



An Improved Algorithm for Facial Feature Location by Multi-template ASM

Li Benfu^(✉)

Southern Medical University, Guangzhou, China
da3831182neixi@163.com

Abstract. In order to improve the accuracy of the Shape Model Active method, we propose a new method to improve the accuracy of ASM (ASM) algorithm in face detection, and propose a new method to construct the local template. In the process of local localization, the paper uses form Closed-algorithm to segment the texture segmentation. Information is effectively improved the performance of the ASM method. The results show that the proposed algorithm can extract the feature points of most forward faces correctly. The proposed algorithm has a wide range of applications in image understanding of face tracking, recognition and facial expression analysis.

Keywords: Active shape model (ASM) · Multi-template · Narrow strip-map Closed-form algorithm

The active shape model (active shape model, ASM) algorithm [1] is one of the active vision algorithms. Because of its flexible adaptation to target contour, and to some extent, it can be used to detect the flexible object. It is the main face feature detection algorithm. The ASM algorithm [2] is based on the original ASM algorithm, and then uses the local template to achieve global and local optimization. The point Prof (ILE) is used to match the pixel value gradient information and training template. Therefore, it is difficult to make accurate segmentation of the region of the edge of the feature, such as the fuzzy edge region of the jaw and neck region or other features. For this problem, the Liuai ping et al. [3] uses the discrete cosine transform to model the feature points of the training image, and makes full use of the 2D texture information near the feature points; Yuhua et al. [4] use Gabor transform to model the local texture around the feature points and improve the robustness of the algorithm to illumination and noise; Li [5] et al. In the training process of multi template ASM algorithm, the Gabor feature information is used in the same direction as the feature edge, and the different states of the eye and mouth are set up respectively; Cristinacce et al. [6] The improved ASM algorithm is used to improve the search efficiency of the model. The improved model can improve the accuracy of the feature point location and reduce the influence of illumination and noise. However, the texture feature information is not rich, so they have little effect on the detection accuracy. Toth et al. [7] is used to implement the segmentation of magnetic resonance image(magnetic resonance image, MRI) first, and then use the ASM algorithm to extract the feature contour, which is suitable for the segmentation of the texture features of the smooth region, but only for the segmentation of MRI.

In this paper, we improve the traditional ASM algorithm for locating face image texture smoothing region. The 1 global templates and 7 local templates are established. The feature points are located near the feature region. Then, the -form Closed algorithm is used to segment the Profile vertices. Matching Because the -form Closed algorithm has the characteristics of smooth region segmentation, the feature edge of the smooth region is highlighted, which can improve the efficiency and accuracy of the algorithm.

1 Closed-Form Algorithm Analysis

Closed-form algorithm proposed by [8] and other Levin can be used to solve the problem of image segmentation. The algorithm is described as the I of any image, the Ii of any of its pixels are composed of foreground and background, and the proportion of the foreground pixels is α_i , that is $I_i = \alpha_i F_i + (1 - \alpha_i) B_i$. Thus available $\alpha_i \approx \alpha_i I_i + b_i, \forall i \in W$; Among them $\alpha_i = 1/(F_i - B_i), b_i = -B_i/(F_i - B_i)$, W for the image window. By minimizing

$$J(\alpha, a, b) = \sum_{j \in I} \left(\sum_{i \in W_j} (\alpha_i - a_j I_i - b_j)^2 + \epsilon a_j^2 \right) \quad (1)$$

Can be obtained from the ratio of the foreground pixels A, Wj for the first j pixels of the image window.

By the segmentation results and the analysis of the [8], the Closed-form algorithm can be seen in the segmentation of the image smooth part. By the segmentation results and the analysis of the [8], the Closed-form algorithm can be seen in the segmentation of the image smooth part, (1) before the segmentation, the user interaction is needed to draw the background and foreground of the graph Fig. 1b (white as the foreground, black background), and the automation can not be completed; (2) Because of each pixel are required in to the center of the image window W formula (1), so the computing complexity is high, the high resolution image computation for a long time; (3) the sensitivity of the foreground and background edge noise is high, except for the special setting parameters, it is easy to produce burr or hole in the edge.



Fig. 1. Form-closed algorithm partition diagram

To solve the first problems, this paper applies the ASM feature extraction to find the 2 point (Profile) as the initial marker, which is not required to be used as a result of the second questions, which have little effect on the computation time. Because the ASM can maintain a certain topology, it can better resist the segmentation edge. Based on the above 3 points, we can effectively integrate the -form Closed algorithm into ASM.

2 The Establishment of Multi-template ASM

In this paper, a multi-template is made up of 1 global templates and 7 local templates. The global template is used to locate all the feature points of the face image, and the local template is used to locate the local features after global positioning.

2.1 Selection of Feature Points of Multi-template ASM

The feature points are selected from the-3 CANDIDE model [9], which is used to establish the global template, which is based on the global template, which is based on the global template, which can be used to establish the local template, which is based on Fig. 2b. A feature point is formed between the 2 feature points on the line, and the local template is formed by 100 C, which can maintain the information of the feature points between the global template and the local template, and enhance the prior information of the feature points.

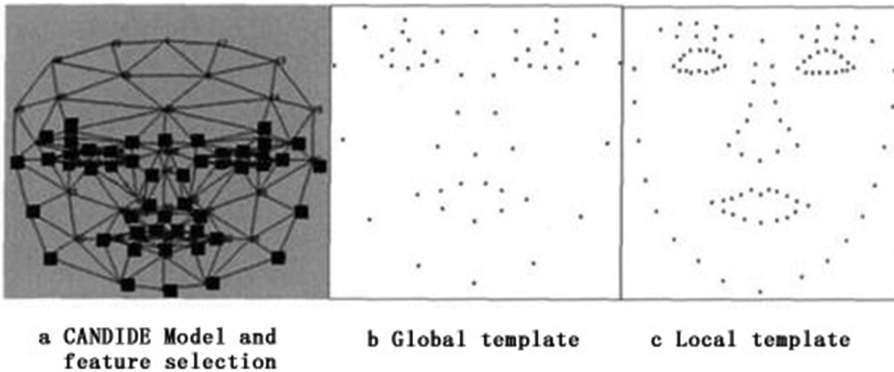


Fig. 2. ASM template

2.2 Feature Points of the ASM Feature Points

According to Fig. 2b, the template of the Fig. 2c to the corresponding position of the training image is manual calibration, and the training image after the following standard processing: (1) Eliminate the bad quality of the image; (2) Select the face relative to the lens deflection is not more than 10b of the image; (3) Select the face image without wearing glasses, earrings and other accessories; (4) segmentation process, using the image of the training only to retain the face, neck and other background are

set to 0, so as to better achieve the model and narrow band segmentation area on the Profile pixel matching.

After manual calibration of the image as shown in Fig. 3c, d, the global template feature points are obtained $S_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{iN}, y_{iN})^T$ and local template feature points $S_{LOCAL_i} = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{i2N}, y_{i2N})^T, i = 1, 2, \dots, M$. Among them M, where I represents the first I image, N indicates the number of feature points. N indicates the number of feature points. Divide the feature points of the 7 into 7 regions S_1, S_2, \dots, S_7 according to the feature points of the face in $S_{LOCAL_i}, i = 1, 2, \dots, M$ regions. Respectively represent the left and right eyebrow, right and left eyes, nose, lips and facial contour. Through principal component analysis, to obtain the shape of the global template and seven local template description (feature sub space). Unified mind $M = \bar{S} + Pb$, among them, \bar{S} is the average shape of the template, $P = (p_1, p_2, \dots, p_t)$ is a set of template shape feature subspace, $b = (b_1, b_2, \dots, b_t)$ is a shape parameter, different B corresponding to different shapes.

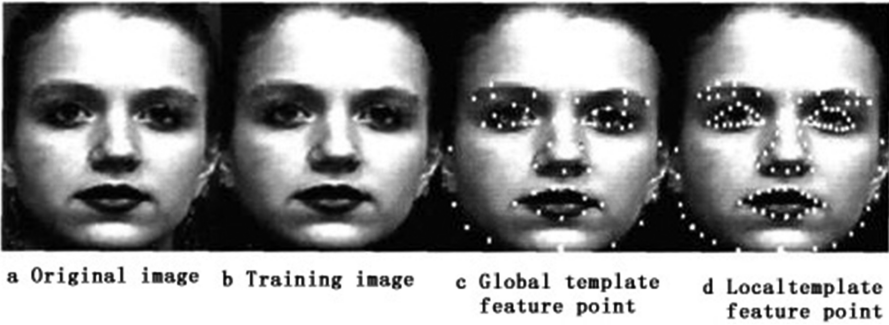


Fig. 3. Original images and training images

Because the shape description is not always consistent with the true shape, it is required to obtain the local gray level information near the training process in order to adjust the shape during the matching process. Take the $2K + 1$ gray level sample $G_{ij} = [g_{ij0}, g_{ij1}, \dots, g_{ij(2K)}]$ on the j Profile of the image i. The number of pixels selected from the K (K = 3) to the normal or outward direction of the normal, G_{ij} for differential, $dg_{ij} = (g_{ij2} - g_{ij1}, g_{ij4} - g_{ij3}, \dots, g_{ij(2K+1)} - g_{ij(2K)})^T$, and its standardization $y_{ij} = dg_{ij} / \left(\sum_{l=1}^{2K} |dg_{ijl}| \right)$.

Calculate the gray difference mean and covariance matrix of the j feature points $\bar{g}_j = \frac{1}{N} \sum_{i=1}^N y_{ij}, \sum_j = \frac{1}{N} \sum_{i=1}^N (y_{ij} - \bar{y}_j)(y_{ij} - \bar{y}_j)^T$, the mean \bar{g}_j and covariance matrix \sum_j are preserved as the training results of this point.

3 Improved Multi-template ASM Algorithm

Specific steps of the algorithm as shown in Fig. 4. First of all to face feature using global template gives an overall positioning, again on the feature regions using template partial local location. In the process of localization, first find out where the profile vertex of the template feature points of a narrow zone; then will all feature points according to the order of the head and tail connected form a polygon, point to the inside the polygon normal to normal with profile, the profile vertex as a seed point initialization position; then with profile pointing outside the polygon normal to normal, the p As the initial position of the background seed point, the Closed-form algorithm is used to segment the foreground pixels in the narrow strip region and the background pixels are 0. Finally, the training template and the segmented images are matched. If the convergence condition is satisfied, the narrow strip is constructed by the new feature points Profile.

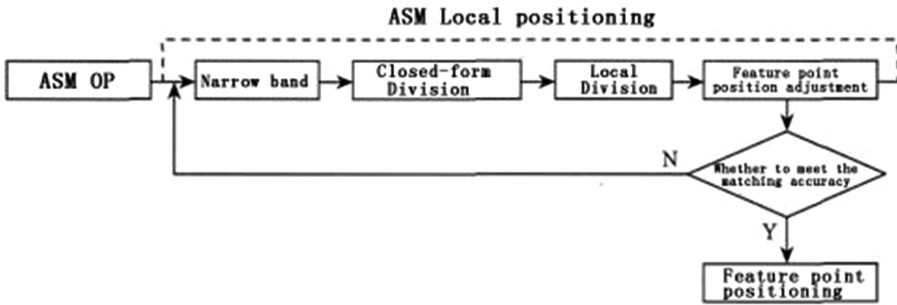


Fig. 4. Algorithm implementation block diagram

3.1 Overall Positioning

First overall positioning, so that local positioning stage of regional features of foreground and background markers initialization is more accurate and rational. First the ASM to initialize the location, according to the center of the left eye, right eye and lip center 3 respectively form model of triangle positioning and image positioning triangles, as shown in Fig. 5. The triangles of the model and the target image of the triangle, using triangular geometric features will align the model triangle and the triangle image, model is placed in the most close to the position of the target image to complete the initialization steps, then ASM overall positioning.

ASM overall positioning process is shown in Fig. 6. In Fig. 6, square as feature points, dots to the midpoint of the characteristic line, dotted lines feature points on the line for the profile of the i feature points on the $2K + 1$ gray sampling $G_{ij} = [g_{ij0}, g_{ij1}, \dots, g_{ij(2K)}]$. K is the number of pixels selected from the normal or the outward direction of the normal. The 3–4 is usually the one that affects the topological properties of the model. The triangle at the top of the dotted line is a Profile vertex, which is used to construct the narrow band and is used as a marker of the foreground and background. The I of the Profile of the feature points on the gray level sampling $G_{ij} = [g_{ij0}, g_{ij1}, \dots, g_{ij(2K)}]$ for differential and normalized by $c_i = [c_{i0}, c_{i1}, \dots, c_{i(2K)}]$,

By minimizing the objective function $f(c_i) = (c_i - \bar{g}_i)^T - \sum_i^{-1}(c_i - \bar{g}_i)$, the best matching points are obtained by minimizing the feature points on the Profile. In Fig. 6, the cross point is searched to the best matching point, where \sum_i^{-1} is the inverse of the gray covariance matrix of the I feature points obtained from the training set, and GI is the I of the Profile feature points.

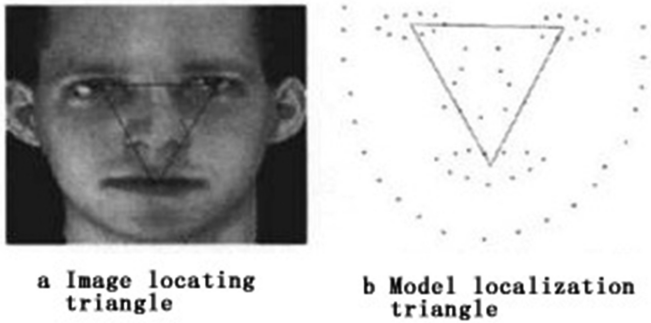


Fig. 5. Triangle registration

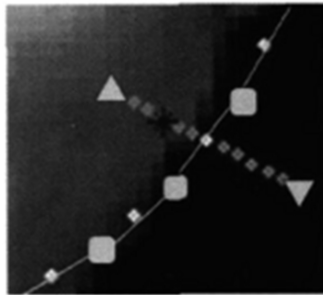


Fig. 6. Positioning process map

3.2 Local Positioning

Overall positioning is a global optimization, when the facial images in the presence of multiple feature details, some local characteristics of regional feature points may due to the role of global optimization and to the lack of accurate positioning, so after a global template search and location can be localized on the basis of global localization, localization process such as dashed lines in Fig. 4 of the box part.

3.2.1 Narrow Band Structure and -Form Closed Segmentation

For Fig. 4 in the narrow strip belt construction such as shown in Fig. 7. Figure 7 cloud form graphics image into foreground, dot in the foreground edge feature points dotted lines for feature points of attachment, and dotted vertical arrow said ASM feature points in the midpoint method to that profile. Profile (figure of hexagon and hexagon) vertex as

a narrow strip of edge point, is connected to all the edge points, get surrounded by two solid narrow band region. In addition, in the foreground and background markers, because there have been overall positioning as a basis, feature point basic positioning in The assumption is that all normal directions point to feature points (Closed) in the -form Profile algorithm, which is based on the assumption that the foreground markers are in the foreground. Even though the above assumption is based on the feature points matching and adjustment, the Profile is not only good. With the target model matching, the narrow band will continue to be iterative tending to the hypothesis.

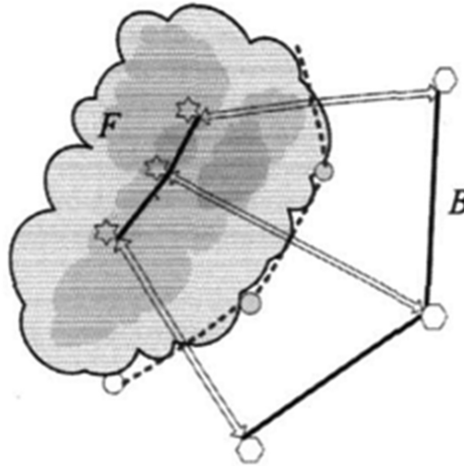


Fig. 7. The composition of the narrow strip

Then the -form Closed segmentation is done for the narrow strip region, pass type (1) solving

$$J(\alpha) = \min_{a,b} J(\alpha, a, b) \tag{2}$$

By type (2), the A, rewrite formula (1) is obtained for the least squares form of a, b that is

$$J(a, b) = \sum_{j \in I} \Pi^{G_j} \begin{bmatrix} a_j \\ b_j \end{bmatrix} - \bar{\alpha}_{j\Pi}^2 \tag{3}$$

One of the first j pixels of the image window W, there is

$$G_j = \begin{bmatrix} I_1 & I_2 & \cdots & I_i & \cdots & \sqrt{\varepsilon} \\ 1 & 1 & \cdots & 1 & \cdots & 0 \end{bmatrix}^T$$

$$\bar{\alpha}_j = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_i \quad \cdots \quad 0]^T, i \in W_j$$

By minimizing (3)a, b. See literature [8].

3.2.2 Local Matching Search

Local positioning algorithm using the local template serial introduction, and through the local matching search, adjust the overall positioning results, steps are as follows:

Step 1. For $j = 1$, the initial shape description of a local template is introduced into the J .

Step 2. Adjust the position, size and angle of the local template, so that it is consistent with the global template corresponding to the global template.

Step 3. The feature points are constructed from the local feature points, and the Closed-form algorithm is used to segment the feature region.

Step 4. Local matching of all the features of a local template for the first j , that is, the first j feature points of the template I , read into the training data (mean \bar{g}_{ij} and covariance matrix \sum_{ij}). Get the feature points corresponding to the Profile of the image gray level sampling, and then the difference is obtained $c_{ji} = [c_{ji0}, c_{ji1}, \dots, c_{ji(2K)}]$. By minimizing the objective function $f(c_{ji}) = (c_{ji} - \bar{g}_{ji})^T - \sum_{ji}^{-1} (c_{ji} - \bar{g}_{ji})$. Get the best matching point on the image, and then adjust the feature points to the best matching point.

Step 5. If the convergence condition is satisfied, the positioning result of the feature region is adjusted to the best matching point, and the next step is to the Step 3. If the convergence condition is satisfied, the positioning result of the feature region is adjusted to the best matching point, and the next step is to the Step 3.

Step 6. Checks if all local templates have been introduced, and if, in the next step, the initial shape of the $j + 1$ local template is introduced, otherwise, the Step 2.

Step 7. Will get the results in the overall positioning of the global template shape feature subspace M to find the optimal shape.

Step 8. Output of the new positioning results.

4 Experiment and Result Analysis

In this paper, the training images are derived from the ORL database. We select 200 face images as the training sample set, which is used to detect the remaining images and some standard video sequences from the ORL database. The size of the unified 100×100 experiments are carried out on the Matlab platform, Some simulation results and experimental data analysis are as follows.

4.1 Simulation Result Analysis

Figure 8 shows the algorithm simulation results, Fig. 8a shows the results of the overall positioning, can be seen, the point of the chin is not convergence in place; Fig. 8b, c is shown as a partial localization process results from Fig. 8c can be seen that the chin point has been converging to the edge of the face. Figure 8 shows the algorithm simulation results, Fig. 8a shows the results of the overall positioning, can be seen, the point of the chin is not convergence in place; Fig. 8b, c is shown as a partial localization process results from Fig. 8c can be seen that the chin point has been converging to the edge of the face.

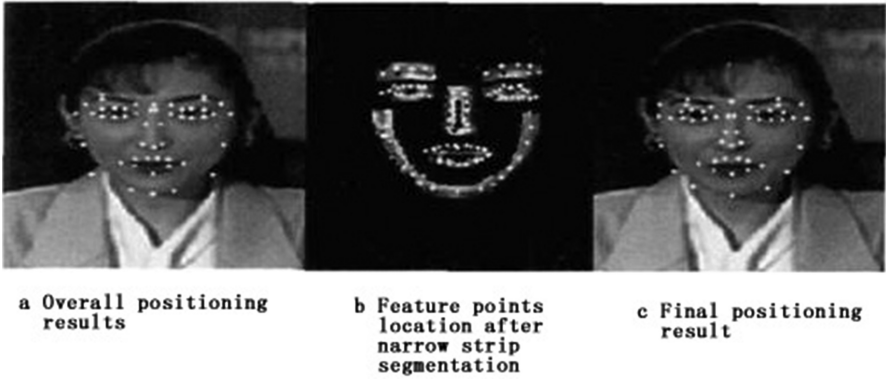


Fig. 8. Simulation results

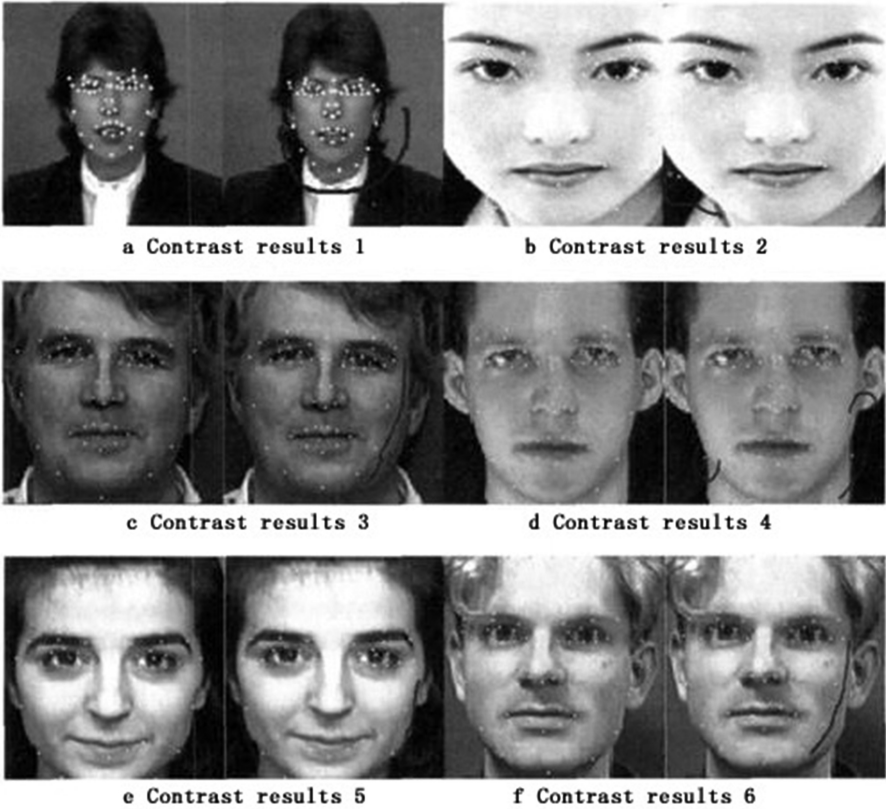


Fig. 9. Simulation results comparison chart

Figure 9 shows for the comparison of the results of this algorithm with the traditional ASM algorithm detection map. In the comparison results, first for the results of this algorithm and the second for traditional ASM algorithm results, in obvious error detection area are painted lines highlighted. By Fig. 9 shows that traditional ASM algorithm of feature point location in the image smooth regions (such as frowning face, chin) exist certain errors, and feature point location algorithm in this paper is the positioning of the closer to real characteristics of the region.

4.2 Comparison of Positioning Accuracy

This paper presents the algorithm to improve the accuracy of the positioning accuracy of the data, the accuracy of the measurement of the accuracy of the Euclidean distance of the average error $E = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{n} \sum_{j=1}^n \sqrt{(x_{ij} - x'_{ij})^2 + (y_{ij} - y'_{ij})^2} \right)$; Among them, N is the total number of test images, n is the number of feature points in a face image, (x_{ij}, y_{ij}) said that the coordinates of the j manual calibration points in the first I test image, (x'_{ij}, y'_{ij}) said that after the convergence of the algorithm.

In order to have a more direct comparison of the 2 algorithms, we further calculate the overall improvement of the algorithm in this paper with respect to the traditional ASM algorithm $I = \frac{E_{ASM} - E'_{ASM}}{E_{ASM}} \times 100\%$. Among them, E_{ASM} represents the average error of the traditional ASM algorithm, and the average error of the E'_{ASM} represents the algorithm.

Table 1 shows that the results of this paper are compared with the traditional ASM algorithm for the remaining 200 images of the ORL database. The average detection error of the traditional ASM algorithm is 15.12 pixels, and the average error of this algorithm is only 7.08 pixels, the improvement of the positioning accuracy is 53.17%.

Table 1. Accuracy comparison based on multi template localization algorithm

Mean error/Pixel		Improvement degree/%
The algorithm of this paper	Traditional ASM algorithm	
7.08	15.12	53.17

5 Concluding Remarks

This paper presents a face feature point localization algorithm based on narrow band Closed-form. The advantages of this algorithm are: (1) Closed-form algorithm is more accurate, so it can be used to segment the training image and to detect the smooth region, while the background pixels are 0, which is advantageous for matching calculation; (2) local template is built on the basis of global template. The construction of the narrow strip can reduce the computation time of the Closed-form algorithm. The experimental results show that the convergence accuracy and the feature points of the

image smoothing region can be improved by incorporating the Closed-form algorithm into the algorithm.

However, there are still some shortcomings in this paper. Firstly, if the overall positioning error is relatively large, it may lead to a narrow strip of Profile vertices, which can converge to the local optimal “false edge”. Secondly, the initial position of the foreground and background marker is not reasonable, and it is easy to lead to segmentation.

In addition, how to apply this algorithm to 3D specific human face model, and improve the similarity between the 3D model and the real human face, is also the subject of further research.

References

1. Cootes, T.F., Taylor, C.J., Lanitis, A.: Multi-resolution search with active shape models. In: Proceedings of the 12th International Conference on Pattern Recognition, Manchester, vol. 1, pp. 610–612 (1994)
2. Jin, W.: Study on Video-Based Face Expression Modeling. Zhejiang University, Hangzhou (2003). (in Chinese)
3. Aiping, L., Yan, Z., Xinpu, G.: Application of improved active shape model in ace positioning. *Comput. Eng.* **33**(18), 227–229 (2007). (in Chinese)
4. Yuhua, F., Jianwei, Ma.: ASM and improved algorithm for facial feature location. *J. Comput. Aided Des. Comput. Graph.* **19**(11), 1411–1415 (2007). (in Chinese)
5. Li, Y., Lai, J.H., Yuen, P.C.: Multi-template ASM method for feature points detection of facial image with diverse expressions. In: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, pp. 435–440. IEEE Computer Society Press, Washington D.C (2006)
6. Cristinacce, D., Cootes, T.: Boosted regression active shape models. In: Proceedings of British Machine Vision Conference, Warwick, vol. 2, pp. 880–889 (2007)
7. Toth, R., Tiwari, P., Rosen, M., et al.: A multi-modal prostate segmentation scheme by combining spectral clustering and active shape models. In: Proceedings of the SPIE, Bellingham: Society of Photo-Optical Instrumentation Engineers Press, vol. 6914, pp. 69144S.1–69144S.12 (2008)
8. Levin, A., Lischinski, D., Weiss, Y.: A closed form solution to natural image matting. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, vol. 1, pp. 61–68 (2006)
9. Ahlberg, J.: CANDIDE-3—an updated parameterized face. Linkping: Linkping University. Image Coding Group, Department of Electrical Engineering (2001)