From Coarse to Fine Correspondence of 3-D Facial Images and its Application to 3-D Facial Caricaturing

Takayuki Kondo  Kazuhito Murakami  Hiroyasu Koshimizu
School of Computer and Cognitive Sciences
Chukyo University
101 Tokodate, Kaizu-cho, Toyota, 470-03 Japan
kondo@koshi-lab.sccs.chukyo-u.ac.jp

Abstract

We propose a method to generate a facial caricature depending on a simple method for corresponding 3-D facial images. In this method, we extract several facial parts regions each of which include facial parts of the face by using both gray and range images. The diagonal corners of the respective extracted regions are used to provide the informations on the correspondence between the faces. Therefore, by using this method, it is expectable to reduce the number of the feature points for the correspondence from more than several hundreds to ten or so. At the same time, this method make it possible to extract boundaries of facial parts.

In order to examine the feasibility of this method, we employ a usual corresponding method of the triangular patch as the reference. We generated the 3-D mean face and 3-D facial caricature to demonstrate experimentally the feasibility of the proposed method. It was clarified that the number of the correspondence points can be reduced to only 10% of the usual method.

1. Introduction

Artist and illustrator are superior in the skill for extracting facial individuality features. In the process of their face observation, it is clear that they exactly catch features of position and shape of the face, and that they exaggerate these features skillfully. Then, they draw them finally as an impressive facial caricature.

Facial caricaturing is a process of the feature extraction and feature exaggeration of faces. Not only gray-level images, we are expected to open the door to utilize more wide informations on the face such as color, motion and range informations in facial caricaturing. Now in this paper, we paid attention to the range information for the extension of the spatial dimensionality of the face. For this extension, we made use of the range image of faces and 3-D facial caricature (or facial 3-D relief) is introduced by using the range image.

Now, in case that we generate such 3-D facial caricature, it is necessary generally to designate many correspondence points in the face image. In this paper, we consider that the number of the correspondence points can be reduced by a simple method using 7 facial parts regions. At first, we introduce a simple method to extract the sub-region of a facial part automatically by using together face range image with gray-level image. Then, we experiment on the robustness of this method by extracting 3-D mean face and by generating 3-D facial caricature. Then, we confirm that 3-D facial caricature can be exaggerated by this simple method.

2. Outline of 3-D Facial Caricaturing

We propose the following fundamental principles to generate 3-D facial caricature.

(1) This proposed method can make the 3-D facial caricature by means of a few facial parts rectangle regions.

(2) This proposed method can automatically provide the correspondence points by using range image and gray-level image.
(3) We can make it clear that the proposed method can extract sufficient facial individuality informations from the 3-D image data.

Now, as shown in Figure 1, let the process of 3-D facial caricature be composed. And in this method the whole face region is divided into seven sub-regions (eye, nose, mouth, head, chin, right cheek, and left cheek). Mean faces for each sub-regions are generated respectively by averaging the respective sub-regions. Facial individuality features are extracted respectively in each sub-region by comparing with the mean sub-regions. These individuality features are exaggerated, and finally, all seven deformed faces of the each sub-region are integrated, then one 3-D facial caricature can be generated.

Original image (Gray and 3-D range images)  
Image processing  
Dividing a face into seven sub-regions  
Extracting facial features  
Deforming by in-betweening method  
Composing seven sub-regions (3-D facial caricature)

Fig.1 Diagram for generating 3-D facial caricature.

3. Regions of Facial Parts

As shown in Figure 2, face is segmented as rectangle regions. Basing on the fact that the highest point of the face of the range image becomes the point of the top of the nose, we propose the following algorithm to extract the respective 7 rectangle regions for nose, mouth, and so on in this chapter. And fine regions for the facial parts are extracted in the rectangle regions.

3.1. Algorithm to extract rectangle nose region

In this algorithm, positive values for x and y are measured to the right on the x-axis and downward on the y-axis, respectively. \( f(x,y) = \{ f(x,y) \mid 256 < x < 240 \} \) denotes face range image, and the maximum \( f_{\text{max}} \) is calculated as \( f_{\text{max}} = \max(f(x,y)) \). The maximum point \( P(x_{\text{max}}, y_{\text{max}}) \) is extracted as the peak of the nose. Investigating the range data around \( P(x_{\text{max}}, y_{\text{max}}) \), the rectangle region of the nose is extracted by means of the horizontal and vertical profiles. Precisely investigating the curvature of the profile around the point \( P(x_{\text{max}}, y_{\text{max}}) \), we extract inflection points \( x_{\text{inf}}, y_{\text{inf}} \) \((t=1,2,3),(x_{\text{inf}}, y_{\text{inf}}) (b=1,2,3,7)\). At that time, 2 horizontal lines passing through the first inflection points \( y_{t1}=y_{t1}, y_{b1}=y_{b1} \) are considered as the boundaries where the nose region is vertically sandwiched.

The gradient (Sobel operator) is calculated from a face range image \( f(x,y) \) as shown in Figure 3(a). From this gradient image, x-coordinates \( x_{t1}, x_{t2} \) between which the gradients are greater than a threshold and which are sandwiching horizontally around the peak of the nose are extracted. These coordinates give the left and right sides of the nose region (Figure 3(b)). Consequently, the nose region is obtained as the region \( \{(x,y) \mid x_{t1} < x < x_{t2} \} \). An example is shown in Figure 4(a).

We introduce an algorithm to extract fine nose region. The fine nose region can be obtained from the rectangle region \( \{(x,y) \mid x_{1} < x < x_{2}, y_{1} < y < y_{2} \} \). Let the boundary image \( f_{\text{boundary}}(x,y) \) of the nose be calculated by differentiating the range data of the rectangle region, and

\[
\begin{align*}
    g_{x} &= \begin{cases} 
        1, & f_{\text{boundary}}(x,y) \leq T \\text{or} \ f_{\text{boundary}}(x,y) > T \\
        0, & \text{otherwise}
    \end{cases} 
\end{align*}
\]

is easily introduced by thresholding. Then, \( g_{x} \) image is processed by thinning and noise reduction procedures. An example is shown in Figure 4(b). This boundary data of the nose is to be provided to generate 3-D surface patches in chapter 5.

\[ y \]

Fig.3 Examples of boundaries extracted as a rectangle region: (a) top point of nose, (b) border points of nose region.
3.2. Algorithm to extract rectangle mouth region

In the same way as chapter 3.1, we extract inflection points \((x_{\text{max}}, y^{(B)}_b)\) \((b=1,2,\ldots,7)\). At that time, 2 horizontal lines which pass through the first inflection points \((x_{(t,m)}, y^{(B)}_1), y^{(B)}_{(t,m)}\) are considered as the boundaries where the mouth region is vertically sandwiched.

\(x_{(t,m)}\) and \(y_{(t,m)}\) are extracted by the gradient and projection. Consequently, the mouth region is obtained as the region \([(x, y) | x_{(t,m)} \leq x \leq x^{(R,m)}, y^{(T,m)} \leq y \leq y^{(R,m)}]\). An example is shown in Figure 5(a).

The fine mouth region can be obtained from the rectangle region \([(x, y) | x_{(t,m)} \leq x \leq x^{(R,m)}, y^{(T,m)} \leq y \leq y^{(R,m)}]\). Let the boundary image \(f^{(x)}_{g_{xy}}\) of the mouth be calculated by differentiating the range data of the rectangle region. As shown in chapter 3.1, \(g_{xy}\) image is processed by thinning and noise reduction procedures. This example is shown in Figure 5(b).

4. 3-D Caricature Generation

This chapter presents a detailed procedures of the diagram for generating 3-D facial caricature of Figure 1. At first, the original face images are normalized in size in this method. Mean face is generated by using that sub-regions explained in chapter 3 before. And 3-D facial caricature is generated by using both the sub-regions and mean face.

4.1. Normalization

Prior to the correspondence, standard length and standard point are decided, and faces are normalized about position, angle and scale of the face. The standard length and the standard point are decided by using three standard points \((x_1, y_1), (x_2, y_2),\) and \((x_3, y_3)\) which are automatically extracted. Two end points of the mouth are \((x_1, y_1)\) and \((x_2, y_2)\). Top of the nose is \((x_3, y_3)\). Their points are extracted by the gradient and projection procedures as shown in Figure 6.

The normal vector of the plane is calculated to pass through the three standard points at first. And face data is rotated so that the direction of the normal vector reaches the specified direction, that is, the face data is rotated so that a line between the standard points \((x_3, y_3)\) and \((x_2, y_2)\) becomes level, and it is enlarged so that a length of \((x_2, y_2)\) from \((x_3, y_3)\) becomes to the specified length. Finally, the free data is moved so that \((x_3, y_3)\) hits the specified point. The face data are normalized by these Affine transform procedures.

4.2. Mean face

In order to generate 3-D facial caricature, it is necessary to generate a 3-D mean face as the reference with which 3-D individuality features are extracted. The mean face of the range image is calculated respectively for each sub-region in which the respective facial part is included. The whole mean face is composed by seven mean faces. The following procedure is proposed to generate the mean face mainly by using Affine transform.
Step 1: One sub-region is selected (i=1,2,...,7).

Step 2: The size of the rectangle sub-region is corresponded to that of the mean sub-region by Affine transform.

Step 3: After normalization of the region, the range data are averaged.

Step 4: Each mean sub-regions (i=1,2,...,7) are collected as one whole face, and one 3-D mean face is generated.

An example of the 3-D mean face which was provided by 4 range images as shown in Figure 7 is shown in Figure 8.

![Fig.7 Examples used for introducing 3-D mean face.](image)

![Fig.8 An example of 3-D mean face.](image)

4.3. Deformation for 3-D caricaturing

Let the 3-D mean face be S, the original 3-D face image be P, and parameter of an exaggeration weight for deformation be b. Basically, we think that the 3-D facial caricature Q can be generated by

\[ Q = P + b(P - S) \]  \hspace{1cm} (2)

However, not only the shape but also the location and range informations of the facial parts are to be exaggerated, we propose the following procedure for deformation.

The deformation procedure is as follows. Here, we denote the seven sub-regions \( \{(x, y) | x_{(L,i)} \leq x \leq x_{(R,i)}, y_{(T,i)} \leq y \leq y_{(B,i)} \} \) (i=1,2,...,7) that are given in chapter 3. This procedure uses the range data shown in Figure 9(a) as an original range image.

Step 1: Facial features are exaggerated by

\[ \begin{align*}
    x^{(Q)}_{(L,i)} &= x^{(P)}_{(L,i)} + b(x^{(P)}_{(L,i)} - x^{(S)}_{(L,i)}) \\
    y^{(Q)}_{(T,i)} &= y^{(P)}_{(T,i)} + b(y^{(P)}_{(T,i)} - y^{(S)}_{(T,i)}) \\
    x^{(Q)}_{(R,i)} &= x^{(P)}_{(R,i)} + b(x^{(P)}_{(R,i)} - x^{(S)}_{(R,i)}) \\
    y^{(Q)}_{(B,i)} &= y^{(P)}_{(B,i)} + b(y^{(P)}_{(B,i)} - y^{(S)}_{(B,i)})
\end{align*} \]  \hspace{1cm} (3)

for corresponding the left-upper point \((x_{(L,i)}^{(P)}, y_{(T,i)}^{(P)})\) and the right-lower point \((x_{(R,i)}^{(P)}, y_{(B,i)}^{(P)})\) of an input face P (Figure 9(a)) to left-upper point \((x_{(L,i)}^{(S)}, y_{(T,i)}^{(S)})\) and the right-lower point \((x_{(R,i)}^{(S)}, y_{(B,i)}^{(S)})\) of a mean face S. The left-upper point \((x_{(L,i)}^{(Q)}, y_{(T,i)}^{(Q)})\) and the right-lower point \((x_{(R,i)}^{(Q)}, y_{(B,i)}^{(Q)})\) of a facial caricature Q are calculated.

Step 2: By means of the range values of with an input face P and the mean face S, a range data of a facial caricature Q is calculated by

\[ z^{(Q)}_{(x_{(L,i)}^{(Q)}, y_{(T,i)}^{(Q)})} = z^{(P)}_{(x_{(L,i)}^{(P)}, y_{(T,i)}^{(P)})} + b(z^{(P)}_{(x_{(L,i)}^{(P)}, y_{(T,i)}^{(P)})} - z^{(S)}_{(x_{(L,i)}^{(S)}, y_{(T,i)}^{(S)})}) \]  \hspace{1cm} (4)

Step 3: Every facial part (i=1,2,...,7) is calculated by step 1 and step 2.

Step 4: Each facial part (i=0,1,2,...,7) is collected, and one 3-D facial caricature is generated. An example is shown in Figure 9(b).

Step 5: Finally, 3-D facial caricature is processed by smoothing filter.

![Fig.9 Examples of 3-D facial caricatures: (a) original range image, (b) 3-D facial caricature of face.](image)

Additionally, this method can exaggerate exclusively one facial part as shown in Figure 9. This can be expressed by

\[ Q = P + b\alpha_i(P_i - S_i) \]  \hspace{1cm} (5)

where \(\alpha_i\) is exaggeration weight, and \(\alpha_i\) for forehead, \(\alpha_i\) for eyes, \(\alpha_i\) for left cheek, \(\alpha_i\) for nose, \(\alpha_i\) for right cheek, \(\alpha_i\) for mouth cheek, and \(\alpha_i\) for chin. Figure 10(a) is a result of all \(\alpha=1.0\). Figure 10(b) is a typical result where the cheek is
especially exaggerated ($\alpha, \alpha_2=1.4$), and Figure 10(e) is a typical result where the nose is extraordinarily exaggerated ($\alpha_2=1.5$).

![Figure 10](image1)

**Fig. 10.** Unique examples of 3-D facial caricature: (a) 3-D facial caricature ($\alpha=1.0$), (a) 3-D facial caricature ($\alpha, \alpha_2=1.4$), (a) 3-D facial caricature ($\alpha_2=1.5$).

---

5. Experiments and Considerations

We compared the proposed simple method with the following reference method. The reference method is the common method in which the faces are corresponded by the triangular patches which are generated on the contour edges of the facial parts as shown in Figure 11.

![Figure 11](image2)

**Fig. 11.** An example of the triangular patches.

Figure 12(a) and (b)(c)(d) show the mean faces which are calculated respectively by the proposed and the reference methods. The each number of the correspondence points was 8, 64, and 205 in the reference method and one of the proposed method was 14. Therefore, it was known that no fatal degradations in the mean face appeared even when more than 90% reduction of the correspondence points was provided.

![Figure 12](image3)

**Fig. 12.** Examples of 3-D mean face: (a) by the proposed simple method (correspondence points=14), (b) by the reference method (correspondence points=8), (c) by the reference method (correspondence points=64), (d) by the reference method (correspondence points=205).

6. Conclusions

In this paper, we proposed a mechanism to generate facial caricature by using face range image and verified to generate the 3-D facial caricatures. In this experiment, the following

---

\[ \text{Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling} \]
\[ 0-8186-7943-3/97 $10.00 \odot 1997 \text{IEEE} \]
the 3-D facial caricatures. In this experiment, the following items were confirmed.

1. Basically the proposed method was able to generate 3-D facial caricature, even when more than 90% reduction of the correspondence points is introduced, and the method for extracting fine regions of the facial part from rectangle regions was also introduced.

2. On the other hand, the mean face of the proposed method is rather degraded. Simultaneously, this degradation did not yield no fatal degradation in 3-D facial caricaturing.

3. Changing independently the exaggeration weight in each facial part provides several interesting results, as shown in Figure 10(b) and (c).

We can use this method as the coming caricaturing tool. Therefore the tool can be used also for many other things than 3-D faces. In the future, we would like to improve the processing speed and would like to apply this method to the motion features of the 3-D faces.

Acknowledgments

This paper was partially supported by Grant-in-Aid for General Scientific Research (No.08207222, No.07680415, No.08780313, 1996) and HUTOP project (MITI, 1996). We would like to express deep thanks to these supports.

References