Performance Evaluation of Facial Caricaturing System PICASSO-2 and Its Revised Methodology

Takayuki Hoshino¹, Takuma Funahashi², Takayuki Fujiwara², Hiroyasu Koshimizu² and Yasuyo Hatano²

¹ Graduate School of Computer and Cognitive Sciences, Chukyo University
² School of Information Science and Technology, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota Aichi, 470-0393, Japan
E-mail: hoshino@koshi-lab.sist.chukyo-u.ac.jp

ABSTRACT

We have not yet completed the automatic recognition of facial parts by image processing. Even if once we get the edge image for the contour of facial parts, due to the existence of the noise, the results are quite frequently not sufficient both in local and global aspects. This is caused mainly by the lack the Top-Down mechanism for compensating the insufficient Bottom-Up mechanism of image processing. Then, in this paper, a curve fitting mechanism in the image processing is introduced as a Top-Down processing for enforcing the total performance of contour extraction of the facial parts. An experiment for extracting the contour line extraction of the jaw was done and some preliminary experimental results were presented in this paper, because jaw contour is especially hard to detect at the field test of EXPO2005 Exhibition. Furthermore the possibility of the new deformation basis for the facial caricature is prospected so that the parameter space of the analytical curve could be introduced.

1. Introduction

Picasso-2 [1], facial caricaturing system, is a system for representing it with the line drawings of the contour lines of the facial parts by image processing [2]. Though extracting natural Line drawings of the facial parts is essential in this system, we have not yet completed the automatic recognition of facial parts by image processing. Even if once we get the edge image for the outline of facial parts, due to the existence of the noise, the results are quite frequently not sufficient both in local and global aspects.

This is caused mainly by the lack the Top-Down mechanism for compensating the insufficient bottom-up mechanism of image processing. Then, in this paper, a curve fitting mechanism in the image processing is introduced as a Top-Down processing for enforcing the total performance of contour extraction of the facial parts.

An experiment for extracting the contour line extraction of the jaw was done and some preliminary experimental results were presented in this paper, because jaw contour is especially hard to detect at the field test of EXPO2005 Exhibition.

Based on the experimental considerations, some future subjects were also pointed, and the possibility of the new deformation basis for the facial caricature where the parameter space of the analytical curve could be introduced is prospected.

On the other hand, there are other systems and products proposed so far that can generate the caricature automatically [3-4], there has been no report on the field test of these systems. Therefore our unique and large trial could be effective for getting facial images of several races, for evaluating the performance of the system and for enforcing the robustness of the system.

An experiment for extracting the contour line extraction of the jaw was done and some preliminary experimental results were presented in this paper, because jaw contour is especially hard to detect at the field test of EXPO2005 Exhibition.

Based on the experimental considerations, some future subjects were also pointed, and the possibility of the new deformation basis for the facial caricature where the parameter space of the analytical curve could be introduced is prospected.

2. Introduction of PICASSO-2

The blue background is removed from original image at first. Skin color region is extracted from its color image. Skin color is obtained from the histogram of Hue. The following Expressions are used:

As for the result, the input image of Fig.1 becomes Fig.2.
\[ H = \arctan \left( \frac{C_1}{C_2} \right) \]
\[ S = \sqrt{C_1^2 + C_2^2} \]
\[ V = -0.3R - 0.59G + 0.89B \]
\[ C_1 = R - Y = 0.7R - 0.59G - 0.11B \]
\[ C_2 = B - Y = 0.3R - 0.59G + 0.89B \]

Approximated circles of eyes are extracted as the irises by the circle Hough transform \[3\] first. Successively the approximated circles of nostrils are extracted almost in the same way by the positional guidance of the irises.

Rectangle regions in which eyes, nose, eyebrows and ears are involved respectively are extracted, prior to the detailed image processing for extracting the contour of the facial parts. Fig. 3 shows an example of these rectangle regions.

In order to make the system robust, the contour of the jaw is extremely important because the quality of the shape of the jaw is dominant in impression.

For eliminating the hair region from the input image, the color image is converted first to a gray image, and the smoothing and thresholding procedures are applied to the gray image. Fig. 4 is an example of the eliminated hair region.

Adaptive Sobel filters are applied to the rest region for extracting the edge image for the jaw recognition by means of the smoothing and thresholding procedures. Fig. 5 shows an example of the edge image for the contour extraction of the jaw.

### 3. Problem

As a result, the robustness of this system against the real world was clarified in the field test at EXPO2005. The primal reason for the failures was caused by the troubles in the detection of the jaw. This experimental fact requires us to introduce some Top-Down procedure to the system for coping with the lack of the quality in the Bottom-Up procedure.

The facial caricatures were visually evaluated and as a result, a lot of failures were found due to the failures in jaw detection. The reason for the cause of the failure is that the detected jaw region frequently contains the exceptional edges. To cope with this problem, we proposed two countermeasures. One is to improve the “bottom-up” image processing procedure, and another is to introduce the analytical curve fitting as the "top-down" procedure.

It is important to assume properly the analytical curves to fit. There are some the methods for selecting the proper analytical curve as shown in the next section. And the method for removing the exceptional point is also described afterwards.

The success rate of the likeness of the caricature exceeded 90% by introducing the fail-safe module. It was known that many of the failure cases with 10% were caused by the failure in the extraction of jaw contour. The success example of the caricature is supported by the properly extracted jaw contour shown in Fig. 5.

On the other hand, the result that the jaw area cannot be specified correctly as shown in Fig. 8 is obtained when the same colors as the skin color of the boundary are in the throat and the hair. Besides this, in the case when a man with the beard, the jaw extraction fails was seen.

### 4. Proposal of Shape Recognition Method

#### 4.1. Outline of the Proposed Top-Down Process

Since it is difficult to complete the processing only by the current bottom up method, we propose a top down procedure as follows:

The current bottom up method of image processing is still used, and then the extracted feature points are intensively investigated by using some analytical curves, and finally we draw a contour curve which is smooth but dependent to the respective feature points.

Many researches for the face contour extraction have been conducted by using a variety of image processing technologies. Though there is a well known method of “snakes” for shape recognition of the jaw contour extraction, “snakes” is expensive in both memory and processing time for the real time implementation. Generalized Hough transform is also promising to utilize for the extraction of the jaw contour, but the cost of GHT is much more expensive.

These are the reason why we must introduce an alternative method which is easy to use and robust in performance.

#### 4.2. Proposed Method, Preliminary Experiment and Preprocessing

The shape of jaw is of course different from one by one, and it is promising to assume there are three typical kinds of jaw in shape. One is an oval shape as shown in Fig. 7. The next is a square shape as shown in Fig. 8, and the last is the round shape as shown in
Fig. 9. Since the contour of the last can not be expressed by the single-valued function along with either horizontal or vertical axis, it is expected to introduce an alternative polar coordinate system to cope with this inconvenience as follows:

Let the maximum of $x_i$ be $x_{\text{max}}$ and the minimum of $x_i$ be $x_{\text{min}}$ ($i=0, 1, \ldots, m$). Let the value $a$ be the mean value of $x_{\text{max}}$ and $x_{\text{min}}$ (eq.(2)). Let $y_i$ at $x_{\text{min}}$ be $y_1$, and $y_i$ at $x_{\text{max}}$ be $y_2$. Let the value $b$ be the mean value of $y_1$ and $y_2$ (eq.(3)).

$$
a = \frac{(x_{\text{max}} + x_{\text{min}})}{2} \quad (2)$$  
$$
b = \frac{(y_1 + y_2)}{2} \quad (3)$$

The parameters $a$ from each feature points in $x$ coordinate and $b$ from each feature points in $y$ coordinate are used to define the value $P$ (eq.(4)).

$$
\begin{align*}
P_{y_i} &= x_i - a \\
P_{y_i} &= y_i - b
\end{align*} \quad (4)
$$

By this transformation, all feature points are shifted to the new axes, and the pattern of Fig.10 is moved neighboring to the new origin as shown in Fig.11.

The next step is to obtain each the maximum $x$ and $y$ in $P$ (eq.(5), eq.(6)).

$$
P_{x_{\text{max}}} = \max \left( \left| P_{x_i} \right| \right) \quad (5)$$

$$
P_{y_{\text{max}}} = \max \left( \left| P_{y_i} \right| \right) \quad (6)
$$

The results $P_{x_{\text{max}}}$ and $P_{y_{\text{max}}}$ are compared with each other by eq.(7), the bigger one is selected as $R$.

$$
R = \begin{cases} 
P_{x_{\text{max}}} & P_{x_{\text{max}}} \geq P_{y_{\text{max}}} \\
P_{y_{\text{max}}} & P_{x_{\text{max}}} < P_{y_{\text{max}}}
\end{cases} \quad (7)
$$

The range where all feature points are within from -1 to 1 is defined by using $R$. This regularization process is defined as follows:

$$
Q_{y_i} = P_{y_i} / R \quad (8)
$$

The reason why we introduced the regularization process is that the rectangular coordinate system used by the regularized $Q$ is transformed to the polar coordinates system. The following expressions are used for the conversion of the coordinate system.

$$
\rho_i = \sqrt{Q_{x_i}^2 + Q_{y_i}^2} \quad (10)
$$

$$
\theta_i = \arctan \left( \frac{Q_{y_i}}{Q_{x_i}} \right) \quad (11)
$$

The feature of jaw contour part is shown in Fig.12 when the polar coordinate system is applied. It is easy to analyze by single-value function by expressing the contour in polar coordinate system. When Fig.12 corresponds to Fig.7, Fig.8 corresponds to Fig.13. The analysis of the jaw contour expressed by the polar coordinate system can be executed definitely easier than those before transformation. In addition, the feature of the size of the jaw can be judged according to the ratio $R$.

Fig.14 is the flow chart of the entire system that was constructed by the above-mentioned procedures.
Here, \( j \) is the serial number from 0 to the size of the image.

Next, the jaw contour of the regularized size is transformed to those of the former size by eq.(13).

The parameters \( R \) requested by prior processing, and \( b \) are used.

\[
\begin{align*}
Ux_j &= Tx_j \cdot R + a \\
Uy_j &= Ty_j \cdot R + b
\end{align*}
\]  

(13)

As a result, it became possible to draw the jaw contour in former size.

4.4. Jaw Shape Recognition

Fig. 14 is a flow chart for the automatic classification of the jaw shape.

5. Proposal Deformation Method

The experiment was conducted by using a set of face outlines formatted in PICASSO format extracted from a part of the HOIP face data base. For example 30 female faces in 20\(^{th} \) were utilized. Linear least square method was used in the experiment. The degree \( n \) of \( AIC \) was estimated by analyzing all samples of which value distribute from minimum to maximum degrees by corresponding to the number of characteristics without fixed value, and by using the parameters of the curves.

The best degree was decided automatically. \( AIC \) is evaluated by eq.(14)

\[
AIC_m = m \cdot \log(S_m) + 2(m+1)
\]  

(14)

The degree of \( AIC \) that reached to the minimum value is adopted.

6. Experiment, Result and Consideration

Fig. 15 (a) - (c) are the results extracted from three type shapes. Fig. 16 (a) - (c) are the results of original method. Fig. 17 (a) - (c) are the results of the proposed method.

By analyzing in the orthogonalized coordinate system, the degree became 2 in the experiment regardless of the difference in jaw shape. As a result, it was known that it became easy and robust to analyze all kinds of jaw shapes. The minimum was 1st degree, the maximum was the 8th, and the average was the 2nd degree. The first was selected when the distance from a center point to each feature point was almost equal, because the inclination was small. When there exists a feature point in the distance from the center point to the top of the jaw, 2nd degree was selected. Degrees that are larger than secondarily were selected for expressing the shape with a lot of features like Fig.8.

7. New Deformation Method and Pretest

We developed a new deformation method based on the pattern space defined on the knots of B-Spline.

7.1. Introduction of new Deformation Method

Fig.1 shows the flow chart of the entire system that was constructed by the above-mentioned procedures. “Normalized” in Fig.18 is the same as those in 4.2.

7.2. B-Spline Curve

The following expression is B-spline formulation:

\[
f(t) = \sum_{j=-k+1}^{n-1} c_j B_{j,k}(t)
\]  

(15)

, where the coefficients \( c \) is the curve parameter, \( k \) is the number of order (number of degree=\( k-1 \)), \( n+1 \) is the number of knots, \( B \) is the basis function of Spline. \( B \) can be uniquely decided, but \( c \) can not be uniquely decided because \( c \) is decided by some methods. \( c \) is decided by using least squares method in this study.
The knots are set up uniformly.

Fig. 18 System flow chart

7.3. Evaluation of Fitted Curve

Fitting is evaluated by two procedures: One is AIC (eq.17), and another is general error analysis (eq.18).

\[
\delta^2 = \sum_{i=1}^{p} (y_i - f(x_i))^2 \tag{16}
\]

\[
AIC = p \cdot \log \delta^2 + 2(n + k - 1) \tag{17}
\]

\[
\sigma^2 = \frac{\sum_{i=1}^{p} \delta^2}{(p - n - k)} \tag{18}
\]

where the parameter \( p \) is the number of facial parts points. Obtained curve was fit, with an estimated error of around \( \pm 0.3 \% \) or less.

7.4. Deformation Method

Deformation in the former PICASSO-2 uses the feature points and the in-betweening method of eq.(19), but the new deformation uses B-spline curve parameter and knot. The following two expressions are the new expression of the deformation:

\[
Q_j = P_j + b(P_j - S_j) \tag{19}
\]

\[
C_j = c_j + \alpha(c_j - \beta_j) \tag{20}
\]

\[
T_j = t_j + \alpha(t_j - \gamma_j) \tag{21}
\]

where the parameter \( \alpha \) is the ratio of exaggeration, \( \beta \) and \( \gamma \) are the average face of curve parameters and knots. The number of parameters is influenced by the number of knots and the order.

7.5. Pretest and Consideration

We experimented on the deformation under the following three conditions. The value of \( \alpha \) is fixed to be 1.0, the number of knots is 5, and degree is 3. Polar (a) shows the result of analysis in polar coordinates. Original (b) shows the image of translated result of analysis in polar coordinate to Cartesian coordinate. Deformation (c) shows the image of exaggerated curve.

Table 1 Experimental Results

<table>
<thead>
<tr>
<th></th>
<th>Oval</th>
<th>Square</th>
<th>Round</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_3 )</td>
<td>0.9983</td>
<td>0.9953</td>
<td>0.9977</td>
</tr>
<tr>
<td>( c_2 )</td>
<td>0.9920</td>
<td>1.0232</td>
<td>0.9894</td>
</tr>
<tr>
<td>( c_1 )</td>
<td>1.0081</td>
<td>1.0702</td>
<td>0.9455</td>
</tr>
<tr>
<td>( c_0 )</td>
<td>1.2885</td>
<td>1.0356</td>
<td>1.0583</td>
</tr>
<tr>
<td>( c_1 )</td>
<td>0.8811</td>
<td>1.0368</td>
<td>0.9515</td>
</tr>
<tr>
<td>( c_2 )</td>
<td>0.9982</td>
<td>1.0451</td>
<td>1.0286</td>
</tr>
<tr>
<td>( c_3 )</td>
<td>0.9878</td>
<td>0.9901</td>
<td>0.9900</td>
</tr>
<tr>
<td>error</td>
<td>3.29e-3</td>
<td>4.18e-3</td>
<td>2.90e-3</td>
</tr>
</tbody>
</table>

Our proposed method was successful to exaggerate each different jaw shape feasibly. The part that had to be exaggerated was successfully exaggerated. The size of the original x-y data was 42, on the other hand, data size of the proposed method is the value of \( c \)’s and knots. The value of \( c \) is 7, and the number of knots data is 5, therefore the total data size became 12 in the proposed method. The number of jaw data was able to be reduced to those below the half of the previous.

8. Conclusion

It was known that the accurate and good shape of the jaw was able to be recognized by the proposed method.

Moreover, by conducting the further experiments with the more face data, more robust and smart
processing could be established in the near future. Simultaneously it would be promising to put a new parametric space used in the proposed method be a new deformation space of the facial caricaturing system.

We succeeded in the recognition of the shape preliminarily, and in addition, the improvement of the accuracy of shape recognition is enforced by the top down method proposed in the paper.

The knots are set up in the place where the area was taken at equal intervals in this paper. But the exaggerated knots should be closer to both the top and the valley of the feature points in the polar coordinates. It is expected to design some auxiliary method for automatically deciding knots by using the feature points.

Reference


