Sep. 1993, pp.16-22.

[4] R.M. Koch, M.H. Gross and others, "Simulating Facial Surgery Using Finite Element Models", SIGGRAPH 96, New Orleans, Louisiana, August 4-9, 1996, pp.421-429.

[5] T. Kaneko and Y. Yamamoto, "A Surface Generation Method with Volume Preservation", Trans. of the IEICE, Vol. J81-D-II, No.1, January 1998, pp.101-108 (in Japanese).

[6] Calvin R. Maurer, Jr., Georges B. Aboutanos, Benoit M. Dawant, Robert J. Macuiunas, J. Michael Fitzpatrick, "Registration of 3-D Images Using Weighted Geometrical Features", IEEE Trans. on Medical Imaging", Vol. 15, No. 6, Dec. 1996, pp. 836-849.