I.J. Information Technology and Computer Science, **2010, 1, 1-9**
Published Online November 2010 in MECS (http://www.mecs-press.org/)

An Improved Active Shape Model Application on Facial Feature Localization

Min.Wang
The School of Information and Control Engineering
Xi'an University of Architecture & Technology
Xi'an, China
Email: wangmin1329@163.com

Yi ling.Wen，Li.Fang and Wei ping.Sun
The School of Information and Control Engineering
Xi'an University of Architecture & Technology
Xi'an, China
Email:{wenyilingwenzi,fangli1231, pp143sun}@163.com

***Abstract*—An improved active shape model is proposed in this paper. The proposed algorithm includes the following four aspects. Firstly, this paper adopts a semi-automatic facial feature points marking tool. Secondly, this paper proposes to extract 2D gradient feature on the highest level and the higher level of multi-resolution pyramid images, and use Gabor wavelet transformation to extract the lowest level's 2D texture feature. Thirdly, this paper adopts a new method of decomposition of multi-resolution pyramid. Pyramid images are got by wavelet transformation. Finally, this paper uses an improved searching scheme of multi-resolution pyramid. The length of 2D profile searching rectangle is changed according to different pyramid levels. Experimental results demonstrate that the proposed algorithm exhibits better performance than the original ASM.**

***Index Terms*—active shape model, facial feature localization, texture profile, texture model, multi-resolution pyramid**

I. Introduction

Facial feature extraction is an important process in facial image analysis. It can be used to face recognition and verification, facial animation, face image compression, etc.Important facial features include the face contour, eyes, eyebrow, nose, mouth, etc [1]. Many approaches have been proposed for the extraction of these facial features in recent years, and the highest attention has been given to active shape model (ASM) and active appearance model (AAM). The active shape model (ASM), proposed by Cootes and Taylor, has attracted much attention in different areas of feature extraction, such as for medical images, face images, hand gestures, etc. Because ASM captures the specific characters of a shape and its variations are denoted by a statistical shape model, it can therefore adapt to any predefined shapes more effectively and accurately.

Although original ASM algorithm can get good convergence results in many cases, it strongly relies on the position of initialization of mean shape model. When the mean shape model is far from the real face contour, it is prone to get a bad local convergence result. At the stage of searching for target points, the original ASM algorithm only uses the local gray profile model around landmark to ensure the accuracy of localization of feature points, which makes the localization result easily occur error. Besides, if the light condition of target images is largely different from that of training images, it will cause the built gray profile model can't guide the move of feature points rightly, which makes the algorithm not converge, and the localization result of feature points not ideal, even localization failure. On account of those above defects, Du Chen introduced the edge information and part information of face to the matching process of ASM which improved the performance of ASM [2].Fan-yuhua used log-Gabor coefficients to describe the local texture distribution and built models for each feature point to increase the robustness to illumination change and other noises [3].

In this paper, an improved active shape model is proposed after deeply researching the original active shape model. The proposed algorithm includes the following four aspects. Firstly, this paper adopts a semi-automatic facial feature points marking tool. Adoption of the marking tool greatly reduces the burden of marking points. Secondly, this paper proposes an improved 2D texture model, which is based on Gabor wavelet coefficients and gradient feature. We propose to extract 2D gradient feature on the highest level and the higher level of multi-resolution pyramid images, and use Gabor wavelet transformation to extract the lowest level's 2D local texture feature. This method makes facial feature points localization more accurate. Thirdly, we improve the multi-resolution pyramid's decomposition method of the original active shape model. This paper proposes to use wavelet transformation to get pyramid sub-images. Finally, this paper modifies the multi-resolution searching scheme of the original active shape model. We change the length of 2D profile searching rectangle according to different pyramid levels in the proposed algorithm.

The organization of this paper is as follows. In Section 2, we provide a brief introduction of the original active shape model. Our improved ASM is introduced in Section 3. Experimental results and the performance comparisons are given in Section 4, and a conclusion is drawn in Section 5.

II. Brief Introduction of the Original Active Shape Model

As a kind of statistical model, the original ASM algorithm includes three procedures: feature points mark, construction of the ASM, and localization of contour. The following gives a brief introduction of the original ASM [1] [4] [5] [6].

A. Feature Points Mark

The model-construction procedure of ASM requires marking n key facial feature points on each training image manually. N facial images are chosen, and each image is marked n feature points as training data manually. The marked feature points can be represented as the following vector.

$$\mathrm{X}_i = [x_{i1},...x_{in}, y_{i1},...y_{in}]^T \quad i = 1,2...N \tag{1}$$

where (x_{ij}, y_{ij}) are the horizontal and vertical coordinates of the jth feature point of the ith training image. T represents the transpose operation. The role of landmark points is controlling the shape of model contours. A typical setup in our system consists of 58 manually assigned landmark points $(n = 58)$. Fig.1 shows a training image with its landmark points marked. All the horizontal and vertical coordinates are arranged of a single group, and are matched with each other. Here, N represents the number of training images, and n represents the number of feature points of each image .Each X_i represents a shape vector.

Figure 1. Locations of the points used to represent a face

B. Construction of The Active Shape Model

The model-construction procedure of active shape model is separated to the following two steps.

The first step: In order to make the shape model independent of the position, size and orientation of the face, we align these shape vectors by rotation, scaling and translation. The measure of the alignment is to minimize the sum of the squared distances of the feature points [6].

The second step: After the alignment, principal component analysis (PCA) is then applied to capture most of the shape variations. The PCA algorithm is as follows.

The mean of the shape can be expressed as

$$\overline{\mathrm{X}} = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} \mathrm{X}_i, \tag{2}$$

and the corresponding covariance as

$$C = \frac{1}{N} \sum_{i=1}^{N} (X_i - \overline{X})(X_i - \overline{X})^T. \tag{3}$$

The eigenvalues $(\lambda_0, \lambda_1, \cdots \lambda_{2n-1})$ and the corresponding eigenvectors $(p_0, p_1, \cdots p_{2n-1})$ of the covariance matrix C, where $\lambda_i \geq \lambda_{i+1}$, can then be computed. Because the higher eigenvalues corresponding to eigenvectors contain much information of shape change, we can use some higher eigenvalues corresponding to eigenvectors to represent any shape vector approximately. We select the first j eigenvalues satisfying the following equation:

$$\sum_{k=0}^{j} \lambda_k \Big/ \sum_{k=0}^{2n-1} \lambda_k \geq t \tag{4}$$

The determination of the value of t varies from different applications. Here, we set t be 0.98. The first j eigenvectors with the largest eigenvalues arranged in descending order can be formulated as $P_s = (p_0, p_1, \cdots p_{j-1})$ .Therefore, a shape model can be approximated as follows:

$$X = \overline{X} + P_s b_s, \tag{5}$$

where b_s is a weight vector for the j eigenvectors , referred to as the shape parameters , which can be computed as follows:

$$\mathrm{b}_s = (b_0, b_1...b_{j-1}) = P_s^{-1}(X - \overline{X}) \tag{6}$$

When fitting the shape model to target shape, the value of b_i is constrained to lie within the range $\pm 3\sqrt{\lambda_i}$,where λ_i is the eigenvalue corresponding to the ith principal component, and λ_i match with b_i .This can ensure that this range of shape parameters can represent most of the shape variations in the training set.

According to PCA, $\overline{\mathrm{X}}$ is the mean shape vector, and P is corresponding eigenvectors matrix. Any facial shape X can be represented as $\mathrm{X} = \mathrm{T}_c(\overline{\mathrm{X}} + \mathrm{Pb})$,where c is geometry parameter, and b is shape parameter of PCA. Geometry parameter c includes the horizontal and vertical offset $\mathrm{X}_t, \mathrm{Y}_t$,the scale variable s and the angle

variable θ . T_c can be represented as follows, which expresses the geometry variation of the shape.

$$\mathrm{T_c}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X_t \\ Y_t \end{bmatrix} + \begin{bmatrix} s\cos\theta & -s\sin\theta \\ s\sin\theta & s\cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$
$$\mathrm{c} = X_t, Y_t, s, \theta \quad (7)$$

Facial shape X is represented as a function, which is based on geometry parameter c and PCA shape parameter b. So, the problem of facial feature localization is changed to the problem of assessing the optimum geometry parameter c and the PCA shape parameter b.

C. Localization of Contour

ASM adopts gray match method which is based on profile. Its main idea is that we get the local texture feature around each feature point, which is the normalized derivative of the profiles sampled perpendicular to the feature point contour and centered at the feature point.

Given a point of (x_j, y_j) ,the sampling of l pixels on two sides of this point along the normal direction of the contour leads to the profile length being $2l+1$.We use gray value to formulate the following vector:

$$\mathrm{g}_{ij} = (\mathrm{g}_{ij1}, g_{ij2} \cdots \mathrm{g}_{ij(2l+1)})^T. \quad (8)$$

where g_{ij} represents the gray-level profile of the jth feature point of the ith training image, and $g_{ijk}(k = 1,2 \ldots 2l+1)$ is the gray-level intensity of a corresponding pixel. The derivative profile of g_{ij} has a length of $2l$,and is given as follows:

$$\mathrm{dg}_{ij} = (g_{ij2} - \mathrm{g}_{ij1}, g_{ij3} - \mathrm{g}_{ij2}, \ldots \mathrm{g}_{ij(2l+1)} - \mathrm{g}_{ij2l})^T. \quad (9)$$

To reduce the impact of illumination and contrast, we normalize the above vector, which is given as follows:

$$\mathrm{y}_{ij} = \frac{dg_{ij}}{\sum_{k=1}^{2l} \left| dg_{ijk} \right|}. \quad (10)$$

where $dg_{ijk} = g_{ij(k+1)} - g_{ijk}$ .So, the mean gray-level profile of all the images at the jth feature point is represented as follows:

$$\overline{g_j} = \frac{1}{N}\sum_{i=1}^{N} g_{ij}. \quad (11)$$

The corresponding covariance is represented as the following formula.

$$\overline{G_j} = \frac{1}{N}\sum_{i=1}^{N} (g_j - \overline{g_j})(g_j - \overline{g_j})^T \quad (12)$$

The mean gray-level profile and covariance are the matching features of the feature point. We make the same sampling analysis of each feature point, and then we can get the statistical property of the whole contour. Those statistical properties have the distribution of Gauss function. We adopt the Mahalanobis distance measure to compute the difference between the new profile and the mean profile, which is defined as the following formula.

$$f(g_{si}) = (g_{si} - \overline{g_i})^T G_i^{-1} (g_{si} - \overline{g_i}) \quad (13)$$

ASM adopts gray-level profile of the contour to iterate. In each iteration step, the current model's location and shape is changed by changing the PCA shape parameter b to finish the match of model and the test image's contour.

III. The Improved Active Shape Model Algorithm

A. Adoption of A Semi-Automatic Facial Feature Points Marking Tool

The ASM algorithm needs to be marked a number of training images manually at the training stage. Each training image should be marked a number of feature points. Thus, feature points marking becomes time-consuming.

In this paper, we adopt a semi-automatic facial feature points marking tool, which greatly reduces the burden of marking points.Fig.2 is our adopted semi-automatic facial feature points marking tool.

Figure 2. A semi-automatic facial feature points marking tool

Our adopted feature points marking tool has the following characters.

- We can load two images at the same time. So, we can mark them at the same time.
- We can mark feature points either on the first image or on the second image.
- The marking tool has the function of position prediction of feature points. If we mark point on one of the two images, the corresponding point will occur on the corresponding image automatically. This function greatly improves marking efficiency. However, this position prediction function is based on the geometrical position of marked points, thus sometimes position prediction is not accurate enough and it needs us to adjust the position manually.
- We can move and delete the marked points.

■We can save the marking state, which allows us to mark the same image discontinuously.

Because the marking tool has the position prediction function, and the marked facial key parts (eyebrow, eye, mouth, face contour) usually have regular geometric shapes, the adoption of the marking tool can greatly improve marking efficiency and reduce marking burden.

B. The Modified 2D Texture Model Based on Gabor Wavelet Coefficient and Gradient Feature

The original ASM uses 1D gray profile of the feature point as the match feature, and its computation complexity is relatively low, but sometimes its localization accuracy is not high. Gabor wavelet coefficient can provide rich local-texture feature information and thus leads to accurate feature point localization, but with the drawback that it may add computation complexity. Reference [7] divided facial feature points into fiducial points and contour points. In that paper, the author extracted Gabor feature for fiducial points, and gray feature for contour points. Its localization accuracy was not high because only the 1D local texture feature was extracted and its feature expression method was relatively single. Reference [8] used gray and Log-Gabor feature to express local texture feature together. Though the author combined 2 different kinds of texture feature, he still extracted 1D feature and could not ensure high localization accuracy. In addition, extracting Log-Gabor feature for all the feature points added computation complexity.

According to the basic idea of multi-resolution pyramid searching strategy of the original ASM, in the upper levels of the pyramid, only estimated values for model fitting are determined. At the lowest level of the pyramid, the final model fitting values for the original images are computed. That is to say, a coarse-to-fine searching strategy is adopted. In the upper levels of the pyramid, the feature points are coarsely localized .And at the lowest level of the pyramid, the feature points are finely localized. So, we adopt the following two steps to build local texture feature model.

The first step: Extraction 2D gradient feature – The original ASM algorithm always builds local gray statistical property of feature points' neighbor area .That is to say, the gray-level model is the normalized derivative of the profiles sampled perpendicular to the feature point contour and centered at the feature point. Experimental results show that to obtain sufficient statistical information needs a large number of training images if we extract gray feature as local feature. In addition, gray feature's robustness to illumination is worse than gradient feature [8]. To reduce the influence of illumination to localization accuracy and reduce the number of training images, this paper proposes to extract 2D gradient feature on the highest level and the higher level of multi-resolution pyramid images, because gray feature is very sensitive to illumination and the number of training images is limited.

The details of profile extension from 1D to 2D can be seen in reference [8].A 2D gradient profile at a feature point is created in the following three steps: convolution, normalization, equalization [9]. These three steps are repeated at each feature point.

■Convolution

In the convolution step, we move a 3×3 gradient mask over each pixel in the area around the feature point to generate a profile matrix. Each element of the profile matrix corresponds to a pixel. The gradient mask we used is as follows:

$$A=\begin{bmatrix}0 & 0 & 0\\ 0 & -2 & 1\\ 0 & 1 & 0\end{bmatrix}. \tag{14}$$

We can get the gradient profile matrix by convolution. The profile matrix is two dimensional. We treat the profile matrix as a single long vector to get its mean value and covariance matrix.

■Normalization

To eliminate the influence of illumination and contrast, we normalize the profile matrix by dividing it by the absolute sum of its elements.

■Equalization

We equalize the normalized profile matrix by applying a sigmoid transform to each element x of the matrix. More details about the sigmoid transform can be seen in reference [8].

The second step: Extraction 2D Gabor feature – According to a coarse-to-fine searching strategy, this paper proposes to use Gabor wavelet transformation to extract the lowest level's 2D texture feature of each feature point because Gabor wavelet coefficients contain rich local-texture feature information.

Gabor wavelet transformation has good space local character and orientation character. It can extract space frequency and local feature of muti-orientation in a local image area. Two dimension Gabor wavelet is shown as the following formula.

$$W(x,y,\theta,\lambda,\varphi,\sigma,\gamma)=e^{-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}}\cos(2\pi\frac{x'}{\lambda}+\varphi) \tag{15}$$

where $x'=x\cos\theta+y\sin\theta, y'=-x\sin\theta+y\cos\theta$.

From (15), we can see that 2D Gabor wavelet is the convolution result of triangle function and Gaussian function. Equation (15) has five parameters which can control the 2D Gabor wavelet. The following will discuss these five parameters' functions.

θ controls the orientation of Gabor wavelet, which makes the wavelet circumrotate around its center.

λ controls the wavelength of triangle function, which determines the frequency of Gabor wavelet.

φ controls the phrase of Gabor wavelet, which is chosen as 0 and $\pi/2$ usually.

σ controls the radius of Gaussian function, which determines the size of Gabor wavelet. It is usually proportional to wavelength.

γ controls the length and width ration of Gabor kernel.

Different parameters can generate Gabor wavelet with different shapes. We adopt five wavelengths and eight orientations in this paper, which is shown as the following formula:

$$\begin{cases} \lambda = 2^{\frac{4+i}{2}}, i = 0 \sim 4 \\ \sigma = \dfrac{3\lambda}{4} \\ \theta = \dfrac{i\pi}{8}, i = 0 \sim 7 \\ \varphi \in \{\pi/4, -\pi/4\} \\ \gamma = 1 \end{cases} \tag{16}$$

By this way, 40 Gabor coefficients in the complex form are used to represent the pixel and its vicinity. Specifically, a jet vector J is shown as follows:

$$\begin{aligned} Jet &= (a_1, \phi_1, a_2, \phi_2, ...a_{40}, \phi_{40}) \\ J_j &= a_j \exp(i\phi_j) \end{aligned}. \tag{17}$$

Where a_j and ϕ_j are the magnitude and phase of the jth Gabor coefficient. a_j varies slowly, and ϕ_j varies quickly.

We suppose there exist two points P and P' with relative small displacement, whose Gabor Jets are $\vec{J}, \vec{J}'$.The similarity between two jet vectors is measured with the following two functions:

$$s_a(\vec{J}, \vec{J}') = \frac{\sum_j a_j a_j'}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}} \tag{18}$$

$$s_\phi(\vec{J}, \vec{J}') = \frac{\sum_j a_j a_j' \cos(\phi_j - \phi_j' - \vec{d}\vec{k}_j)}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}} \tag{19}$$

where $\vec{k} = \begin{bmatrix} \dfrac{2\pi\cos\theta}{\lambda} \\ \dfrac{2\pi\sin\theta}{\lambda} \end{bmatrix}$.

Phase-insensitive distance function (18) doesn't include phase parameter. If there are two points with small displacement, then their phase-insensitive distance is large. A range difference of the two jets can be determined. Phase-sensitive distance function (19) includes phase difference. So, if there are two points with small displacement, then their phases vary largely. So, phase-sensitive distance function can measure difference of two points and can accurately localize feature points. So, we choose phase-sensitive distance to measure the similarity of two points in this paper. The displacement between two locations with small displacement can be approximately estimated by maximizing the phase-sensitive distance $s_\phi(\vec{J}, \vec{J}')$ in (19).If we want to maximize $s_\phi(\vec{J}, \vec{J}')$,we can suppose

$\dfrac{\partial s_\phi}{\partial d_x} = \dfrac{\partial s_\phi}{\partial d_y} = 0$,then

$$d(\vec{J}, \vec{J}') = \begin{pmatrix} d_x \\ d_y \end{pmatrix} = \frac{1}{\Gamma_{xx}\Gamma_{yy} - \Gamma_{xy}\Gamma_{yx}} \begin{pmatrix} \Gamma_{yy} & -\Gamma_{yx} \\ -\Gamma_{xy} & \Gamma_{xx} \end{pmatrix} \begin{pmatrix} \Phi_x \\ \Phi_y \end{pmatrix} \tag{20}$$

If $\Gamma_{xx}\Gamma_{yy} - \Gamma_{xy}\Gamma_{yx} \neq 0$, where

$\Phi_x = \sum_j a_j a_j' k_{jx}(\phi_j - \phi_j')$,

$\Phi_y = \sum_j a_j a_j' k_{jy}(\phi_j - \phi_j')$, $\Gamma_{xy} = \sum_j a_j a_j' k_{jx} k_{jy}$.

and $\Gamma_{xx}, \Gamma_{yx}, \Gamma_{yy}$ are defined accordingly.

Extracting 2D Gabor texture feature may add the computation complexity, but we only extract this kind of feature at the lowest level of pyramid images, so the computation complexity is not very high.

We can localize feature points accurately in the case that the computation complexity is not very high, since we effectively combine 2 different kinds of texture feature.

C. A Modified Multi-Resolution Pyramid Decomposition Method

At the training and searching stage of ASM on muti-resolution frame, original ASM algorithm adopts the following method to get pyramid images of training image. First, let training image pass a Gaussian filter, and then sample the filtered image in interval. The resolution of these images decreases gradually. The size of the latter scale image is half of that of the former scale image. The pyramid sub-images obtained by this method, whose fine feature will get less and less as the resolution decreases. So, it counts against extracting rich texture feature and influences the feature localization accuracy.

Wavelet transformation has multi-resolution character. It can extract rich information from images effectively and realize adaptation analysis of images. Using wavelet transformation to build pyramid images has many advantages [10].

- Wavelet decomposition is non-loss focus change. Its direct frequency is smooth, which is good to implement image correlation.
- Each layer of wavelet decomposition is constituted by a direct frequency weight and three detail weights. So we

can implement character match using these character weights.

■Wavelet transformation has Mallat quick algorithm, thus its computation complexity is low and computation speed is fast.

When we choose wavelet decomposition to implement character match, we should consider the following problems: choice of wavelet function, decomposition level L.

■Choice of wavelet function

The chosen wavelet function should satisfy the following conditions. ①Information of original image should be protected as much as possible in low frequency image. ② Computation formula should be as simple as possible. It's better to have corresponding quick algorithm to implement using hardware.

■Determination of decomposition level L

Determination of decomposition level is relevant to the size of template image. Experimental and statistic results show that the size of template image shouldn't be small than 8×8 when correlation match is implemented. When it is small than this size, its provided quantity of information is limited. Thus it is easy to get error matching result. In addition, the decomposition level shouldn't be large. Many times low-pass filter will also lose information of original image, thus make matching result on higher level incredible. Usually, the maximum of decomposition level is $3\sim 5$ .So, the value of decomposition level should satisfy the following formula:

$$L < \min\left\{\log_2{}^{N/8}, 5\right\} \qquad (21)$$

where N is the minimum of the width and height of template image.

According to the above rules and the size of training and testing images, this paper uses three levels to decompose the training and testing images. The adopted wavelet function is db1 wavelet. The decomposition results are shown in Fig.3.In Fig.3, picture (a) is the original ASM's decomposition result, and picture (b) is our method's decomposition result. In picture (a) and (b), there are the first level (the original image), the second level and the third level pyramid sub-images from left to right. We can clearly see that, the pyramid sub-images obtained by wavelet transformation contain richer information than that of original ASM as the resolution decrease gradually. Rich fine-information is good to extract rich local texture-feature and build stable texture-profile model, and thus increases the accuracy of feature points localization.

(a) Decomposition result of original ASM

(b) Decomposition result of our proposed method

Figure 3. Multi-resolution pyramid sub-images

D. A Modified Pyramid Searching Strategy

In traditional ASM algorithm, the profile searching length is kept the same for different pyramid levels when multi-resolution scheme is adopted. In this paper, we set the first level (the lowest level) as the original image, the second level (the higher level) and the third level (the highest level) as pyramid sub-images obtained by wavelet transformation. At the multi-resolution searching stage, we first search feature points at the lower resolution, and then search at the higher resolution gradually. Experimental results show that displacement of feature point varies from different pyramid levels [11]. The lower the level is, the smaller the displacement is. Because we adopt a coarse-to-fine searching strategy, the displacement is very small when the new position of the feature point is near the target position. So, in our proposed algorithm, we change the length of the 2D profile searching-rectangle according to different pyramid levels. At the low resolution, we can get ideal match result of the whole contour because we ignore some details of image and our searching range is relatively wide. Correspondingly, at the high resolution, we may get good match result of local feature because we reduce our searching range. In our proposed algorithm, the length of profile searching-rectangle is reduced adaptively according to the decrease of the level. The value of the displacement of the first level is constrained. In our experiment, we set the profile length of the third level as m, m/2 of the second level, and m/4 of the first level. We set a threshold T for the displacement of the first level. If the value of the displacement is greater than T, it is set to be T.

This modified pyramid searching strategy strengthens the robustness of the algorithm, and the efficiency of the algorithm is improved to a certain extent because the computation complexity is relatively low at the low resolution.

At the searching stage on each pyramid layer, we need an algorithm to determine when we should search at the higher resolution level and when we should stop the whole searching process. In our experiment, we record the number of searched target points within the range of 50% of the 2D profile center. When the number of the points, which satisfy the condition, is 95% of the sum points (the number of marked feature points of the shape

model), we can deem the algorithm converges at that level. Then we can search at the next resolution level .We will not stop searching until it converges at the corresponding level. When the algorithm converges at the last level (level 0), we can stop the whole searching process.

Ⅳ. EXPERIMENTAL RESULTS AND ANALYSIS

Our experiment is based on IMM and ORL face database. We choose 200 images as training images. We choose 100 images as test images. All the images for training and test are neutral expression of frontal face. We mark 58 facial feature points for each training image manually. In this paper, we mark the feature points at the outer contour of face and the edges of each face component.

Usually, a large profile searching length is needed for robust initialization and good feature points localization. However, a too large profile searching length may make matching results worse because an inappropriate face shape can be considered as the original face shape. In addition, it may add feature points localization time if the profile searching length is too large. Based on above causes and according to the practical size of IMM and ORL face database' face images, we test over and over again and finally set the 2D profile searching-rectangle length is 4 of the first level, 8 of the second level, and 16 of the third level. The displacement threshold T is set to be 1 on the first level.

A. Performance Comparison

To evaluate the performance of our modified ASM algorithm, we use two evaluation functions [11]. The first is the average error function which is represented as the following formula.

$$E_{ave} = \frac{1}{N} \times \frac{1}{n} \sum_{i=1}^{N} \sum_{j=1}^{n} abs(x(i,j) - pos(i,j)) \quad (22)$$

where N is the number of test images, n is the number of feature points of each image. $x(i,j)$ is the manually marked position of the jth feature point on the ith image, and $pos(i,j)$ is position localized by the ASM algorithm.

The other one is average computation cost time function which is represented as follows.

$$\mathrm{T}_{ave} = \frac{1}{N} \sum_{i=1}^{N} t_i \quad (23)$$

where t_i is the cost time processing each image.

Tab.1 is the performance comparison result of our proposed ASM and the original ASM. From the comparison result, we can see that our proposed algorithm has greatly improved the localization accuracy and processing time compared with the original algorithm.

TABLE Ⅰ.
COMPARISON PERFORMANCE OF THE MODIFIED ASM AND ORIGINAL ASM

Objective function	The modified ASM	Original ASM
E_{ave} (unit: pixel)	2.09	4.67
T_{ave} (unit:s)	1.87	3.45

B. Experimental Results

Parts of experimental results are shown in Fig.4.The group (a) is our proposed ASM algorithm's localization results, and group (b) is the localization results of the original ASM.Obviously, we can see that our proposed algorithm has made great improvement on localization accuracy than that of the original algorithm.

(a) Localization results of our proposed ASM

(b) Localization results of the original ASM

Figure 4. Localization results of ASM

To evaluate the robustness of our proposed ASM algorithm, we choose several images whose faces have small angular deflection and make experiments. Parts of experimental results are shown in Fig.5. From Fig.5, we can see that our proposed algorithm may fail to locate when faces have small angular deflection.

Figure 5. Some localization failure results

Ⅴ. CONCLUSION

In this paper, a modified active shape model is proposed to improve the performance of the original active shape model. The proposed algorithm includes the following four aspects. Firstly, this paper adopts a semi-automatic facial feature points marking tool. Secondly, this paper proposes to extract 2D gradient feature on the highest level and the higher level of multi-resolution pyramid images, and use Gabor wavelet transformation to extract the lowest level's 2D texture feature. Thirdly, this paper adopts a new method of decomposition of multi-resolution pyramid. Finally, this paper uses an improved searching scheme of multi-resolution pyramid.

Experimental results show that our proposed ASM algorithm has made grate improvement on localization accuracy and processing time than that of the original ASM.However, because the training images are neutral expression of frontal face, so the experimental results also show that our proposed algorithm may fail when faces have small angular deflection. How to make the proposed algorithm more robust of face's angular deflection is an important direction for our future study.

REFERENCES

[1] Kwok-Wai Wan, Kin-Man Lam,Kit-Chong Ng,"An accurate active shape model for facial feature extraction", Pattern Recognition Letters ,vol.26,pp. 2409–2423,May 2005.

[2] Du Cheng,Su Guangda,Lin Xinggang ,"A modified muti-scale active shape model for face align",Journal of Optoelectronics Laser,vol.15,pp. 706-709,June 2004.

[3] Fan Yuhua, Ma Jianwei,"ASM and its modified facial feature localization algorithm", Journal of Computer-Aided Design & Computer Graphics, vol.19, pp.1411-1415, November 2007.

[4] Liu Aiping, Zhou Yan, Guan Xipu,"Improved ASM method of face localization", Computer Engineer, vol.33, pp. 227-229. June 2007.

[5] Wang Liting,Ding Xiaoqing,Fang Chi,"A robust automatic facial feature localization method",Acta Automation Sinica ,vol.35,pp.9-16, January 2009.

[6] T.F.Cootes, C.J.Taylor, D.H.Cooper, J.Graham, Active shape models-their training and application, Computer Vision and Image Understanding 61(1995)38–59.

[7] Yan Tong,YangWang,Zhiwei Zhu ,Qiang Ji,"Robust facial feature tracking under varying face pose and facial expression", Pattern Recognition, vol.40,pp. 3195 – 3208,July 2007.

[8] Fu Yunshu, Fu Zhonghua, Zhang Yanning,"Facial feature extraction based on a modified active shape model", Application Research Of Computers, vol.11, pp. 255-257, May 2006.

[9] Stephen Milborrow. Locating facial features with active shape models. Master paper of Faculty of Engineering, University of Cape Town, 2007.11.

[10] Jingmin Zhang , Zhijia Zhang, Dongshu Wang, "A pyramid rapid image matching algorithm based on wavelet decomposition", Information Sciences,vol.17,pp. 1724–1730, January 2007.

[11] Zhonglong Zheng ,Jia Jiong ,Duanmu Chunjiang ,XinHong Liu ,Jie Yang,"Facial feature localization based on an improved active shape model", Information Sciences,vol.17,pp. 2215–2223,August 2008.

[12] Andreas Koschan ,Sangkyu Kang, Joonki Paik ,Besma Abidi,Mongi Abidi,"Color active shape models for tracking non-rigid objects",Pattern Recognition Letters ,vol.24,pp.1751-1765,August 2003.

[13] Van.Ginneken.B, Frangi, A.F, Staal, J.J.Active shape model segmentation with optimal features.IEEE Transactions on Medical Imaging, 2002, 21(8):924-933.

[14] Chengzhi Sun, Mei Xie.An enhanced active shape model for facial features extraction. The eleventh IEEE International Conference on Communication Technology, Nov, 2008,661-664.

[15] Shiguang Shan, Wen Gao, Wei Wang, Debin Zhao.Enhanced active shape models with global texture constraints for image analysis. The 14th International Symposium on Methodologies for Intelligent Systems, 2003, 593-597.

[16] Z.Zhao, E.K.Teoh, A new scheme for automated 3D PDM construction using deformable models, Image and Vision Computing 26(2) (2008)275–288.

[17] Stephen Milborrow and Fred Nicolls.Locating facial features with an extended active shape model. www.milbo.users.sonic.net.

[18] Cootes, T.F., Taylor, C.J., 1992.Active shape models.In: Third British Machine Vision Conf., pp.266–275.

[19] Cootes,T.F.,Taylor,C.J.,Lanitis,A.,Cooper,D.H.,Graham,J., 1993.Building and using flexible models incorporating grey-level information.In:Fourth Internat.Conf.Computer Vision.pp.242–246.

[20] Cootes,T.F.,Hill,A.,Taylor,C.J.,Haslam,J.,1994.The use of active shape models for locating structures in medical images.Image Vision Comput.12(6),355–366.

[21] Froba, B., Kastner, T., Zink, W., Kublbeck, C., 2001.Real time active shape models for face segmentation.In:Internat.Conf.on Image Processing.pp.205–208.

[22] Ginneken,B.,Frangi,A.,Staal,J.,Romeny,B.,Viergever,M.,2001.A nonlinear gray-level appearance model improves active shape model segmentation.In:Proc.Mathematical Methods in Biomedical Image Analysis.pp.205–212.

[23] C.Davatzikos, X.Tao, D.Shen, Hierarchical active shape models, using the wavelet transform, IEEE Transactions on Medical Imaging 22(3) (2003)414–423.

[24] G.Hamarneh, T.Gustavsson, Deformable spatio temporal shape models: extending active shape models to 2D+time, Image and Vision Computing 22(2004)461–470.

[25] W.Wang,Shiguang Shan,Wen Gao,Bo Cao,An improved active shape model for face alignment,in:The Fourth International Conference on Multi-modal Interface,Pittsburgh,USA,2002,pp.523–528.

[26] S.C.Yan,C.Liu,S.Z.Li,H.J.Zhang,H.Y.Shum,Q.S.Cheng,Face alignment using texture-constrained active shape models, Image and Vision Computing 12(2003)69–75.

[27] M.Rogers, J.Graham, Robust active shape model search, Proceedingsof ECCV, vol.4, 2002, pp.517–530.

[28] S.Yan,X.Hou,S.Z.Li,H.Zhang,Q.Cheng,Face alignment using view-based direct appearance models, Special issue on facial image processing, analysis and synthesis,Int.J.Imaging Syst.Technol.13(1) (2003)106–112.

[29] Y.Tian,T.Kanade,J.F.Cohn,Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity, Proceedings of FGR02,2002,pp.218–223.

[30] F.Jiao,S.Z.Li,H.Y.Shum,D.Schuurmans,Face alignment using statistical models and wavelet features, Proceedings of CVPR03,vol.1, 2003,pp.321–327.

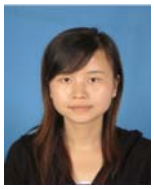
Min.Wang was born in Chang Zhou city of Jiang Su province in Oct, 1959. He obtained a B.S.degree of communication engineering in the Department of The School of Information and Control Engineering, Xi'an University of Architecture & Technology, China in June 1982.

He worked at iron and steel company in Tai Yuan in 1975. He is currently an associate director of communication engineering scientific research center in the Department of The School of Information and Control Engineering, Xi'an University of Architecture & Technology. He has published over 30 papers in peer-reviewed journals and

conferences. His current research interest focuses on multimedia communication, digital homestead and digital amusement, intelligent information processing.

Prof.Wang is currently a member of Chinese Association of Automation. He was also a member of Image and Graphic Academy of Shan'Xi.

Yiling.Wen was born in Han Zhong city of Shan'Xi province in Sep, 1985. She is currently pursuing her M.S.degree of signal and information processing in the Department of The School of Information and Control Engineering, Xi'an University of Architecture & Technology. Her research interest is image processing. She obtained a B.S.degree of Electron Information Engineering in the Department of The School of Information and Control Engineering, Xi'an University of Architecture & Technology, China in June 2008.



Li.Fang was born in Hou Ma city of Shan'Xi province in Dec, 1984. She is currently pursuing her M.S.degree of computer application technology in the Department of The School of Information and Control Engineering, Xi'an University of Architecture & Technology. Her research interest is graphics and image processing. She obtained a B.S.degree of Information and Computation Science from Department of Computer Science and Technology, Shaanxi University of Technology, China in June 2006.

Weiping.Sun was born in Shang Nuo city of Shan'Xi province in Jan, 1984. She is currently pursuing her M.S.degree of signal and information processing in the Department of The School of Information and Control Engineering, Xi'an University of Architecture & Technology. Her research interest is image processing. She obtained a B.S.degree of Electron Information Engineering in the Department of Computer Technology, northwestern polytechnic University, China in June 2007.

