MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

FEATURE EXTRACTION FROM FRONT AND SIDE VIEWS OF FACES FOR 3D FACIAL MODEL CREATION

Taka-aki Akimoto, Richard Wallace and Yasuhito Suenaga

NTT Human Interface Laboratories
Nippon Telegraph and Telephone Corporation
1-2356 Take Yokosuka kanagawa 238-03 Japan
e-mail: akimoto%nttcvg.ntt.jp@relay.cs.net

ABSTRACT

This paper describes a method to extract facial features from front and side views for 3D facial model creation by modifying a 3D generic head model. Front and side views of a person's face are taken simultaneously by two cameras under the same geometric condition. The side view (profile) is first analyzed using a DP template matching to extract vertical locations of nose and mouth. Then, by using the vertical location data, the existing areas of the eyes, nose and mouth are estimated on front and side view images, and such facial features are extracted as strong edges in those areas. The extracted facial features, including the outline of the hair and face, are used to create a 3D facial model for the person by modifying a generic head model.

1. Introduction

The generation of human images[1,2,3,4] is an attractive subject in computer graphics because of its wide applications in various fields. However, several problems must be solved in establishing a human image generation technique. One important problem is to create a 3D model of a specific person.

The authors have been developing an automatic 3D facial model creation system[5]. In this system, the data which expresses the shape of a face are acquired from front and side views. A generic head model, which has been created in advance, is then modified based on the acquired data. As a result, the 3D facial model of a specific person is obtained. The details of the method are described in the following sections.

2. Automatic 3D facial model creation system

Figure 1 shows a block diagram of our automatic 3D facial model creation system. The entire system is divided into two parts; feature extraction part and generic model modification part.

In the feature extraction part, front and side view images are segmented according to pixel color, and the background, hair and face region are identified. Next, the outline of the face from a side view, called a profile, is analyzed to estimate the position of facial features such as eyes, mouth or chin. The position data allows the eyes, nose and mouth to be extracted from both images. The chin outline from a front view image is also estimated using the position data.

In the generic model modification part, the generic head model is modified according to the position and shape of the facial features acquired in the feature extraction part. The generic head model is a 3D model that has a typical face shape. It is created in advance by using a cylindrical laser scanner[6].

This paper focuses on the feature extraction part of this system. The characteristics of our approach are as follows.

(1) Feature extraction is done on the front and side views of a human face that are taken simultaneously under the same geometric condition. Therefore, information acquired from the side view can be used for feature extraction from the front view.

(2) A DP template matching is used to reliably identity features on the profile such as chin tip, mouth, nose tip and nose bridge.

3. Preprocessing

3.1 Front and side view images

Front and side views are taken simultaneously by two TV cameras, and digitized to 512 by 480 pixel color images. The cameras are arranged so that the size and position of the face in both front and side views are the same. In order to make background extraction easier, a white background was used in the experiments. Figure 2 shows the typical captured images.

3.2 Segmentation of images using color

To extract large regions such as background, hair or face (skin area), the digitized images are segmented into regions based on pixel color. First, all pixels in each image are classified by the MDL clustering method[7] in the RGB color space. Then, regions are formed by gathering pixels of the same cluster. Very small regions or regions whose average color is very close to that of neighboring regions are merged to those regions. Figure 3 shows the segmented images.

3.3 Identification of background, hair and face regions

Since background, face and hair regions normally occupy the largest image areas, such regions can be identified easily and reliably. After that, small facial features such as eyes or mouth are extracted from the face region with empirical knowledge.

First, the region which has the biggest sum of height and width is identified as a background. Then, the two regions excluding the background are picked up, and their relative positions are checked whether they satisfy predefined conditions or not.

Figure 2. Input images.

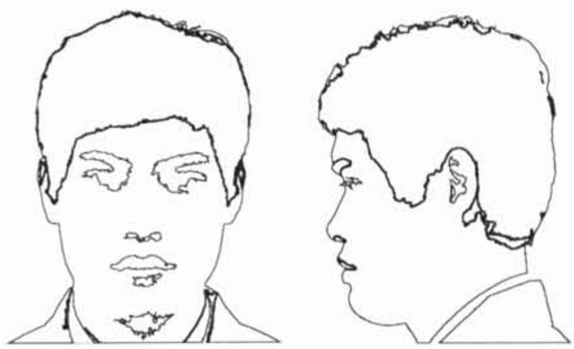

Figure 3. Results of image segmentation.

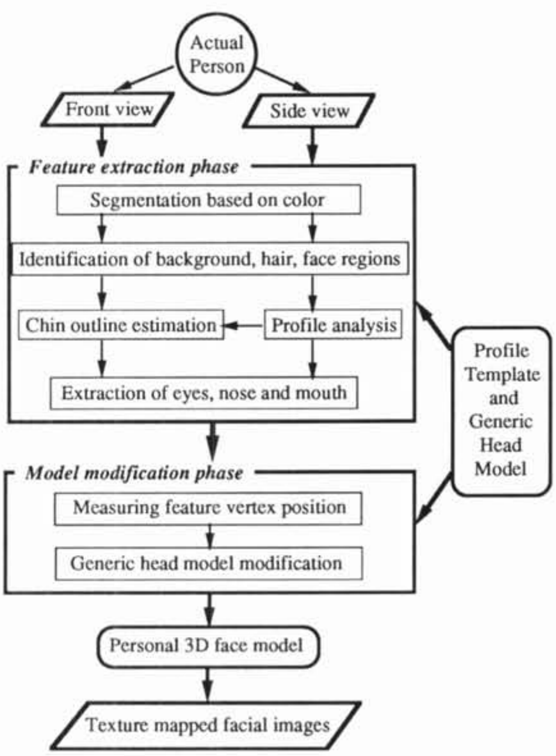

Figure 1. Automatic 3D facial model creation system.

In the region pairs that pass the above check, one pair having the largest area is identified as hair and face regions.

4. Profile analysis and facial feature extraction

4.1 Profile analysis by template matching

A facial profile contains useful information[8]. The positions of the facial features can be estimated from the profile curve. Moreover, the profile can be extracted easily, because it is the boundary between the face region and the white background. Thus, the vertical positions of facial features are estimated by analyzing the profile curve. To do this, we employed the template matching approach. The template consists of about 50 points and line segments, and represents a generic profile shape. The template also includes the position data of the bridge of nose, the tip of nose, the bottom of nose, mouth and chin.

The profile in the side view image is extracted as a chain of pixels. Therefore, template matching is accomplished by determining the pixel p_i' in the profile corresponding to each point p_i in the template as shown in Figure 4.

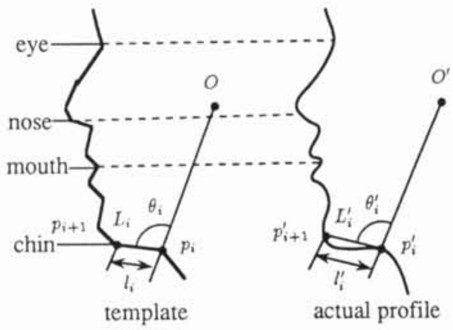

Figure 4. profile analysis.

Figure 5. Result of profile analysis.

This problem is solved by finding the set of pixels $P' = \{p'_1, p'_2, \ldots, p'_n\}$ which gives a minimum value to the function $f(P')$ shown in equation (1).

$$f(P') = \sum_{i=1}^{n}(w_1(\theta_i - \theta'_i)^2 + w_2(l_i - l'_i)^2). \qquad (1)$$

In this equation, as shown in Figure 4, l_i and l'_i are the length of segment L_i and L'_i. L_i is the line between the vertex p_i and p_{i+1} in the template, and L'_i is the line between the pixel p'_i and p'_{i+1} in the profile. θ_i and θ'_i are the angles between segment L_i and line $\overline{p_i O}$, and between segment L'_i and line $\overline{p'_i O'}$, respectively. O' is the center of the hair and face area in the side view. O is a predefined center point in the template. w_1 and w_2 are constants.

Dynamic programming is used to find the global minimum of the function $f(P')$. Figure 5 shows the template matching result. The curve in the left hand side of this figure indicates the template. Lines connecting the template and the profile indicate the corresponding points.

4.2 Facial feature extraction

By using the vertical positions of the eyes, nose, mouth, and chin obtained by the profile analysis, we can limit the search area for facial feature extraction. The rectangular windows in Figure 6 show the search areas. The position and size of each area is determined based on the vertical position and the typical size of the facial features.

Facial features are extracted by finding strong edges in the search areas. First, the region in each area is filtered with the Sobel operator. Then, the strong edges are extracted by thresholding the pixel values against a predefined threshold. Lines within the boxes in Figure 6 show the extracted facial features.

The chin outline illustrated in Figure 6 is estimated by drawing a spline curve which passes through the bottom of chin, right and left cheek outlines, and pixels having higher edge strength.

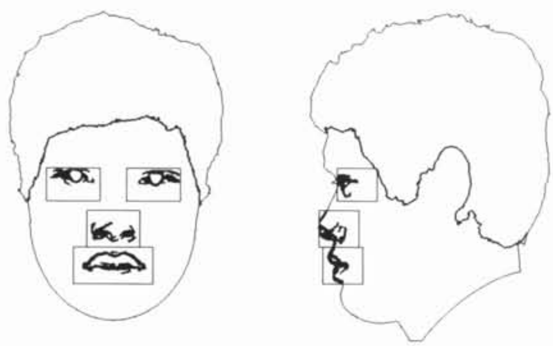

Figure 6. Search areas and extracted features.

5. Experimental results

An experiment was done for eight pairs of front and side views of different people. All the people used in this experiment were men with short hair, no glasses, and no mustache.

Figures 7(1-a), (2-a) and (3-a) show the results of the profile analysis, and Figures 7(1-b), (2-b) and (3-b) show search areas and extracted facial features. Figure 8 shows a 3D facial model and texture-mapped facial image created from the results shown in Figure 6 by our automatic 3D facial model creation system.

As shown in Figure 7, the profile analysis by template matching worked well for all images. Therefore, search areas for finding the eyes, nose and mouth could be determined correctly for all images.

6. Conclusions

This paper described a method of extracting facial features from front and side views in our automatic 3D facial model creation system. By analyzing a profile curve by template matching, the position of the eyes, nose, mouth and chin in facial images are estimated. Then, the areas for searching such facial features are determined using this position data, and facial features are extracted as strong edges in the search areas. These extracted features are used to modify a generic head model. As a result, a 3D facial model for the person is obtained.

Our current program worked quite well for all images if the face areas were identified correctly and the edges of the eyes, nose and mouth were strong enough to extract. However, for vague images taken under bad conditions, more precise methods like deformable template[9] may be needed. Finding more sophisticated methods for extracting exact outline of each facial feature is an important future work.

Acknowledgements

We would like to thank Dr.Yukio Kobayashi, the executive manager of Visual Perception Laboratory in NTT Human Interface Laboratories, and all the members of our laboratory for their encouragements and valuable discussions. We also thank Mr. Akihiko Hashimoto and Mr. Takashi Yamana for their cooperation in inputting facial image data.

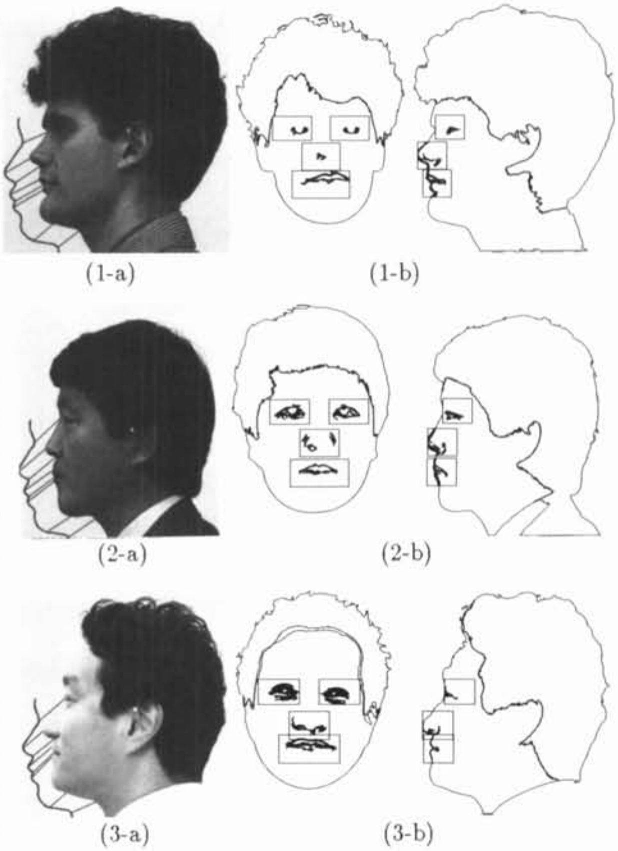

Figure 7. Experimental results for other person.

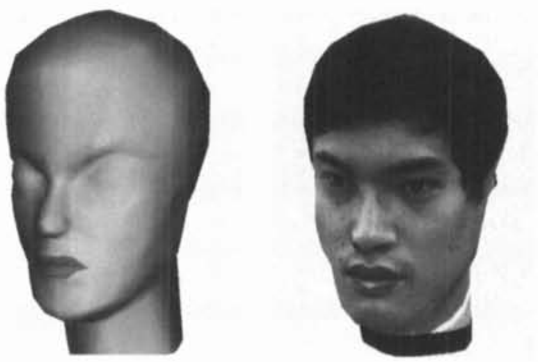

Figure 8. 3D facial model and texture-mapped facial image created by our automatic 3D facial model creation system.

References

[1] F. Parke: "Parameterized Models for Facial Animation", IEEE, CG & A, Vol.9, No.2, pp.61-68 (1982).

[2] S. Platt and N. Badler: "Animating Facial Expressions", Computer Graphics, Vol.15, No.3, pp.245-252 (1981).

[3] J. Lewis and F. Parke: "Automated Lip-Sync and Speech Synthesis for Character Animation", CHI+GI conference proceedings (1987).

[4] L. Williams: "Performance-driven Facial Animation", Computer Graphics, Vol.24, No.4, pp.235-242 (1990).

[5] T. Akimoto, R. Wallace and Y. Suenaga: "Automatic Creation of Face Model for Human Image Generation from Front and Side Views", Proc. ImageCom'90, (November. 1990).

[6] Y. Suenaga and Y. Watanabe: "A Method for the Synchronized Acquisition of Cylindrical Range and Color Data", IAPR Workshop on MVA (November 1990).

[7] R. Wallace and T. Kanade: "Finding Natural Clusters having Minimum Description Length", 10th-ICPR, pp.438-442 (June 1990).

[8] L. Harmon: "Automatic Recognition of Human Face Profiles", Computer graphics and image processing, Vol. 6, pp.135-156 (1977).

[9] A. Yuille, D. Cohen and P. Hallinan: "Feature Extraction from Faces using Deformable Templates", Proc. 1989 CVPR , pp.104-109.