# A Method for Exaggerative Caricature Generation from Real Face Image

Cheng Li[1]
Institute of Information Science
Beijing Jiaotong University
Beijing , China
11120370@bjtu.edu.cn

Zhenjiang Miao[2]
Institute of Information Science
Beijing Jiaotong University
Beijing , China
zjmiao@bjtu.edu.cn

***Abstract*—The generation of caricaturing portrait with exaggeration from real face image is one of the hot spots in the field of animation creation and digital multimedia entertainment. Based on improvement of traditional ASM facial feature points, this paper makes detailed definition of the human facial features and describes them in a way of proportion. Then, we propose a method which is based on the facial features and the relationship between them to generate exaggerated portrait from real face image. This method also introduces "contrast principle" while getting the exaggerated shape of face from two main aspects—facial form exaggeration and five senses organs exaggeration. At last, this method combines MLS image deformation method which is based on feature points to generate the exaggerated portrait of face image. Our experiments show that this method is practicable and can finally get results with good effect.**

***Keywords—exaggeration; facial animation; caricaturing portrait generation; ASM; image morphing;***

## I. INTRODUCTION

Since the 21st century, with the rapid development of the computer graphics technology and the rapidly growing demand for giving consideration to both aesthetics and artistry in computer graphics creation, Non-photorealistic rendering (NPR) technology arises at the historic moment. Most NPR methods process images as a whole mainly from its color, texture and other aspects but don't change the geometric shape or the relative position of its local content, that is, they don't need to morph during image processing. But when creating caricaturing portrait from real face images, people should analyze the facial features of face images and morph the images in an exaggerated way, then combined with the artistic effect of some other NPR methods we can finally get the caricaturing portrait with some particular visual effect.

From now on, people have already finished some work in this study. Brennan[3] proposed a system to produce exaggerated portrait with interaction in 1982. But this system needs users to have some foundation in exaggerated portrait creation, and won't provide the most obvious features and its orientation to exaggerate. Koshimiz etc.[4] summarized the creation method of capturing the difference between an average face and an exaggerated face in animation creation, and proposed the PICASSO caricature system. They did some research on how to capture the obvious features of a face, but their result is not very natural. Chen etc.[1] proposed an example-based facial sketch generation system. They asked the animators to create a large amount of facial sketch works based on which to build a sample library that composed of face images and its sketch works. Then they let the computer to learn the creation method of the animators from a large amount of data. At last, they adopted the nonparametric estimation method to obtain the sketch of target face image which had good effect. But this method needs professional animators to create large amount of sketch works for learning for each exaggeration style, and the exaggeration style totally depends on the animators who create it. This makes the style of results monotonous and the process time-consuming. After that, Liang, Chen etc.[5] improved their previous work. They analyzed the original training set with PLS and divided it into several subsets which represented different exaggeration types and then they classified the input image into some appropriate exaggeration types and exaggerated it with the parameters learned from these types. This approach is able to capture the obvious features among different faces in an objective way but it requires a very huge sample library to provide support so it's not very practical. Yan etc.[6] did much work in feature extraction and image morphing, and proposed to adopt fish-eye distortion in image morphing method. But it's hard to measure the degree of deformation, so a lot of adjustment may be needed to achieve good visual effect. This makes this method time-consuming and requires much intervention from users.

Through observing and analyzing the way of caricaturing portrait creation, this article firstly gives detailed definition and description of human facial features based on ASM. Then, a method to achieving exaggeration style from aspects of both facial form and five sense organs is proposed. Meanwhile, specific rules of exaggeration are also made to obtain exaggerated shapes. We also introduce the concept of "public face" to replace the traditional way of calculating average face from a very huge sample library and thus make this method more practicable and simplify the process of implementation. At last, we use the MLS image deformation method which is based on feature points to morph the input face image in an exaggerative way to finally get the portrait with exaggeration.

## II. DEFINITION AND DESCRIPTION OF PRINCIPA FACIAL FEATURES

### A. Definition of Facial Features Based on ASM

The active shape model (ASM) is one of the common and effective ways to obtain facial features of face image. It can

978-1-4799-2190-4/13 $31.00 © 2013 IEEE
DOI 10.1109/ACPR.2013.150



IEEE computer society

usually well obtain the shapes of the facial form and the five sense organs from the face images with front face. But affected by different posture, facial expressions, environment light and the view of camera, sometimes the result of this method is not very exact. To overcome this problem, we can adjust the position of feature points manually to insure its accuracy.

Through ASM method, we can finally obtain 68 facial features points. These points can give the shapes of eye, eyebrow, nose, mouth and face. How to use these only 68 points to describe each feature of human face as objective and accurate as possible will directly affect the final results. Based on these 68 feature points, this article makes detailed definition of the principal facial features and lists them as below:

TABLE I. DEFINITION OF PRINCINPAL FACIAL FRATURES BASED ON ASM

| *Definition Name* | | *ASM Number* | *Description and Calculation* | *Measurement Scale* |
|---|---|---|---|---|
| **Facial Form Features** | Zhong Ting Length | 15,20,41 | vertical distance between eyebrow bottom and nose bottom | Face Width |
| | Xia Ting Length | 7,41 | vertical distance between nose bottom and jaw top | Face Width |
| | Face Length | 21,15,7 | average of vertical distances between point pairs P21,P7 and P15,P7 | Face Width |
| | Forehead Length | Non | average of zhong ting and xia ting | Face Width |
| | Face Width | 0,14,1,13 | average of horizontal distances between point pairs P0,P14 and P1,P13 | Face Length |
| | Standard Length | 31,36 | Horizontal distance between P31,P36 | Non |
| | Jaw Length | 7,57 | Vertical distance between P7 and P57 | Xia Ting Length |
| | First Angle | 1,2,12,13 | Average of angle between horizontal line and P1-P2 line and angel between horizontal line and P12-P13 line | Non |
| | Second Angle | 1,2,12,13,3,4,10,11 | Average of angle between P1-P2 line and P3-P4 line and angel between P10-P11 line and P12-P13 line | Non |
| | Third Angle | 4,7,10 | angle between P4-P7 line and P7-P10 line , represent whether the jaw is peaked or round | Non |
| **Mouth Features** | Mouth Width | 48,54 | Horizontal distance between P48 and P54, represent the width of mouth | Jaw Width |
| | Upper Lip Thickness | 53,64,66 | Average of vertical distances between the point pairs P53,P66 and P64,P66 | Under Lip Thickness |
| | Under Lip Thickness | 57,66 | Vertical distance between P57 and P66 | Upper Lip Thickness |
| | Jaw Width | 3,11 | Horizontal distance between P3 and P11 | Face Width |
| **Nose Features** | Nasion Width | 37,45 | Horizontal distance between P37 and P45 | Standard Length |
| | Nose-bridge Width | 36,44 | Horizontal distance between P36 and P44 | Nasion Width |
| | Nose-wing Width | 39,43,40,42 | average of horizontal distances between point pairs P39,P43 and P40,P42 | Nasion Width |
| | Nose Length | 37,45,41 | Vertical distance between middle point between P37,P45 and P41 | Distance between Upper lip and |
| **Eyes Features** | Eye Length | 27,29, 34,32 | Horizontal distance between P27 and P29,P34 and P32 | Eye Height |
| | Left Eye Height | 28,30, 33,35 | Vertical distance between P28 and P30, P33 and P35 | Eye Length |
| **Eyebrow Features** | Brow Length | 21,24;18,15 | Horizontal distance between P21 and P25 in the left, right brow is same | Face Width |
| | Brow Thickness | 22,26,23,25 17,19,16,20 | Average of distances between the point pairs P22-P26 and P23-P25 in the left, right brow is same | Brow Length |
| | First Brow Angle | 21,22,23; 15,16,17 | angle between P21-P22 line and P22-P23 line in the left, right brow is the same position | Non |
| | Second Brow Angle | 22,3,24; 16,17,18 | angle between P22-P23 line and P23-P24 line in the left, right brow is the same position | Non |

As is shown in the table, this article divides facial features into five parts---- facial form, mouth, nose, eye, and eyebrow and then makes detailed definition and description of them. The ways to measure different features is mainly Euclidean distance and angle values. But due to the different resolution of different images and different proportion of face in the whole image, it's unable to measure the real differences between different face images just by Euclidean distance or such absolute measurement. For this reason, this article uses a way of proportion description to measure the same features of different face images. The contents of sub head "Measurement Scale" of each feature in the table above is to be used for this kind of measurement. For example, for the pair of features face width and face length, due to their relativity, the man with a

long face usually seems with a narrow face at the same time; however, a man with a wide face usually seems with a short one, therefore, we measure this pair of features with the proportion of face length and face width. For another example, for the feature of mouth size, people usually judge this feather through the proportion between mouth and cheek, so the "measurement scale" of mouth width is the feature "cheek width".

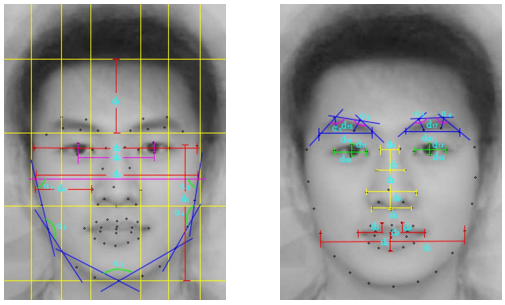

Fig. 1. Description of facial features and five sense organ features based on ASM

### B. *Inprovement of ASM Feature Points*

Traditional ASM method can only extract 68 facial feature points from face image. Although it's no problem to obtain the principal facial features through these points, it's hard to make accurate adjustment while doing the image morphing because lacking of enough points. For this reason, this article improves traditional ASM feature points through the way of "proportion interpolation" and increases the number of points from 68 to 144 to enhance the control of local features of a face image.

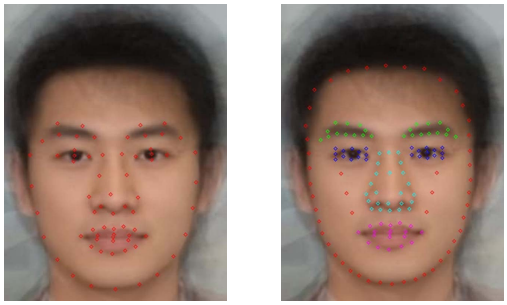

Fig. 2. Traditonal ASM feature points and the improved result in this article

## III. Inplementation of Portrait with Exaggerative Style from Face Image

### A. *System Framework of Exaggerated Portrait Generation*

Fig. 4 shows the flow chart of implementing the system of exaggerated portrait generation. As is shown in the figure, all the data are expressed in parallelogram; data processing is expressed in rectangular. In most of existing methods, people often use statistical methods to calculate the average shape of face features from large amount of face images and regard the average shape as the standard reference of exaggeration differences. However, because of different ethnic and geographical distribution, in different countries or even different areas of the same country, there are big differences of facial features among people. Especially in countries like China, there is very complex population and gene distribution. So it's very hard to get the real sense of average face feature but we may choose an average shape as long as it can meet our needs.

South African photographer Mike took tens of thousands of pictures in the global and synthesized dozens of pictures of male and female "Public Face" in different countries using software. We think Mike's "Public Face" has the significance of statistics, so we use it as the "Standard Face" in our system. Of course, we can also use the traditional method to calculate the average to get "Standard Face", but in here, we are aimed at introducing a method of generating exaggerated portrait which is practical, convenient, fast and objective as possible. Through experiments, we can finally get satisfactory results with our method, thus proves the feasibility of this attempt.

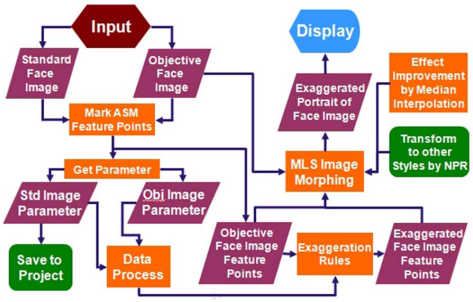


Fig. 3. Syetem Framework of exaggerated portrait generation

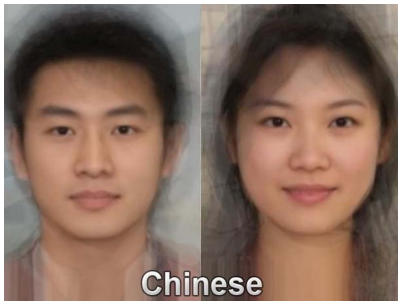


Fig. 4. Standard face image
(All rights reserved by Mike)

### B. *Exaggeration Rules and Exaggerated Shape Generation*

Through observation of the portrait works created by animators, the author finds that whatever the face shape is, the animator will certainly make some kind of exaggeration of it. They often shrink the area above cheekbone to some extend while the main area to exaggerate is from cheekbone below to jaw. Forms of exaggeration are manly shrinkage, sharpening, inflation, stretching and so on. With the exaggeration of the face shape, the animators will also focus on to exaggerate the obvious features of the five sense organs.

Therefore, this article makes such kind of exaggeration rules as exaggerate the face image from two aspects---- facial

form and five sense organs. For the facial form features, this article mainly concerns on the size, length of face or whether the jaw is round or not and so on while for the features of five sense organs, this article will find out the most obvious feature to exaggerate through calculation, compare and analysis.

In addition, the author finds that the animators usually consider the facial features as a whole while creating portraits. That is, they not only consider certain features of facial form or five sense organs, but also their relationship. For this reason, this article proposes "Contrast Principle" and applies it into the algorithm of exaggerated shape generation. Specifically, if someone has a long face, we exaggerate this feature by stretching it and meanwhile we shrink the width of face. On the contrary, if someone has a wide face, we exaggerate this feature and shorten its length at the same time. For another instance, if someone has a long nose, we exaggerate this feature and shorten the length of nasion and eye distance in the meantime; otherwise, we increase the eye distance and the length of nasion.

To get exaggerated portrait from face image, we should first get the exaggerated facial shape which consists of exaggerated feature points. This needs to do some exaggeration-processing to the original feature points. We use affine transformation to complete this. Assume that the position of the original point is（$X_0$,$Y_0$）, the position of the reference point is （ $X_c$,$Y_c$ ）,the position of point after transformation is （$X_1$,$Y_1$）, then, the process of transforming can be expressed with the equation below:

$$(X_1,Y_1,1)^T=\begin{pmatrix}1-\cos\theta*s & \sin\theta*s & (X_c*\cos\theta-Y_c*\sin\theta)*s\\ -\sin\theta & 1-\cos\theta*s & (X_c*\sin\theta+Y_c*\cos\theta)*s\\ 0 & 0 & 1\end{pmatrix}(X_0,Y_0,1)^T \quad (1)$$

As is shown in the equation, s is the scale factor, θ is rotation factor, and we use these factors as the main parameter of exaggeration.

*1) Algorithm of Facial Form Exaggeration*

For the facial form features, the general thought of exaggeration in this article is from global to local, that is, firstly to make a judgment of the face width and process it, then, to judge whether the face is a square face or a round face and whether the jaw is thin or round to further process it, the specific algorithm is shown below:

```
float k = inputface/stdface;
//k is the ratio of length-width ratio of input face image and standard image
if (k>1)   // the input face is a long face
{
    exagfaceftpt[i].y +=α1*pow(basiclength,1/k)*atan(k);
    // vertical adjustment of facial form feature points
    exagfaceftpt[i].x ±=α2 *pow(basiclength,1/k)*atan(k);
    //horizontal adjustment of feature points based on "Contrast principle"
}
else       //k<1, so the input face is a wide face
{
    exagfaceftpt[i].x±=α3*1/k*basiclength;
    // horizontal adjustment of facial form feature points
     exagfaceftpt[i].y-=α4*1/k*basiclength;
    // vertical adjustment of feature points based on "Contrast principle"
}
k = GetFirstAng()/stdface.GetFirstAng();
//k is the ratio of the feature-first angle of input image and standard image
float theta = atan(GetFirstAng()*pow(k,2))-atan(GetFirstAng());
//obtain the exaggeration parameters based on the differences
exagfaceftpt[i] = AffineTransform(theta,1);
//adjust the feature points near the first angle by affine transformation
k = GetThirdAng()/stdface.GetThirdAng();
// k is the ratio of the feature-third angle of input image and standard image
theta = atan(GetThirdAng()*pow(k,2))-atan(GetThirdAng());
//obtain the exaggeration parameters based on the differences
exagfaceftpt[i] = AffineTransform(theta,1);
//adjust the feature points near the third angle by affine transformation
k = GetSecondAng()/stdface.GetSecondAng();
// k is the ratio of the feature-second angle of input image and standard
image theta = atan(GetSecondAng()*pow(k,2))-atan(GetSecondAng());
//obtain the exaggeration parameters based on the differences
exagfaceftpt[i] = AffineTransform(theta,1);
//adjust the feature points near the third angle by affine transformation
```

In the algorithm above, we adjust the position of points mainly through these relations:

$$\begin{cases}\Delta_{xy}=\alpha B_0^{1/k}\arctan(k)\\ \Delta_{xy}=\alpha B_0 k^{-1}\\ \Delta_{\theta}=\theta_0(k^{\alpha}-1)\end{cases} \quad (2)$$

Where, $\Delta_{xy}$ represents for the exaggeration adjustment of distance, $\Delta\theta$ represents for the exaggeration adjustment of angle. $B_0$ is the standard length, α is the experimental parameter, k is the difference of feature. Based on these relations, while k>1, the difference can be amplified positively, and the degree of amplification is positively related to k; while 0<k<1, the difference can be amplified to the opposite side, and the degree of amplification is negatively related to k.

*2) The Discovery of Obvious Features of Five Sense Organs and Its Exaggeration*

For the exaggeration of features of five sense organs, we should firstly find out the most obvious features. In this article, we firstly find out the most obvious feature area through dividing the five sense organ features into 4 feature areas and calculating their variance. Then we find out the feature with largest positive difference and the feature with largest negative difference in this area and exaggerate them.

Specifically, assume that the feature set of five sense organs of the input image is $\{X_i, i=1,\ldots,4\}$, accordingly, the feature set of the standard image is $\{X_i', i=1,\ldots,4\}$,where, $X_i$ and $X_i'$ is separately the subset of each feature set. And these subsets represent the feature of eye, mouth, nose and eyebrow. In each subset, there are some specific features $F_{ik}$, $F_{ik}'$ (k = 1,…, ni), calculate the variance of each subset:

$$D_{X_i}=\sum_{k=1}^{n_i}(F_{ik}-F_{ik}')^2 \quad (3)$$

Compare these values of variance, the bigger the value is, the more obvious difference between the feature of input image and standard image that this subset represents for. Then compare each feature in this subset to find out the feature with

largest positive difference $F_{imax}^{+}$ and the feature with largest negative difference $F_{imax}^{-}$, while this pair of features is the ones we will exaggerate. At last, we exaggerate these features base on "Contrast Principle" to get the exaggerated shapes of the five sense organs. If we only find one of $F_{imax}^{+}$ and $F_{imax}^{-}$, then we just exaggerate the one we find.

Like the facial form exaggeration, this article makes different exaggeration rules for each area of five sense organs. Limited by space, we will not give the detail of algorithm one by one.

Through the exaggeration of facial form and five sense organs, we can get the exaggerated shape of the input face image. Then based on the relationship between the original shape and the exaggerated shape, we morph the image to get the final portrait with exaggeration.

### C. *Image Morphing and Result Generation*

We select the MLS image deformation method in this article as our image morphing method, because this method obtain the deformation parameters based on the position of feature points, this can just fit the method based on the shape of feature points we proposed previously very well. In this method, feature point is also called "control point", while the feature point whose position is changed is also called "target point", and the image will deform towards the direction of "target point". But there is also a shortage of this method that when the image is stretched, there will be black cracks and holes in the image. For this problem, we adopt the conditional median interpolation method to improve it. Experiment shows this method can repair the image defects made by MLS very well. Compare to the traditional median filter method, this improved conditional median interpolation method won't makes image blurry and can reserve more details of image, simple and effective.

We also add the transformation from original image to sketch style to our system while generating the result. Thus, through processing, we can get the portrait with exaggeration in different visual effect. The result generated with the method of this article is shown as below:

Fig. 5. The original face images for test (All from internet)

Fig. 6. The portrait with sketch style generated by our method

Fig. 7. The original face images for test (All from internet)

Fig. 8. The portrait with exaggration generated by our method

## IV. CONCLUSION AND PROSPECT

Based on ASM facial feature points, this paper makes detailed definition and description of the principal facial features and proposes a method to implement exaggeration from the aspects of both facial form and five sense organs. Considering the importance of interrelation of facial features in caricature portrait creation, we propose "Contrast Principle" and apply it into the making of exaggeration rules. We choose the MLS image deformation method while doing the image morphing, and combined with the NPR algorithm of sketch stylization, we finally get the portrait with exaggeration in different visual effects.

We will improve our work and do some further research in the aspects below: 1. we should study more creation styles of different animators to improve the existing algorithms and make more kinds of exaggeration rules to get various effects with different style. 2. Study image morphing methods further to improve the morphing speed and enhance the control and quality of image morphing. 3. Try more texture effect when generating portrait with exaggeration besides of sketch style.

## REFERENCES.


[1] Chen Hong, Zheng Nanning, Xu Yingqing., "An example-based facial sketch generation system", Journal of Software, vol. 3, pp. 202-208, 2003.

[2] Brennan S *et al*, "Caricature generator. Cambridge", Massachusetts Institute of Technology, 1982.

[3] Koshimizu H, Tominaga M, Fujiwara T, *et al*, "On Kansei facial processing for computerized facial caricaturing system Picasso", Proceedings of IEEE International Conference on Systems , Man , and Cybernetics , Tokyo , 294 ~299, 999.

[4] Lin Liang, Hong Chen, *et al*, "Example-based Caricature Generation with Exaggeration", Proceedings of the 10th Pacifc Conference on Computer Graghics and Applications (PG'02)

[5] Yan Fang, Fei Guangzheng , Liu Tingting, *et al,* "A Ggeneration Algorithm of Caricarure Portrait", Journal of Computer and Aided Design & Computer Graphics, vol. 19, No. 4, pp. 442-448, Apr, 2007

[6] S. Schaefer, T. Mcphail, and J. Warren., "Image deformation using moving least squares", SIGGRAPH 2006, ACM Press, New York, NY, USA.