13-20

Facial Caricaturing Robot COOPER Exhibited at EXPO2005 and

Its Improvements

Naoya Tokuda,
Chukyo University
101 Tokodachi, Kaizu-cho,
Toyota, 470-0393, JAPAN
tokuda@koshi-lab.sist.chuyko-u.ac.jp

Takayuki Hoshino,
Chukyo University
101 Tokodachi, Kaizu-cho,
Toyota, 470-0393, Japan
hoshino@koshi-lab.sist.
chukyo-u.ac.jp

Takashi Watanabe,
Chukyo University
101 Tokodachi, Kaizu-cho,
Toyota, 470-0393, Japan
nabe@koshi-lab.sist.
chukyo-u.ac,jp

Takuma Funahashi
Chukyo University
101, Tokodachi, Kaizu-cho,
Toyota, 470-0393, JAPAN
takuma@koshi-lab.sist.chukyo-u.ac.jp

Takayuki Fujiwara,
Chukyo University
101 Tokodachi, Kaizu-cho,
Toyota, 470-0393, Japan
tfuji@sist.chukyo-u.ac.jp

Hiroyasu Koshimizu,
Chukyo University
101 Tokodachi, Kaizu-cho,
Toyota, 470-0393, Japan
hiroyasu@sist.chuyko-u.ac.jp

Abstract

We developed the facial caricaturing robot "COOPER" that was exhibited at the Prototype Robot Exhibition of EXPO 2005, Aichi Japan during 11 days from Jun.9 to Jun.19. COOPER watches the face of a person seated at the chair, obtains facial images, and analyzes the images to extract 251 feature points to generate his facial line drawings with deformation. It is noted that the caricature was drawn on the specialized "Shrimp rice cracker" in 4 minutes. To do this we customized the original system PICASSO by coping with the illumination circumstances in EXPO pavilion. This paper illustrates the outline of the COOPER and the details of the image processing in it. And we discusses on the prospects of the future subjects based on more than 352 facial caricatures obtained at EXPO2005.

1. Introduction

Facial caricaturing robot "COOPER" was developed, and exhibited at the Aichi Expo EXPO2005 (Prototype Robot Exhibition) [1] This robot acquires the face image of the visitor, extracts the facial features automatically, deforms them as his caricature [2-4], describes it on the shrimp rice cracker, and finally presents the facial caricature to the visitor. Fig 1 shows an external view of COOPER.

COOPER outputs the caricature shown in Fig 2 by scorching the surface of the shrimp rice cracker with the semiconductor laser pen for representing it as the connected line drawings.

The image processing system for COOPER takes a couple of facial images with CCD camera, extracts the facial features from the image, and generates the caricature. The facial image processing system is designed for recognizing irises, nostrils from the facial images first, and defines the respective regions involving eyes, nose, mouth, and ears guided by the positional relation among irises and nostrils. The hair region and skin region surrounding these regions of the face are also defined, and the shape features of the each facial parts

are detected. In the final result, this system generates the caricature comprising by 251 feature points that are defined originally as the cooper-PICASSO format. This system evaluates simultaneously the quality of the intermediate results including caricature by using "fail-safe modules". Thus, we have succeeded in designing the robust performance even in the unconditioned circumstance such as EXPO site.

For obtaining the images, the visitors are asked to sit at a chair with the blue backrest, as shown in Fig. 3. This system extracts the facial features from input images, and generates the caricature in the same way as the deformation method of PICASSO system [5]. This system uses a couple of facial images captured by CCD camera with 1 fps. If this system fails to process the first image, the same procedure is applied again to the second image. Then, COOPER generates the caricature, converts it into the robot control language and finally controls the laser pen. These industrial-use robot arms were assembled together with the head with cameras, body and legs. Since COOPER has both tilting and rotating mechanisms of the head, we could design the motion of the robot head to be performable like a human caricaturist. We also designed the motion of arms to reduce its loading weight as small as possible and then to realize smoother movement of the arms. We took also the safety conditions into consideration in order to cope with some abnormal operations of the laser pen.

We took some auxiliary information of the respective visitor by using touch-sensitive panel. The contents of this information consist of a kind of facial expression (normal, sad and smiling), sex, age(less than 10, 10's, 20's, 30's, 40's, 50's, more than 60), and his authorization for the usage of his face data for further researches on facial caricaturing.

Fig.1 Exterior view
of COOPER

Fig. 2 Example of the facial
caricature
on the shrimp rice cracker

Fig. 3 Blue backrest chair and touch-sensitive panel

2. DETAILS OF IMAGE PROCESSING SYSTEM

2.1. Detection of skin region

As the preprocessing for extraction of facial features, this system detects skin color region from RGB image. In the preprocessing, blue region of the background is eliminated from the input image. And therefore the skin color region is detected from the input image as shown in Fig. 4 based on the hue discrimination as shown in Fig. 5. This skin color region is defined and used for the successive image processing.

2.2. Extraction of irises and nostrils

In this system, irises are first extracted by using Hough transform [6] for leading other hierarchical processing modules. Secondly nostrils are extracted in the same way of irises at the nose region. The results of irises and nostrils are shown in Fig. 6.

2.3. Facial parts detection

The regions of eyes, nose, mouth and ears are defined by using the information on irises and nostrils. As defined in each facial parts region, outlines of eyes, nose, mouth and ears are detected from the input gray image by using smoothing, contrast enhancement, thresholding and thinning procedures, as shown in Fig. 6.

2.4. Contour detection

We basically designed that the caricature of COOPER is represented with a set of line drawings. This means that the face of line drawings is less informative than the original image in physical meaning, but that the face of line drawings is more effective than the face image in impression. In this sense, the shape feature of the face contour, hair and jaw is more dominant than the gray image. Moreover the fact that the face of line drawings is easier to realize the correspondence among faces than the face images is one of the technical advantages.

The outline of hair is detected from the binary image by the method of smoothing, contrast improvement and thresholding, as shown in Fig. 7.

The outline of jaw is detected from R image of RGB color image by using Sobel operator and thresholding, as shown in Fig. 8.

2.5. Fail-safe principle and its implementation

At the same time of the extraction of facial parts, this system evaluates how feasible the result is, and modifies the result, if necessary, according to the statistical standard for the positional relationship among facial parts. This fail-safe system evaluates the result by the estimation function preliminarily prepared [7] which was defined by the difference between the result of the input face and mean face. If this system rejects the result, it is replaced by the corresponding facial parts of the mean face and fitted it as the facial parts.

2.6. Caricature generation

COOPER system inherits the basic mechanism of deformation from the original system PICASSO. Facial caricaturing system PICASSO which extracts some facial individuality features from the input face and deforms these features to generate a caricature. The facial caricature Q is generated by comparing the input face P with the mean face S, which is defined by averaging input faces as shown in Fig. 9 and Eq. (1). This system introduces the exaggeration rate b for adjusting the deformation of the caricature to each visitor
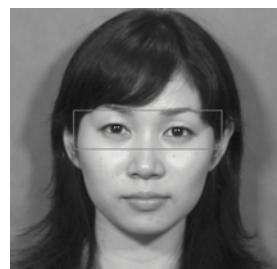
$$Q = P + b(P - S) \qquad (1)$$

Fig. 4 Input face

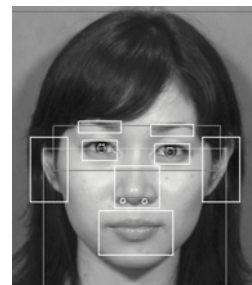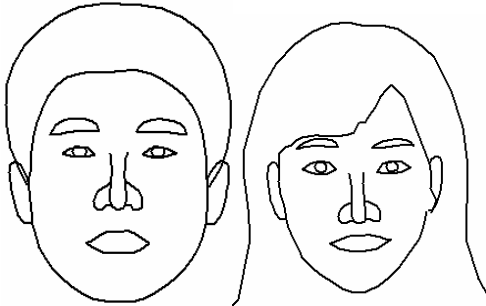Fig. 5 Skin color region

Fig. 6 Example of facial features extraction

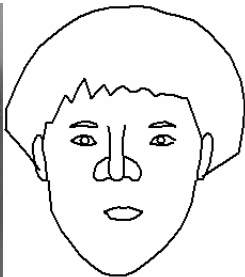Fig. 7 Hair region

Fig. 8 Pre-processing of jaw extraction

(a) Mean face (b) Caricature 1

(c) Input face (d) Caricature 2

Fig. 9 Example of caricature generation

3. EXPERIMENTS

During 11 days in the Aichi Expo2005, COOPER manufactured 352 facial caricatures almost successfully and presented them to the visitors. Afterwards, the total evaluation of these caricatures by the intensive observations was executed, and the result shown in Table 1 was obtained.

As the result of the above detailed evaluations, the lack of quality in jaw parts detection caused a lot of failure examples in facial caricature generation. Some intensive investigation lets us know that a person with unclear

boundary between mandible and neck, with a thick beard and/or with the dark shadow on the cheek are likely to be failed in caricature generation.

Inspired by these considerations, it is expected to enforce this system from both bottom-up and especially top-down procedures. As the top-down procedure, we introduced the curve fitting methods by means of several analytical curves such as polynomial and B-Spline functions. Fig. 10 shows a couple of examples of the preliminary experiments. As known from these results, it would be promising to reduce the side effect of the noise, and further more to make the facial caricaturing robot COOPER be automated.
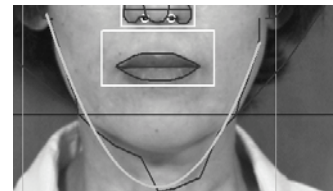

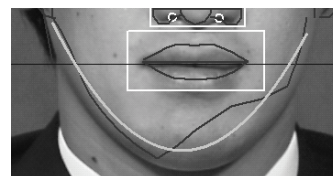(a) Least mean square approximation
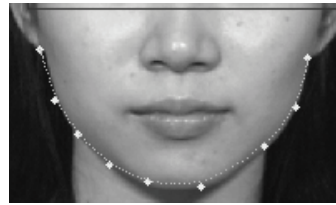
(b) B-Spline

Fig. 10 fitted contour line

In Fig. 11 and Fig. 12, we show some experimental examples of the curve fitting by analytic functions to the jaw contour which was detected as the typical one of the failures. This system is likely to fail to extract a natural jaw contour as shown in Fig. 11 (a) and/or to succeed as shown in Fig. 11 (b) when least mean square approximation is used. On the other hand, this system failed both Fig. 12 (a) and Fig. 12 (b), when 'B-Spline' fitting is used.
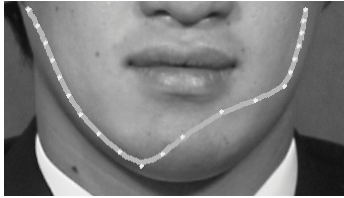

(a) case1

(b) case 2

Fig. 11 Least mean square approximation

(a) case1

(b) case2

Fig. 12 B-Spline

4. SUMMARY AND FUTURE TASKS

In this paper, we investigated intensively the performance of the facial caricaturing robot COOPER exhibited at EXPO2005, and based on the investigations, in order to enforce the bottom-up image processing we introduced a top down curve fitting architecture to extract the natural contour of the jaw and a few preliminary experiments were executed. Then we could experimentally show the possibility to introduce a curve fitting method as one of the top-down procedures. As the future tasks, it is expected to remove the random noise prior to the proposed top down procedure, and it is also expected to make the gradient operators be more adaptive to the detailed shape of the facial parts. Moreover, this system must be more robust to the unexpected changes of the illumination and to the spontaneous movement of the head of a person sitting in front of the camera.

5. REFERENCES

[1] Robot Project: Prototype Robot Exhibition: http://www.expo2005.or.jp/jp/C0/C3/C3.8/C3.8.2/C3.8.2.6

[2] Caricaturing robot, EXPO'85, Matsushita pavilion:http://www.mid.co.jp/mid/shirahama/nigaoe.html

[3] E. Takigawa, H. Kishiba and M. Kawade: @faceAutomatic Gender and Age Estimation with Face, Proc. of the Conf. on Japanese Academy of Facial Studies 2002, p171 (2002)

[4] K. Teranishi, N. Kotani and M. Shinya: Chara-Face:Aportrait Caricaturing System, Proc. of General Conf. on IEICE, A-14-5 (2000)

[5] H. Koshimizu: Computer Facial Caricaturing, Trans. The Institute of Image Information and Television Engineers, Vol.51, No.8, pp. 1140-1146 (1997.8).

[6] T.Funahashi, T.Fujiwara, M.Tominaga, and H.Koshimizu: Hierarchical Face and Facial Parts Tracking and Some Applications, Prod. of 7th International Conference on Quality Control by Artificial Vision, pp. 305-310, Japan (2005)

[7] T. Fujiwara, R. Ushiki, M. Taga and H. Koshimizu: A Method of Facial Attribute Classification based on Statistic Analysis of the Relationship among Facial Parts, Journal of Japanese Accademy of Facial Studies, Vol.2 No.1, pp. 25-38 (2002)

Table 1 Experimental result

A: The number of generated caricatures,
B: The number of successful generation of caricature,
C: The number of unsuccessful extraction of irises and nostrils and successful extraction from the 2nd frame,
D: The number of the fitting by mean face,
E: The number of cases that were recovered by FFS,
F: The number of failures below the constant standard in the visual inspection,
G: The number of cases that failed fatally,
H: The ratio of the successful cases to the total number of the caricatures,
I: The ratio of the successful cases to the total number of the caricatures by using FSS,

	6/9	6/10	6/11	6/12	6/13	6/14	6/15	6/16	6/17	6/18	6/19	total
A	46	33	30	29	29	31	32	33	35	34	20	352
B	34	25	22	21	17	19	26	30	17	28	14	253
C	4	0	2	0	2	1	1	1	1	0	1	13
D	4	8	4	8	8	8	4	0	14	3	3	64
E	8	8	6	8	10	9	5	1	15	3	4	77
F	3	0	1	0	2	3	1	2	2	3	2	19
G	1	0	1	0	0	0	0	0	1	0	0	3
H	73.91	75.76	73.33	72.41	58.62	61.29	81.25	90.91	48.57	82.35	70.00	71.88
I	91.30	100.00	93.33	100.00	93.10	90.32	96.88	93.94	91.43	91.18	90.00	93.75