International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 22, No. 8 (2008) 1647–1668

A NEW METHOD FOR FACIAL FEATURES QUANTIFICATION OF CARICATURE BASED ON SELF-REFERENCE MODEL

FEI NI, ZHUANG FU*, QIXIN CAO and YANZHENG ZHAO

State Key Laboratory of Vibration
Shock & Noise, Shanghai JiaoTong University
ShangHai 200040, P. R. China
**zhfu@sjtu.edu.cn*

Some facial features that differ from an ordinary face should be identified by a computer when generating a facial caricature. These distinctive facial features are called self-features. Compared with traditional Mean Face Model (MFM) that is unable to quantify these self-features well, a Self-Reference Model (SRM) is presented in this paper. Firstly, based on the physiology structure of a front face, a self-reference is found, and this reference is used to measure the self-features. According to the self-reference, some standard facial parameters are worked out by collecting statistic data of many facial images. Then, in an input face image, by evaluating some differences between the input face and the standard facial parameters, the self-features are properly estimated and quantified. Finally, by analyzing some caricatures produced by caricaturists, the SRM can prove the validity of the proposed Algorithm.

Keywords: Facial features recognition; caricaturing method; self-reference model.

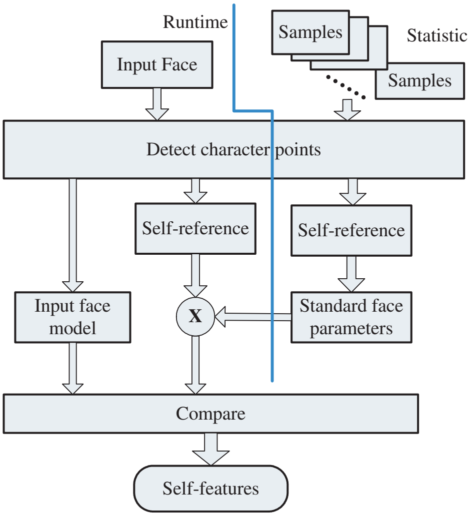
1. Introduction

There are almost 6.5 billion people on the planet. Each face is made up of several main parts: forehead, eyebrows, eyes, nose, mouth, jaw and cheeks. Their relative positions are basically fixed, but the patterns of these features are so complex that no people have the same face around the world even between a pair of twins. So, it is obvious that the self-features can be representative, and these features should be distinct enough to distinguish him or her from others. Therefore, as for face searching, identification and automatic caricature generated by computers, to extract some self-features should be very helpful.

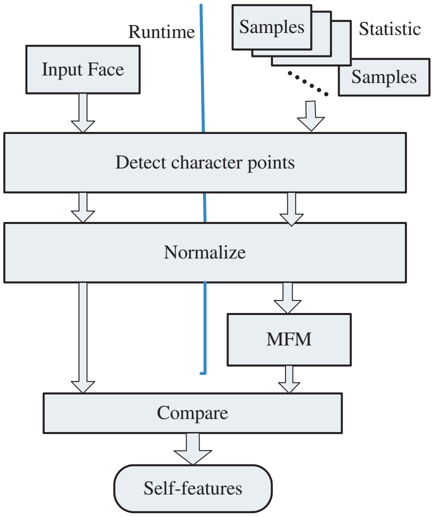
Caricature is an art processing that conveys humour to people via drawing an exaggerated portrait. A successful caricature should display an object's self-features exactly. For instance, as a caricaturist, when he draws a person with a short jaw, the jaw should be much shorter than the actual. In this way, the short jaw can be regarded as a self-feature. Generally, to exaggerate self-features is the most important method for caricaturing.[15] It is a common point that a caricaturist always compares an object face with a standard face and draw the caricature by exaggerating some distinctive self-features.[18] Unfortunately, even a professional caricaturist would not be able to quantify all the exaggerations that he or she wants to introduce.[18] Therefore, there are many difficulties in quantifying self-features. Obviously, these difficulties lead to problems of "how to define these positions, sizes, and forms of the self-features?", "How to estimate and quantify them?" and "How can a jaw be regarded as a short jaw or a long jaw after the length of the jaw is obtained?" In fact, some approaches have been proposed (see Refs. 1, 3, 6, 10, 14, 16, 18 and 19), and these methods have discussed how to generate a caricature by computer. In most of these references, Exaggerating the Difference From the Mean (EDFM) has been widely accepted as the driving factor behind the generation of caricature.[3] By collecting some feature points from numerous face-samples, we can build up a Mean Face Model[6] (MFM). The exaggerations are determined by some differences between an input face and the MFM. In Ref. 6, an exaggeration rate controls the exaggerating degree of these self-features. Moreover, Shet *et al.*[18] stated that the Cascade Correlation Neural Network (CCNN) can be used for capturing the drawing styles of an artist, and generating the realistic automatic caricature. But they do not discuss how to find and estimate these self-features in detail. In Ref. 9, some anthropometrics features are employed to assist in detecting some feature points of the face. Authors have placed emphasis on the problem of seeking these feature points but the self-features. In Ref. 12, a method based on examples is proposed, and many training samples will be offered to the computer in order to learn some skills and styles of some painters. So, these proposed approaches could be summarized as follows:

1. These presented approaches involve discussion on how to draw caricatures (see Refs. 1, 3, 6, 9, 10, 12, 14–16, 18 and 23), and how to extract the drawing styles (see Refs. 1, 3, 6, 9, 10, 12, 14, 16, 18, 19 and 23) but few papers focus on the driving factor of caricature — finding and quantifying the self-features.
2. Though some literature[16,19] have discussed how to identify self- features, the self-features are derived from the MFM. The MFM is to spot some feature points from numerous samples and calculate their averages. These averages are regarded as a standard face model (or the standard face parameters). An input face should be compared with the MFM. However, the MFM has some natural demerits (see Sec. 2 for details).

According to the above comments, this paper presents a novel method for finding and quantifying self-features based on the SRM. First, we will introduce the system overview in Sec. 2. Second, some anthropometric features of head are analyzed and a self-reference is extracted to describe the self-features. Subsequently, statistic data are used to figure out the standard face parameters (standard face model) in Sec. 3. Then, in Sec. 4, some experiments are discussed to prove that the SRM can reflect drawing skills of artists properly. Finally, conclusion and discussions are drawn at the end of Sec. 4.



(a)



(b)

Fig. 1. The frames of the SRM and the MFM.

2. System Overview

Differently from the MFM, the paper presents a Self-Reference Model to quantify the self-features of a given face. The frames of the SRM and MFM are shown in Figs. 1(a) and 1(b). Obviously, the MFM needs two stages: runtime and statistic (or training). In the MFM, an input face will be verified by an invariable reference: the mean face. So, the MFM has some demerits:

1. The MFM requires normalizing. Generally, under different poses and scales, to normalize a face must be very difficult. Moreover, to normalize an image would result in some errors inevitably.
2. The MFM describes the whole features only, so these corresponding relations between the MFM and an input image are not legible. In Refs. 1, 3, 6, 10, 14, 16, 18 and 19, the processing of quantification is not mentioned, and in different references, different self-features have been selected. So far, there is not a set of uniform self-features.
3. According to some drawing skills, the portrait composition is the most important one. The composition involves arrangement of some elements which make up a portrait, such as eyes, eyebrows, nose, etc.[15] But the MFM is hard to adopt and quantify these relations.

The SRM also involves runtime and statistic. In the statistic stage, some standard facial parameters can be worked out. In the runtime stage, a self-reference is revealed to measure the facial features and generate the standard face model. Supported by the self-reference, in an input image, the input face model is generated, and it will be compared with the standard facial parameters. So, the SRM can make a dynamic reference instead of an invariable reference used by the MFM. Therefore, several advantages of the SRM are listed as follows:

1. Different input faces are verified by different self-reference, and the input face is estimated by the self-reference of the input face.
2. Normalizing is not required. In the MFM, the process that transforms the input faces to the normal size is complex and it would lead to heavy costs of calculation. So the SRM eliminates some errors which emerge out during the normalizing.
3. The SRM can define the positions and sizes of the facial features evidently. Especially, the SRM can express intuitionistic and quantificating composition legibly.

3. Building up the Self-Reference Model

3.1. *The principle of composition*

According to Polyclitos rule, a head is 3.5 times the height of a forehead. To draw four horizontal lines, the head should be made up of 3.5 units:

(a) The top head (the region of hair line).
(b) The line between the eyebrows and the top of the ears.
(c) The line at the bottom of the nose.
(d) The line at the bottom of the jaw.

If we use the same unit to measure the width of the head, the head can be divided into two and a half units in the horizontal direction.

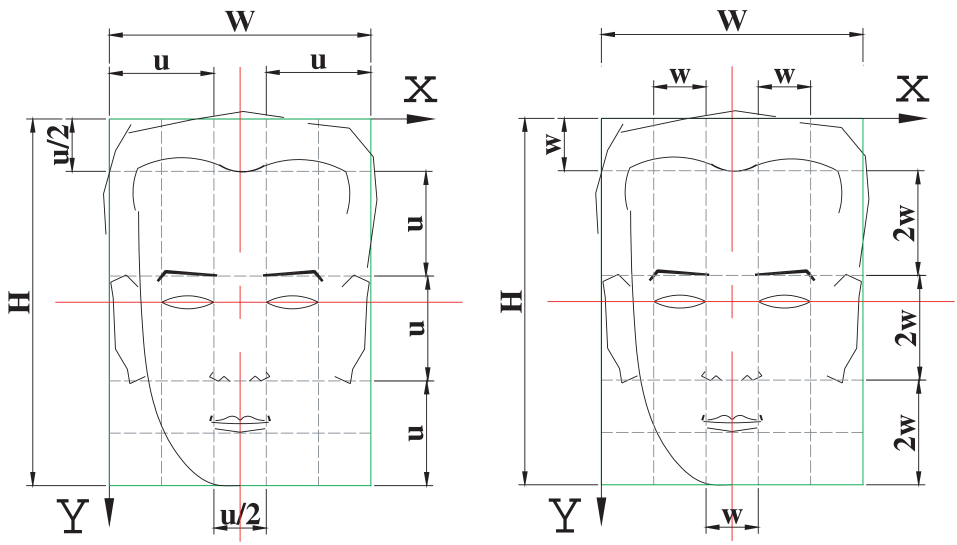
So, to assume the height of a forehead is u, the width and height of the head: W, H can be calculated by Eq. (1):

$$H = 3.5u, \quad W = 2.5u. \tag{1}$$

According to u, H, W, a whole composition can be arranged as one shown in Fig. 2(a). It is apparent that, as far as Ployclitos Drawing Rule is concerned, once u has been fetched, the whole composition can be completed. Thereby u is regarded as a self-reference. Unfortunately, because the forehead is covered by hairs, it is difficult to design an algorithm for discerning the high of the forehead properly. Therefore, for automatic caricature, we need a self-reference that can be recognized by certain algorithm exactly. As we all know, in the field of facial recognition, eye location problem has attracted significant interests over the past decades. Many efforts have been exerted to detect eyes.[2,5,11] So, we can assume that the width of the eye, which can be easily identified by an algorithm, should be regarded as a self-reference. In fact, in some literatures of drawing skills, the width of an eye is an auxiliary reference.[4] Now, the eye width is indicated by w. H and W are determined by:

$$H = 7w, \quad W = 5w. \tag{2}$$

Based on w, H, W, the composition is shown in Fig. 2(b).



(a) (b)

Fig. 2. The facial composition.

Obviously, the relation between w and u is simple:

$$u = 2w. \tag{3}$$

So, w is called as the self-reference which can determine the facial composition and size [see Fig. 2(b)]. In the following discussion, all self-features are founded on the self-reference, and some experiments prove that the self-reference is valid in Sec. 4.

3.2. *To deduce facial parameters using the self-reference*

In this paper, the self-features of a face are divided into three types: composition feature, size feature and form feature. Table 1 lists these self-features. The composition feature, size feature and form feature describe positions, sizes and forms of these self-features, respectively.

Both the composition and size features are closely connected with the self-reference although it is hard to picture their relations in our minds visually. Obviously, even the width of a nose has been detected, you cannot tell that the nose is wide or narrow because there is no reference. This is the opportune reason why we need the SRM or MFM. Compared with the composition and size features, the form feature is easy to obtain, and it is independent of the self-reference. In fact, once some feature points are picked up, the form feature can be determined immediately. For example, even without the self-reference, the self-feature — the external canthus is higher or lower than the internal canthus can be estimated directly. So our system focuses on obtaining the composition and size features in an input face. The

Table 1. The facial self-features.

Feature Name	Symbol	Comment	Feature Type
Face form	c_1, c_2, c_3	See Fig. 3	Composition Features
Hair Line Position	N/A	The distance between hair line and eye line	
Eyebrow Line Position	c_4	The distance between eyebrow line and eye line	
Eye Line Position	N/A	It is a reference position	
Nose Line Position	c_5	The distance between nose line and eye line	
Low Lip Line Position	c_6	The distance between low lip line and eye line	
Jaw Line Position	c_7	The distance between Jaw line and eye line	
Height of Eye	s_1	See Fig. 3	Size Features
Width of Nose	s_2	See Fig. 3	
Width of Mouth	s_3	See Fig. 3	
Thickness of Low Lip	s_4	See Fig. 3	
Thickness of Up Lip	s_5	See Fig. 3	
Form of Nose	f_1	No Concern	Form Features
Form of Eye	f_2		
Form of Eyebrow	f_3		
Form of Mouth	f_4		

composition and size features are shown in Fig. 3(a). The face is divided into 30 small regions marked by a, b, c, d, e and 1, 2, 3, 4, 5, 6 [see Fig. 3(b)]. In an input image, if the four canthuses are discovered, other parts of the face can be anchored on these regions roughly. Table 2 lists these regions.

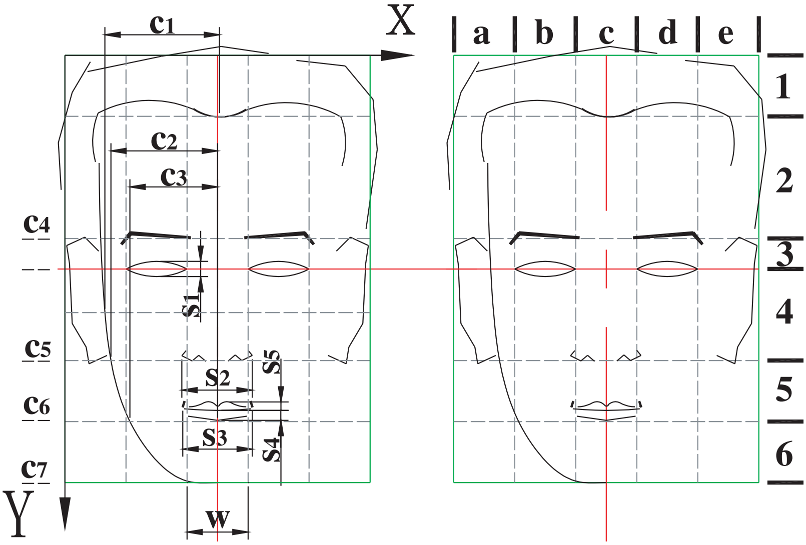
Now, the standard facial parameters k can be calculated by Eq. (4):

$$k = \frac{p}{w} \tag{4}$$

where $p = [c_i, s_j]$ $i = 1, \ldots, 7$ $j = 1, \ldots, 5$. p is a vector including the composition and size features. k_i is one of the standard facial parameters, which correspond to p_i. In fact, k_i can be regarded as an expectation value. Thereby, in order to reckon k_i, we need to collect the statistic of p_i. So far, the standard facial parameter has been connected to the self-reference: w. The next step will be how to estimate the standard face parameters.

3.3. *To estimate the standard facial parameters*

For obtaining p, it is necessary to detect some feature points (see Fig. 4). These feature points are indicated by a series of successive numbers in Fig. 4, and x, y coordinates of these points are denoted by (x_i, y_i) $i = 1, \ldots, 14$. The reason why we select these feature points is that they are related to the 12 self-features mentioned in Table 1. For instance, (x_1, y_1) indicates the position of the eyebrow: c_4; and (x_6, y_6) stands for c_1.



(a) (b)

Fig. 3. The facial composition and size features.

Table 2. The regional positions of the facial features.

Facial Features	Regions
Left eye	3b, 4b
Right eye	3d, 4d
Left eyebrow	2b, 3b
Right eyebrow	2d, 3d
Nose	4b, 4c, 4d, 5b, 5c, 5d
Mouth	5b, 5c, 5d, 6b, 6c, 6d
Left zygomatic	4a
Right zygomatic	4e
Mandible (symmetry)	5a, 5b, 5c

References 8, 17, 20 and 21 have proposed some valid and reliable methods to spot these feature points. Specially, in Ref. 13, Active Shape Model (ASM) is very appropriate to detect these points. Once these feature points have been detected, we can estimate p by:

$$w = x_5 - x_4, \quad ry = \frac{(y_3 + y_4)}{2}, \quad rx = \frac{(x_3 + x_4)}{2} \tag{5}$$

in which, ry, ry are x, y coordinates of the middle point of the line between two internal canthuses respectively. Then:

$$p = Tr \tag{6}$$

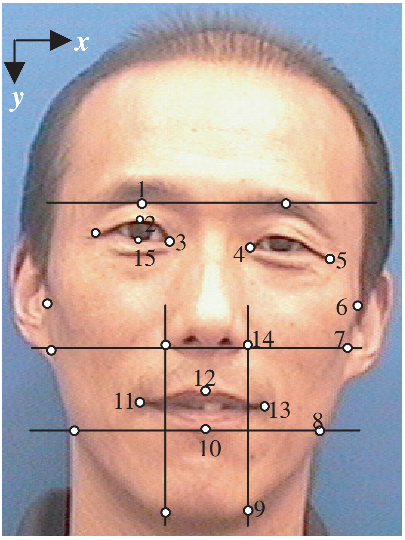


Fig. 4. The facial feature points.

where $r = [1 \;\; rx \;\; ry \;\; -1]^T$, T is a transform matrix, in which these coordinates of the feature points are filled like this form:

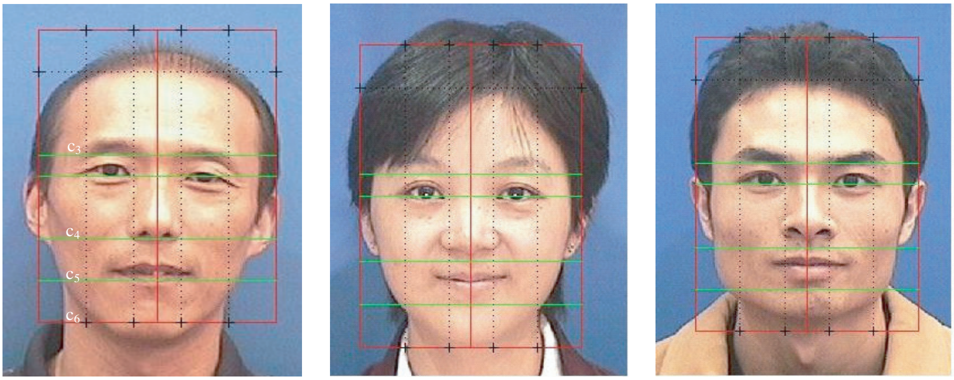
$$T = \begin{bmatrix} y_6 & y_7 & y_8 & y_1 & y_{14} & y_{10} & y_9 & y_{15} & x_{14} & x_{13} & y_{12} & y_{10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & y_2 & 0 & 0 & y_{13} & y_{13} \end{bmatrix}.$$

The form of T is decided by these feature points in Fig. 4. T has 12 columns, and each column is corresponding to a self-feature described in Table. 1. Additionally, four elements in each column imply the coordinates of these concerned points. For example, the first column indicates c_1, so these elements in the first column involve y_6 and -1 (the third element will be multiplied by ry, and ry is correlative with c_1).

Now, we give an example of calculating parameter k_3. k_3 is one parameter of the standard face model, which reflects the facial form: c_3. Firstly, 102 facial images are involved into the statistic. In each facial image, we can calculate its k_{s3} $s = 1 \cdots 102$ by Eqs. (4)–(6), respectively. Thereafter, k_{s3} of all samples are collected as a set of statistic data. Figure 5 shows some samples with different k_{s3}. In Fig. 5, some frames are drawn based on a hypothesis:

$$k_4 = 0.5 \quad k_5 = 1.5 \quad k_6 = 2.5 \quad k_7 = 3.5.$$

Obviously, a wider jaw has bigger k_{s3}, in other words k_{s3} can account for the width of a jaw. But, a more important problem is how to know that the width of



(a) $k_3 = 1.66$ (b) $k_3 = 1.6$ (c) $k_3 = 2.02$

Fig. 5. Some samples with different k_3.

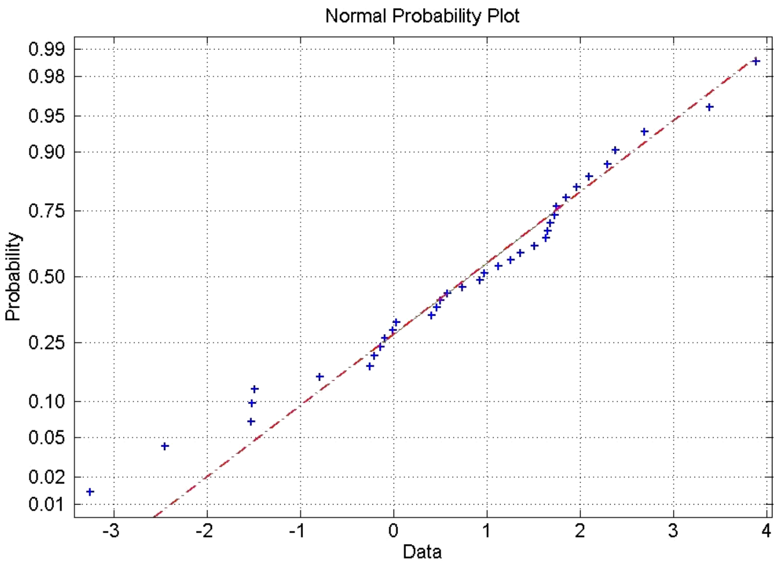


Fig. 6. The normal probability of k_{s3}.

the jaw is big or small. We consider that k_{s3} are a set of biological feature data. Therefore, to assert that k_{si} submits to a Gauss distribution should be reasonable. Now, we need to verify whether these samples come from a Gauss distribution or not. Figure 6 displays a normal probability plot of the data in k_{s3}. Superimposed data on the plot is a line joining the first and third quartiles. The line is extrapolated out to the ends of the samples to help evaluate the linearity of the data. If the data does come from a normal distribution, the plot will seem linear. In Fig. 6, the plot is linear, so we can ensure that these samples are from a Gauss distribution.

More formally, K–S test can be employed to estimate a Gauss distribution. Equations (7) and (8) give the equations for calculating skewness: S and kurtosis: K:

$$S = \frac{E(x-\mu)^3}{\sigma^3} \tag{7}$$

$$K = \frac{E(x-\mu)^4}{\sigma^4} \tag{8}$$

where μ and σ^2 are the expectation and variance of these samples respectively. So, k_{s3} is tested by Eqs. (7) and (8), and the results are

$$S_3 = 0.031, \quad K_3 = 2.896.$$

Under the confidence: $p = 0.9$, the confidence interval of S_{t3} and K_{t3} are $[-0.3, 0.3]$ and $[2.89, 3.06]$, respectively. And:

$$S_3 \in S_{t3} \quad K_3 \in K_{t3}.$$

So, we believe that the samples k_{s3} do come from a Gaussian distribution set. Now, by employing minimum-variance unbiased estimator, μ and σ of k_{s3} are estimated and the results are:

$$\mu_3 = 1.7408 \quad \sigma_3 = 0.1497.$$

Therefore,

$$k_{s3} \sim N(1.7408, 0.0224).$$

The confidence interval of μ_3 is: [1.6728, 1.8089].

The same goes for other k_{si}, and the same conclusions have been drawn. Table 3 lists these Gaussian distribution parameters of k_{si}. It can be found that the variance of c_1 is smaller than the variances of c_2 and c_3. From this information, we can draw a conclusion that c_2 and c_3 have more diversity, and the phenomena is consistent with a fact that a jaw takes more responsibility for different facial form than a zygoma does. More typically, c_7 that indicates the length of a jaw has more distinct diversity. So, the length of a jaw is also an important self-feature. The standard facial parameters k are the expectation of k_{si}:

$$k_i = \mu_i.$$

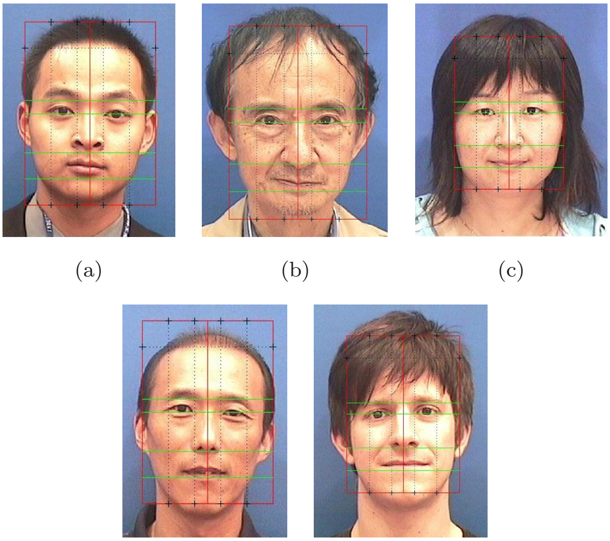
3.4. *To estimate self-features of an input image*

To identify the self-features of an input image is to calculate its k_{si}, and compare it with the standard facial parameters k. Equation (9) defines some feature scales o_i as

$$\begin{aligned} o_i &= \frac{k_{si}}{k_i} \quad i = 1, \ldots, 12 \\ o_i &= \frac{k_{si} - 3\sigma_i}{k_i} \quad \text{if } o_i < \frac{k_{si} - 3\sigma_i}{k_i} \\ o_i &= \frac{k_{si} + 3\sigma_i}{k_i} \quad \text{if } o_i > \frac{k_{si} + 3\sigma_i}{k_i}. \end{aligned} \tag{9}$$

Table 3. The Gaussian distribution parameters of k_{si}.

Features	μ_{si}	σ_{si}	Confidence Interval of μ_{si}
c_1	2.0782	0.0952	[2.0349 2.1215]
c_2	2.0081	0.1113	[1.9575 2.0588]
c_3	1.7408	0.1497	[1.6728 1.8089]
c_4	0.4776	0.0750	[0.4564 0.4987]
c_5	1.4948	0.1064	[1.4649 1.5248]
c_6	2.5147	0.1388	[2.4757 2.5538]
c_7	3.5054	0.3527	[3.4061 3.6047]
s_1	0.2954	0.0347	[0.2856 0.3052]
s_2	1.2408	0.0655	[1.2223 1.2593]
s_3	1.4293	0.1154	[1.3968 1.4618]
s_4	0.2410	0.0878	[0.2163 0.2657]
s_5	0.2744	0.0676	[0.2554 0.2935]



(a) (b) (c)

(d) (e)

Fig. 7. Some samples with their standard facial parameters frame.

As a feature scale, o_i describes the ith feature: k_{si}. For instance, in an input face, the face has a long jaw if $o_7 > 1$, otherwise the face has a short jaw when $o_7 < 1$. There are some samples shown in Fig. 7. In the figures, the frames are made up of these standard facial parameters based on their self-reference w, respectively.

Table 4 lists o_i of the faces in Fig. 7. In Table 4, to compare these data with the impressions derived from our minds, we find that these feature scales represent the self-features properly. For example, the face (a) has a short jaw, so its o_7 is 0.81. The process of obverting is also similar: o_4 of the face (c) is 1.27. The value tells us that the distance between the eyebrow and the eyes is longer than the distance

Table 4. The feature scales of the samples in Fig. 7.

	Face				
Features	(a)	(b)	(c)	(d)	(e)
o_1	0.94	0.933	1.10	1.08	0.959
o_2	0.92	0.902	1.01	1.03	0.952
o_3	0.99	0.862	1.17	0.93	1.06
o_4	1.25	1.07	1.27	0.94	0.967
o_5	0.88	1.03	1.05	1.04	1.03
o_6	0.84	0.98	1.00	1.04	0.891
o_7	0.81	0.87	1.03	1.1	1.12
o_8	1.05	1.14	0.98	0.95	0.97
o_9	0.93	1.11	0.98	0.92	1.02
o_{10}	0.96	1.17	0.95	1.05	1.16
o_{11}	1.13	0.77	0.87	1.08	0.86
o_{12}	1.08	0.83	0.87	0.95	0.79

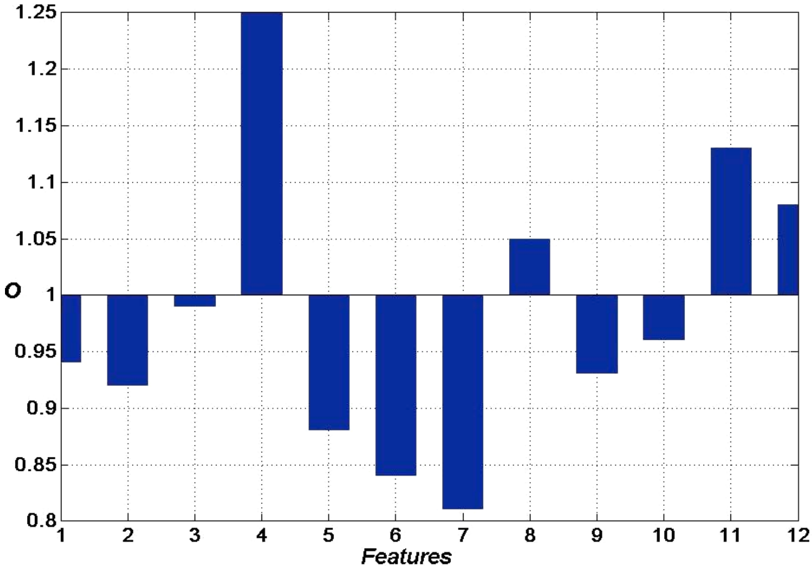


Fig. 8. The distribution of the feature scales.

in the standard face model. Additionally, o_4 is not intuitionistic enough to make a person who is not a caricaturist pay attention to the self-feature, but the SRM can find it out and output the information.

The o_i of Fig. 7(a) are plotted in Fig. 8. Each feature distributes on both sides of 1, and 1 is called as the feature line. If a feature scale locates above the feature line, the corresponding self-feature should be exaggerated to a bombastic tendency. On the contrary, if a feature scale locates beneath the feature line, the corresponding self-feature should be exaggerated to a contractible tendency.

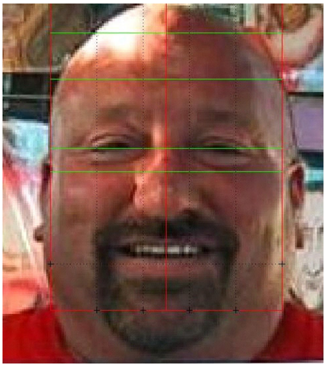
4. Experiments and Conclusions

4.1. *Experiments*

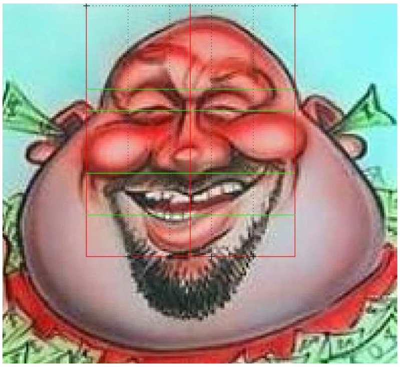
In order to test the SRM, we downloaded 27 caricatures[7] worked by some caricaturists from the Internet to verify the validity of the SRM.

Some facial pictures and their caricatures are shown in Figs. 9–12(a) and Figs. 9–12(b), respectively. These original and caricatured feature scales: o_o, o_c are plotted in Figs. 9–12(c). These feature scales present some consistent drawing skills:

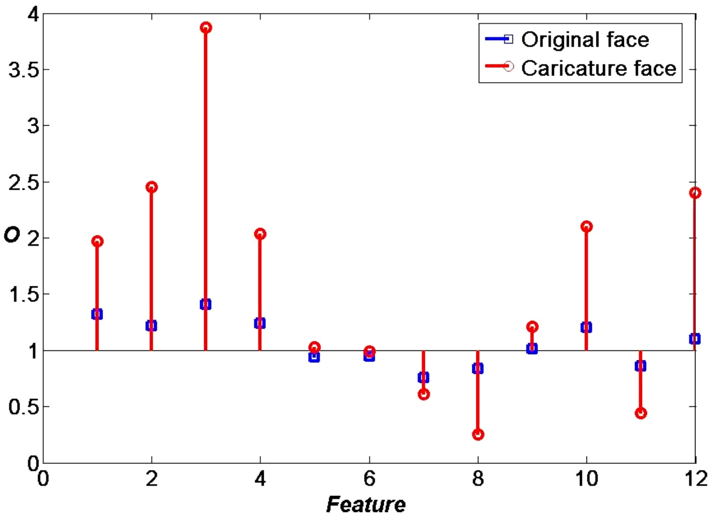
1. Generally, 98.6% of o_{oi} and corresponding o_{ci} locate on the same side of the feature line. The difference should be less than 0.05, even when they located on different sides.



(a)

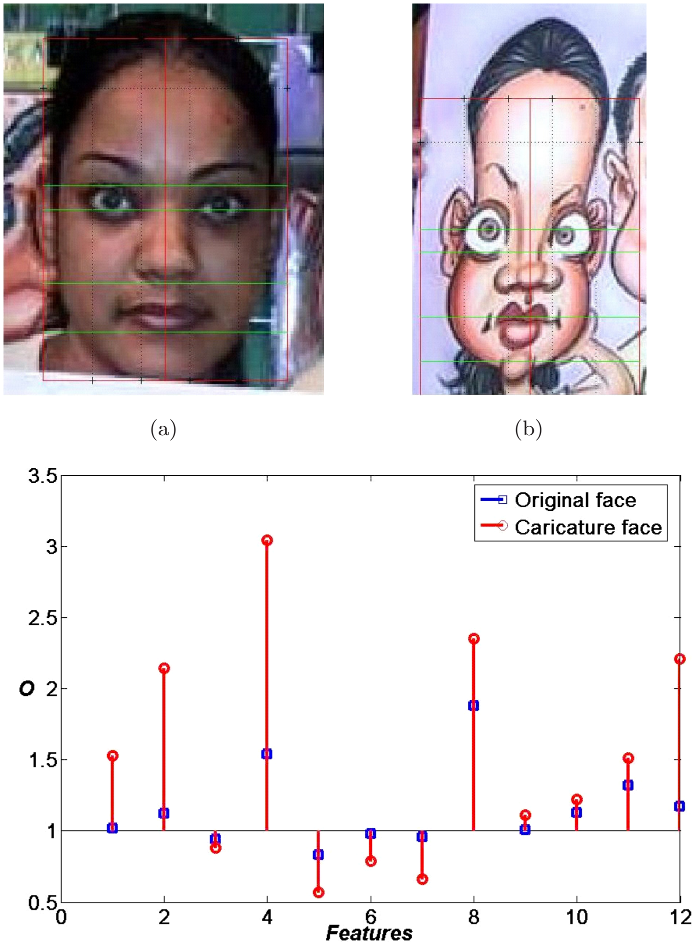


(b)



(c)

Fig. 9. The original and caricatured face, and their feature scales.



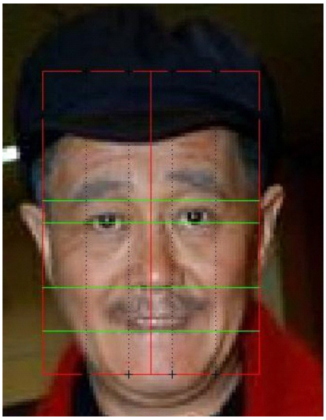
(a) (b)

(c)

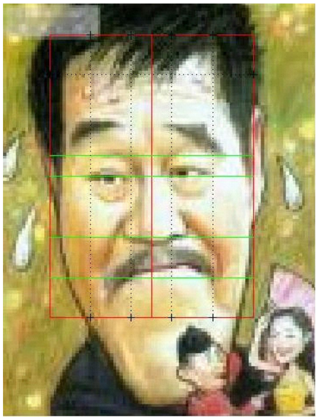
Fig. 10. The original and caricatured face, and their features scales.

2. If a feature scale is far away from the feature line in the original face, the distance between the feature scale and the feature line in the caricatured face is longer than the corresponding distance in the original feature scales. The relation complies with the drawing skill discussed in Ref. 15.
3. Caricaturists usually exaggerate these features with more distinction, and these features are equipped with greater standard deviation, such as: $k_{3,}k_{7,}k_{10}$.

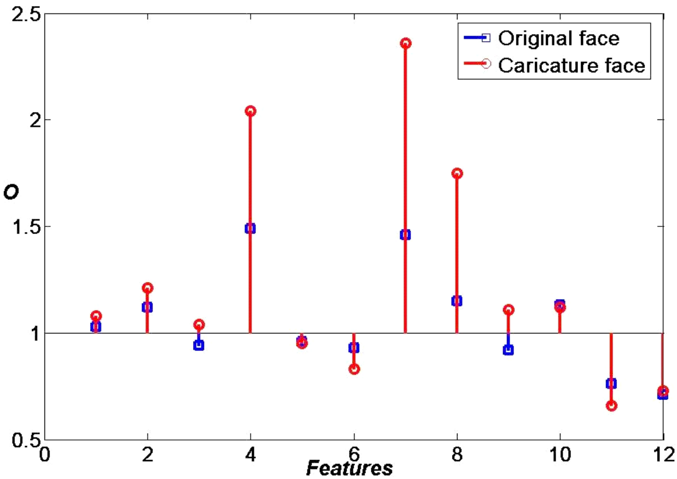
Going forward, we even find some bugs that do not comply with the drawing rules in these images drawn by caricaturists. For instance, in Fig. 12, o_{o7} and o_{c7} are located on different sides of the feature line, and the difference between them is 0.5. The distribution means that the length of jaw is longer than the



(a)



(b)



(c)

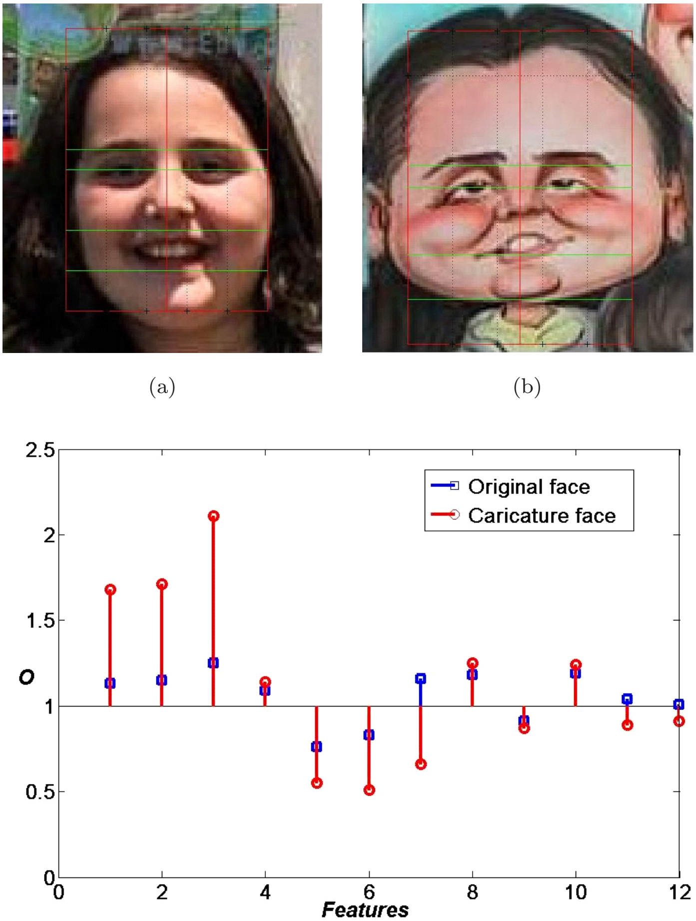
Fig. 11. The original and caricatured face, and their feature scales.

standard facial model, but in the caricatured face the jaw is exaggerated to be contractible. In fact, ordinary people can also see that the face has a long jaw.

Consequently, we state that the SRM can properly perform the drawing skills of caricaturists. In other words, the SRM can effectively estimate the facial features. Based on the SRM, we expect a caricatured face to be generated by exaggerating these self-features, the relation between the original and caricatured feature scales may be linear or not. Equation (10) shows an example of nonlinear relation:

$$o_{ci} = o_{oi}^2. \tag{10}$$

Figure 13 shows the relation between o_{ci} and o_{oi}.



(a) (b)

(c)

Fig. 12. The original and caricatured face, and their feature scales.

So, we propose the following method for automatic caricature: firstly, in an input image the original feature scales are evaluated. Then according to the original feature scales, the caricatured feature scales are calculated, such as in Eq. (10). Finally, these caricatured feature scales can be used to build up a caricatured face. As a result, the facial features of the caricatured face will comply with the original face strictly.

Eventually, in order to generate a caricature, it is necessary to spot some drawing points in a face, and these drawing points would be selected according to the principle of the composition. These drawing points are marked in Fig. 14. Obviously, the feature points are a part of the drawing points.

We know then that the B-spline curve is a suitable method for fitting curves. The property of B-spline curve implies that these fitting curves are smooth and

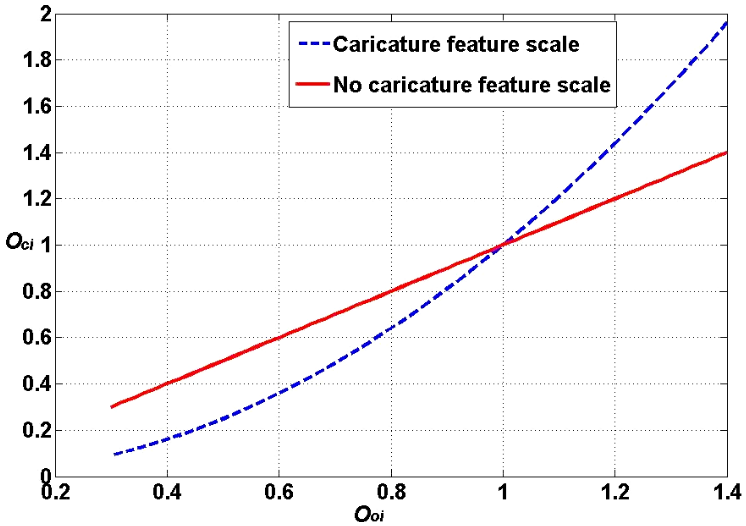


Fig. 13. The relation between o_{oi} and o_{ci} when $o_{ci} = o_{oi}^2$.

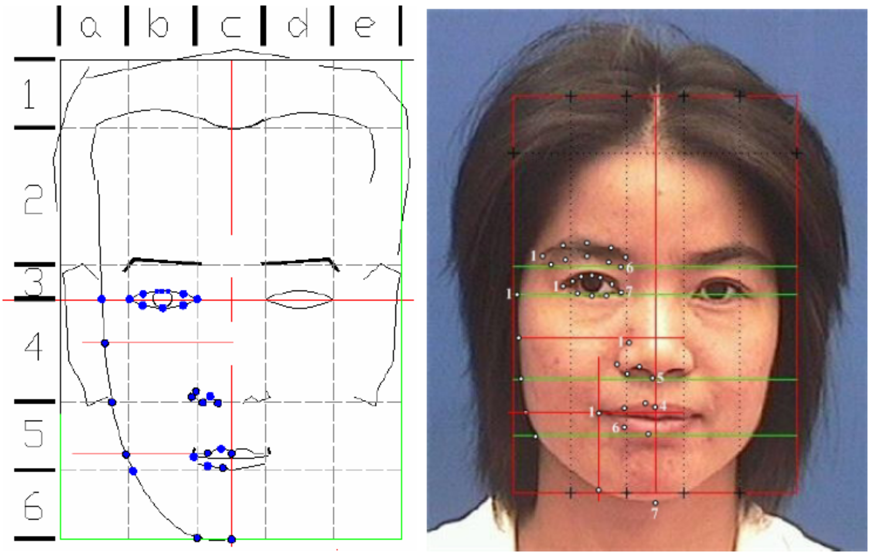


Fig. 14. These drawing points.

independent. Because the hair style is too complex to be spotted exactly by ASM, and what we need to do is only to demonstrate the caricatures with the self-features calculated by the SRM, we do not assign drawing points for the hair. In the following discussion, the hair points are detected manually. Using these drawing points, the portrait can be drawn as shown in Fig. 15.

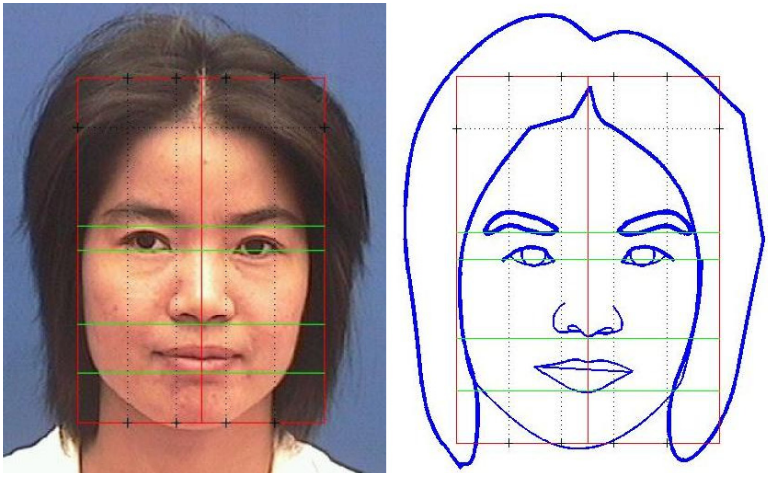


Fig. 15. To draw the portrait by using drawing points.

Now, for generating a caricature, we should move some drawing points according to certain features. For example, if $o_{o1} > 1.2$, $o_{c1} = 1.44$ (according to Eq. (10)), these points that indicate c_1 should be moved toward two sides to form a wider face. These coordinates of new feature points can be calculated by Eqs. (4) and (6) reversely. Figure 16 shows these caricatures generated by the above method.

In Fig. 16, the relation between o_{oi} and o_{ci} is calculated by Eq. (10). Obviously, the caricatures as seen in Fig. 16 can exhibit these self-features well.

4.2. *Conclusions*

The present paper focuses on the method for automatic caricature. A new method called the SRM is proposed. The method manages to quantify the relation between the self-reference and the facial features. In fact, these relations are not clear or visual, but by some skills in drawing a portrait, we find these recondite relations, and express them by the mathematical language and some equations. So, the contributions of this paper are to describe the relations and prove their validation in mathematics. In the SRM, based on Polyclitos Drawing Rule, we divide the facial features into three types — composition, size and form feature. A self-reference is extracted to evaluate and quantify the composition and size features of the input face. By collecting statistics, a set of standard facial parameter k_i is developed, and some feature scales which derived from comparing k_{si} of an input face with k_i are introduced to indicate the self-features quantifiably. Finally, some caricatures drawn by caricaturists and the SRM are involved in analyzing the validity of the SRM. We make sure that the SRM accords with the drawing skills of caricaturists, and it can generate good quality. So, the SRM is an effective method for quantifying the self-features, and it is a reasonable foundation of automatic caricature.

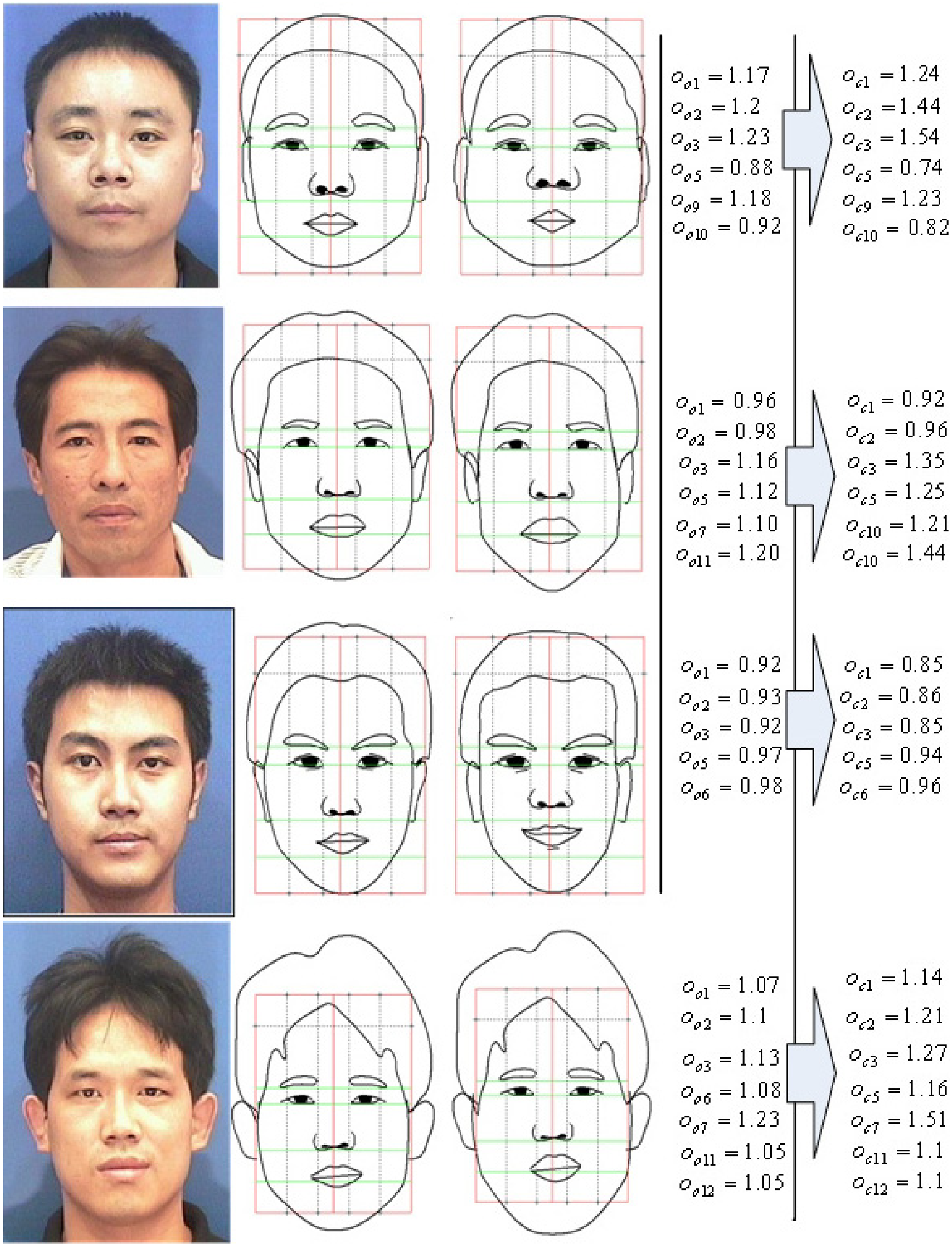


Fig. 16. Some caricatures with the self-features generated by the SRM.

Additionally, in the runtime stage, almost all cost of calculation should be attributed to detecting the feature points by using ASM. We do not plan to discuss ASM in this paper, there are several references[8,20,22] that describes in detail its running-time and cost of memory. In the SRM, as long as these self-feature points are detected, the remaining work is to employ Eqs. (4) and (9) to calculate o_i. There is no iteration in the processing, and the result can be achieved immediately.

4.3. *Discussions*

In fact, the body of people should be the biggest self-feature. We can recognize a person who we know even without looking at his or her face. The facial self-features are a part of all self-features. Some strong self-features include habitual poses, expressions and clothes. But technically, it is hard to extract one's habitual pose, expressions in just one picture. So, we just put emphasis on the facial self-features because these self-features are a visual way to express one's own features.

Two caricaturists certainly produce slightly different caricatured faces for a given face. Actually, aiming at a given face, the SRM can generate two different caricatures by changing the mapping relation between the original and the caricatured feature scales. The relation is described in Eq. (10). So, if we use another mapping relation, the caricature will not look the same. But the mapping relation must obey these following rules:

(1) o_{ci} and o_{oi} must locate on the same side of the feature line.
(2) If a feature scale is far away from the feature line in the original face, the distance between the feature scale and the feature line in the caricatured face should be longer than the corresponding distance in the original feature scales.

There were several students from the Fine Art College of Shanghai University have evaluated these caricatures. They thought that these caricatures aptly express the main self-features. Also, they indicated that some detail textures, such as double-edge eyelid, wrinkles or some facial traces, were too important self-features. These details will greatly vivify these caricatures. So, as a next step, we intend to develop some valid methods for detecting some details, specially, we wish to identify whether a person has a double-edge eyelid.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under grant No. 60675040, 60875058 and 2008 SMC Excellent Youth Teacher Award of SJTU. In Figs. 1–8 and 13–16, these face shots were done at CIIF 2006 (China International Industrial Fair, Shanghai, 2006), when our facial recognition system was exhibited. All of the authors are grateful to the owners of these original images used in this paper.

References

1. P. J. Benson and D. I. Perrett, Synthesising continuous-tone caricatures, *Image. Vis. Comput.* **9** (1991) 123–129.
2. M. Bianchini and L. Sarti, An eye detection system based on neural autoassociators, *Artificial Neural Networks in Pattern Recognition, Proc. Second IAPR Workshop, ANNPR 2006* (2006), pp. 244–252.
3. S. E. Brennan, Caricature generator: the dynamic exaggeration of faces by computer, *Leonardo* **18**(3) (1985) 70–178.
4. G. Civardi, *Drawing Portraits* (Search Press, 2004).
5. M. Dobes, J. Martineka and D. Skoupila, Human eye localization using the modified Hough transform, *Optik* **117** (2006) 468–473.
6. J. Hayashi, K. Murakami and H. Koshimizu, Method for automatic generation of caricatured profile in PICASSO system, *Syst. Comput. Japan* **29**(1) (1998) 12–19.
7. http://bjimg.focus.cn/photoshow/1322/440889.html.
8. J. Hyungkeun, L. Kyunghee and P. Sungbum, Eye and face detection using SVM, *Proc. 2004 Intelligent Sensors, Sensor Networks and Information Processing Conf.* (2004), pp. 577–580.
9. C. J. Kuo, R. S. Huang and T. G. Lin, Synthesizing lateral face from frontal facial image using anthropometric estimation, *IEEE Int. Conf. Image Proc.* **1** (1997) 133–136.
10. J. H. Langlois, L. A. Roggman and L. Mussleman, What is average and what is not average about attractive faces, *Psychol. Sci.* **5** (1994) 214–220.
11. K. H. Lin, K. M. Lan and W. C. Siu, Locating the eye in human face images using fractal dimensions, *IEE Proc. Vis. Imag. Sign. Process.* **148**(6) (2001) 413–421.
12. L. Lin, C. Hong and Y. Xu, Example based caricature generation with exaggeration, *Proc. 10th Pacific Conf. Computer Graphics and Applications*, Beijing (2002), pp. 386–393.
13. A. R. Mirhosseini, H. Yang and K. Lam, Adaptive deformable model for mouth boundary detection, *Opt. Engin.* **37**(3) (1998) 869–875.
14. K. Murakami, H. Koshimizu and A. Nakayama, Facial caricaturing based on visual illusion — a mechanism to evaluate caricature in PICASSO system, *IEICE Trans. Inform. Syst.* **E76-D**(4) (1993) 470–478.
15. L. Redman, *How to Draw Caricatures* (McGraw-Hill Publishers, 1984).
16. G. Rhodes and T. Tremewan, Averageness, exaggeration and facial attractiveness, *Psychol. Sci.* **7** (1996) 105–110.
17. M.-C. Rol and S.-W. Lee, Performance analysis of face recognition algorithems on Korean face database, *Int. J. Patt. Recogn. Artif. Intell.* **21**(6) (2007) 1017–1033.
18. R. N. Shet, K. H. Lai and E. A. Edirisinghe, Use of neural networks in automatic caricature generation: an approach based on drawing style capture, *IEEE Conf. Publication* n 2005–10882, *IEE Int. Conf. Visual Information Engineering*, VIE 2005 (2005), pp. 23–29.
19. M. Shiono, T. Takeda and T. Murayama, System for caricatured portrait drawing from facial photographs, *Terebijon Gakkaishi/J. Institute Television Engineers of Japan* **42**(12) (1988) 1380–1386.
20. H. T. Sugu, I. Yoshio and Y. Masahiko, Parallelization between face localization and person identification, *Proc. Sixth IEEE Int. Conf. Automatic Face and Gesture Recognition* (2004), pp. 183–188.
21. H. Wang, P. Li and T. Zhang, Boosted Gaussian classifier with integral histogram for face detection, *Int. J. Patt. Recogn. Artif. Intell.* **21**(7) (2007) 1127–1139.

22. J. Wang and T. Tan, A new face detection method based on shape information, *Patt. Recogn. Lett.* **21** (2000) 463–471.
23. R. Zhou and J. Zhou, Caricature generation based on facial feature analysis, *J. Comput. Aided Des. & Comput. Graph.* **18**(19) (2006) 45–52.

Fei Ni received his M.S. degree in engineering from the School of Mechanical and Electronic Engineering at WuHan University of Technology, China in 2004. He is currently working on the Ph.D in engineering science at Shanghai Jiao Tong University, China.

His research interests include image processing algorithm and signal processing.

The bios and photos of Q. Cao and Y. Zhao are unavailable.

Zhuang Fu received the Ph.D. in mechanical engineering from Harbin Institute of Technology in Harbin, China in 2000. From 2001–2002, he was a postdoctoral researcher at the Robotics Institute of the School of Mechanical Engineering at Shanghai Jiao Tong University in Shanghai, China, where he is now an associate professor.

His research interests include image processing algorithm, mobile robot and electrical system design.