

© [2008] IEEE. Reprinted, with permission from [Chunhua Du, Qiang Wu, Jie Yang, Zheng Wu, SVM based ASM for facial landmarks location, Proceedings of 2008 IEEE 8<sup>th</sup> International Conference on Computer and Information Technology, CIT 2008]. This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Technology, Sydney's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this document, you agree to all provisions of the copyright laws protecting it

# SVM based ASM for facial landmarks location

Chunhua Du<sup>1,2</sup>, Qiang Wu<sup>1</sup>, Jie Yang<sup>2</sup>, Zheng Wu<sup>2</sup>

<sup>1</sup>Department of Computer Systems, University of Technology, Sydney  
Sydney PO Box 123, Broadway 2007, Australia

<sup>2</sup>Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University  
Shanghai, 200240, P.R. China

dch3482275@163.com, wuq@it.uts.edu.au, jieyang@sjtu.edu.cn, sjtu\_wuzheng@hotmail.com

**Abstract**—Finding a new position for each landmark is a crucial step in active shape model (ASM). Mahalanobis distance minimization is used for this finding, provided there are enough training data such that the grey-level profiles for each landmark follow a multivariate Gaussian distribution. However, this condition could not be satisfied in most cases. In this paper, a new method support vector machine (SVM) based ASM (SVMBASM) is proposed. It approaches the finding task as a small sample size classification problem, and uses SVM classifier to deal with this problem. Moreover, considering imbalanced dataset which contains more negative instances (incorrect candidates for new position) than positive instances (correct candidates for new position), a multi-class classification framework is adopted. Performance evaluation on SJTU face database show that the proposed SVMBASM outperforms the original ASM in terms of the average error as well as the average frequency of convergence.

**Index Terms**—Active shape model (ASM), Support vector machine (SVM), Multi-class classification, Facial landmark, New position

## I. INTRODUCTION

Locating facial landmarks plays a very important role in face research such as face recognition, expression recognition and pose estimation. Approaches for the location of facial landmarks known to date can be divided into two categories—local methods and global methods. Local methods [1], [2], [3] individually detect eyes, mouth and other facial landmarks. Global methods such as deformable templates [4] and Snake [5], Active Shape Model (ASM) [6], Active Appearance Model (AAM) [7] and Lucas AAM [8], on the other hand, simultaneously locate more facial landmarks. Local methods can quickly locate facial landmarks. Nevertheless, the number of facial landmarks located by them is not sufficient enough for further high-level vision applications. In addition, local methods are very sensitive to the changes of lighting and face pose. In contrast to local methods, global methods are more robust. Furthermore, they can detect more facial landmarks. In terms of processing speed, global methods can also be compared with local methods. Thus global methods are more suitable for practical application. This paper focuses on global methods, particularly, on ASM, as it has been successfully used in many real applications such as medical image segmentation [9], hand contour segmentation [6] and location of facial feature [7].

It is well known that the most important step in ASM is finding a new position for each landmark at each iteration. The

accuracy of this finding heavily influences the overall accuracy of the location of landmarks. Normally, a new position for each landmark is located by computing all the Mahalanobis distance between the profiles centered at the candidate new positions and the local structure, the best new position for the landmark is the one with the minimal Mahalanobis distance [10]. The use of Mahalanobis distance minimization is under the assumption that enough training images are available such that the grey-level profiles of the landmark follow a multivariate Gaussian distribution. However, there could not be enough training images in practice. Thus it is very difficult to create enough profiles for each landmark, and the statistical model of profiles can't approximate multivariate Gaussian distribution very well. In this case, using Mahalanobis distance all the same is unreasonable. In this paper, support vector machine (SVM) [11] is proposed to replace the Mahalanobis distance minimization for the finding task. The finding task is coped with in a quite different way and is converted into a small sample size classification problem. SVM classifier is used for the task. Extensive experiments show that the proposed method significantly outperforms the original ASM in terms of average error and average frequency of convergence.

The rest of the paper is organized as follows. Section II and section III give a brief introduction of the original ASM and SVM, respectively. Section IV formulates the proposed SVMBASM in detail. Experimental results and discussion are presented in section V. It is concluded in section VI.

## II. ASM

### A. Construction ASM model

Assume there are  $n$  landmarked face images. On each image,  $k$  landmarks are marked by hand, as shown in Figure 1. The coordinates of all landmarks in a image give a shape vector

$$a_i = x_1^i, y_1^i, \dots, x_k^i, y_k^i, i = 1, 2, \dots, n \quad (1)$$

where  $(x_j^i, y_j^i)$  denote the coordinates of the  $j$ th landmark in the  $i$ th image. All shape vectors are aligned [10]. PCA is then applied on the aligned shape vectors to get the mean shape  $\bar{a}$ , eigenvalues  $(\lambda_1, \lambda_2, \dots, \lambda_q)$  and the corresponding eigenvectors  $(p_1, p_2, \dots, p_q)$ . The first  $t$  eigenvectors  $(p_1, p_2, \dots, p_t)$  are selected such that the shape vector can be

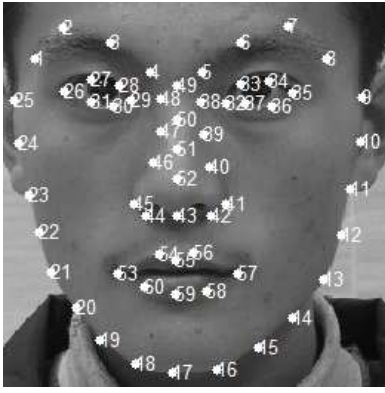


Fig. 1. Face image with marked facial landmarks

reconstructed by

$$a_i \approx \bar{a}_i + Pb \quad (2)$$

where  $b$  is a vector of  $t$  elements containing the model parameters.

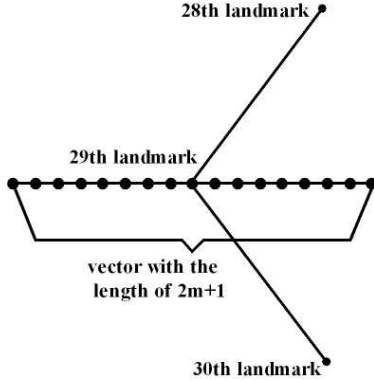


Fig. 2. Profile of the landmark for training

To locate new positions, local structure [10] for each landmark should be constructed in advance. For the  $i$ th landmark in the  $j$ th training image, as shown in Fig. 2, a grey-level vector is obtained by sampling  $m$  pixels on either side along the direction perpendicular to the line connecting its previous and next landmarks. In the original ASM, the normalized first order derivative of this grey-level vector is regarded as a profile  $g_{ij}$  for this landmark. The same operation is repeated for each training image. It thus gives  $n$  profiles  $(g_{i1}, g_{i2}, \dots, g_{in})$  for the  $i$ th landmark. The mean profile  $\bar{g}_i$  and covariance  $S_i$  of these profiles are then computed and treated as the local structure for this landmark. This is repeated for each landmark, giving one local structure for each landmark. Thus, the Mahalanobis distance between a new profile  $g_{iq}$  and the corresponding local structure is computed as follows:

$$f(g_{iq}) = (g_{iq} - \bar{g}_i) S_i^{-1} (g_{iq} - \bar{g}_i)^T \quad (3)$$

### B. Fitting ASM

At this stage, flexible shape model in equation (2) is used to find the new examples of the shapes in images. An example

of the shape in image is denoted by

$$X = M(s, \theta)[a_i] + T_c \quad (4)$$

where  $T_c = (X_c, Y_c, X_c, Y_c, \dots, X_c, Y_c)$ ,  $M(s, \theta)[\cdot]$  means rotation by  $\theta$ , scaling by  $s$ , and  $(X_c, Y_c)$  is the center of the shape in the image frame. In addition,  $(X_c, Y_c)$ ,  $\theta$ , and  $s$  are treated as pose parameters.

ASM fitting is realized by performing two iterative steps:

- 1) Finding a new position for each landmark;
- 2) Adjusting the shape and pose parameters;

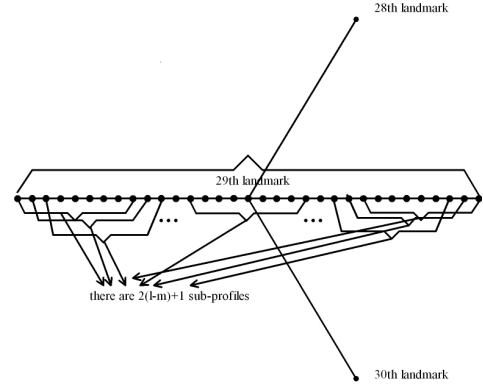


Fig. 3. Profile of the landmark for fitting

During each iteration, as shown in Fig. 3, for  $i$ th landmark,  $l$  ( $l > m$ ) pixels on either side are sampled to produce  $2(l-m)+1$  sub-profiles. The Mahalanobis distances between all these sub-profiles and the corresponding local structure are then computed. The center of the sub-profile with the minimal Mahalanobis distance is set as the new position for the current landmark, then a displacement  $dX_i$  for this landmark can be obtained. This step is repeated for each landmark, all displacements are then put together to form a displacement vector  $dX = (dX_1, dX_2, \dots, dX_k)$ . Then pose parameters displacements  $ds$ ,  $d\theta$ ,  $dX_c$  and  $dY_c$  can be obtained by aligning  $X$  to  $X + dX$ . Then shape parameter  $b$  can be updated as follows:

First equation (4) is extended to:

$$M(s(1+ds), (\theta+d\theta)) [a_i + da_i] + (X_c + dX_c) = (X + dX) \quad (5)$$

Since  $X$  can also take the form in equation (4), the above expression becomes:

$$M(s(1+ds), (\theta+d\theta)) [a_i + da_i] = M(s, \theta) [a_i] + dX + X_c - (X_c + dX_c) \quad (6)$$

Since

$$M^{-1}(s, \theta)[\cdot] = M(s^{-1}, -\theta)[\cdot] \quad (7)$$

Thus

$$da_i = M((s(1+ds))^{-1}, -(\theta+d\theta)) [M(s, \theta) [a_i] + dX - dX_c] - a_i \quad (8)$$

In the meantime, according equation (2),  $db$  is utilized such that

$$a_i + da_i \approx \bar{a} + P(b + db) \quad (9)$$

Substracting equation (9) from equation (6) gives

$$da_i \approx P \times db \quad (10)$$

Thus

$$db = P^{-1} da_i \quad (11)$$

since  $P^T = P^{-1}$  thus

$$db = P^T da_i \quad (12)$$

Combining equation (8) with equation (12),  $db$  can be obtained. The above arguments lead to the following sequence of adjusting:  $dX \rightarrow ds, d\theta, dX_c, dY_c \rightarrow dx \rightarrow db$ . At last, shape and pose parameters can be updated as follows:

$$\begin{aligned} X_c &= X_c + w_t dX_c, Y_c = Y_c + w_t dY_c, \\ \theta &= \theta + w_\theta d\theta, s = s(1 + w_s ds), b = b + w_b db \end{aligned} \quad (13)$$

where  $w_t, w_\theta, w_s$  are scalar weights.

With the new shape and pose parameters, new model shape can be obtained first by equation (2) followed by equation (4).

### III. SVM

SVM [11] is a very useful machine learning algorithm, which looks for an optimal hyperplane for pattern classification with the maximum margin, as shown in Fig. 4. SVM holds

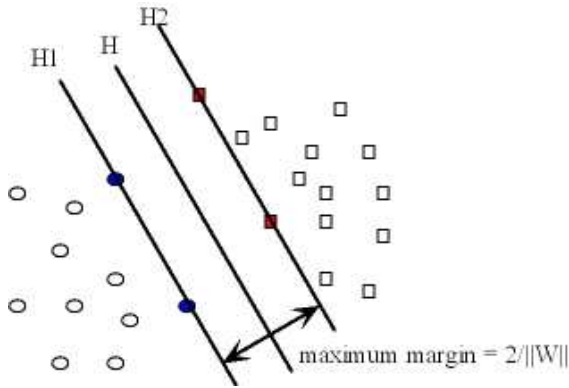


Fig. 4. SVM

several advantages [12] over other classification methods:

- 1) If a solution has been obtained, it is the unique solution;
- 2) Good generalization properties of the solution;
- 3) Common formulation for the class separable and the class non-separable problems as well as for linear and nonlinear problems;
- 4) Reasonable theoretical foundation based on structural risk minimization;

Because these advantages, SVM has been widely used in pattern recognition area [13][14]. A distinguishing characteristic of SVM is that the solution is determined by the pattern vectors closest to the hyperplane, where these vectors are called support vectors. This means that SVM does not

require many training data to train classifier. It thus can solve small sample size classification problem. On the other hand, as described before, finding new position for landmarks is converted into a small sample size classification problem. This makes it possible for us to use SVM to deal with such finding task.

### IV. SVMBASM

This section detailedly explains the SVMBASM. The proposed SVMBASM consists of two steps

- 1) Training a multi-class SVM classifier for each facial landmark
- 2) Fitting the model

In the training phase, suppose for a given landmark in each training image,  $l$  pixels on either side are sampled, it then gives  $2(l - m) + 1$  sub-profiles which are labeled with  $1, 2, \dots, 2(l - m) + 1$ . Since there are  $n$  training images, it will give  $n \times (2(l - m) + 1)$  training instances, in which each class contains  $n$  training instances. A multi-class SVM classifier for this landmark is then trained with these training instances. This is repeated for each landmark. Thus there will be  $k$  multi-class SVM classifiers as there are  $k$  landmarks in each training image.

In the fitting phase, for a given landmark in a test image,  $l$  pixels on either side are sampled as before to form  $2(l - m) + 1$  sub-profiles, which are then fed into the corresponding classifier. The center of the sub-profile which is identified with the label of  $l - m + 1$  (in fact the center of the sub-profile identified with the label of  $l - m$  or  $l - m + 2$  can also be regarded as the new position) is regarded as the new position for the current landmark. It should be noted that if there are several candidates for the new position, their mean is set as the new position. On the other hand, if no instance is identified with the label of  $l - m + 1, l - m$  or  $l - m + 2$ , the center of the sub-profile with the closest label to  $l - m + 1$ , is set as the new position. In fact, such situation seldom occurs. Finding new position for each landmark, it then gives a displacement vector. The pose and shape parameters are then adjusted as the original ASM does.

#### Summary of training and fitting:

##### 1) Training:

- a) Select  $n$  face images whose ground truth coordinates (marked by hand) of landmarks are available, then build the shape model as described in Section II.
- b) For each landmark, generate  $n \times (2(l - m) + 1)$  training instances by sampling operation described before.
- c) For each landmark, use their training instances to train a multi-class SVM classifier.

##### 2) Fitting:

- a) Give a starting position
- b) For each landmark in the test image, generate  $2(l - m) + 1$  test instances by sampling operation. Then feed these  $2(l - m) + 1$  sub-profiles into

the corresponding multi-class SVM classifier and find which sub-profile is classified with the label of  $l - m + 1$  (or  $l - m, l - m + 2$ ), and finally set the center of this sub-profile as the new position of the current landmark. Repeat this operation for other landmarks and obtain a displacement vector  $dX = (dX_1, dX_2, \dots, dX_k)$ .

- c) Align the current positions  $X$  of the landmarks to their new positions  $X + dX$  and obtain the pose parameters displacements  $ds, d\theta, dX_c, dY_c$ .
- d) Compute the shape parameter  $db$  with equation (8) and equation (12).
- e) Adjust the pose and shape parameters with equation 13.
- f) Compute the new model shape first by equation (2) followed by equation (4).
- g) Perform steps *b* and *c*, *d*, *e* and *f* for a fixed number of iterations.

## V. EXPERIMENTS AND DISCUSSION

To evaluate the performance of the SVMBASM and compare it with the original ASM, SJTU(Shanghai Jiao Tong University) face database [15] is used. It contains view-varying face images of 2278 subjects. For each subject, there are 9 different poses from left-side view to right-side view:  $-90^\circ, -60^\circ, -45^\circ, -30^\circ, 0^\circ, 30^\circ, 45^\circ, 60^\circ, 90^\circ$ . All face images have the size of  $640 \times 480$  pixels. In addition, the ground truth coordinates of 60 landmarks in each frontal face image are marked by hand. We used SJTU face database other than other publicly available face databases for the reason that our SJTU face database supplies the coordinates of the landmarks such that we can quantitatively compare the proposed SVMBASM with original ASM. In our experiments, 200 face images are selected to construct the shape model as well as to train a multi-class SVM classifier for each landmark. Other 1000 face images are chosen for testing.

### A. The evaluation criteria

Our evaluation method is similar to those used in [8]. Given the initial, perturbed parameters, both the original ASM and the SVMBASM fitting algorithm should hopefully make the starting positions of the landmarks converge to the ground truth ones. Two criteria are used to quantitatively measure the performance of the two methods. One is average error:

$$E_{ave} = \frac{1}{r} \times \frac{1}{k} \sum_{i=1}^r \sum_{j=1}^k dis(X(i, j) - pos(i, j)) \quad (14)$$

where  $X(i, j)$  is the ground truth coordinates of the  $j$ th landmark in the  $i$ th test image while  $pos(i, j)$  is the coordinates located by the two methods,  $dis(X(i, j) - pos(i, j))$  is the Euclidean Distance between them. The other criterion is average frequency of convergence, which can be expressed by  $\frac{nr}{r} \times 100\%$ ,  $nr$  being the number of fitting for each method has converged,  $r$  is the size of the test set. A method is considered to be convergible if the average error is less than a threshold after a fixed number of iterations.

### B. The first set of experiments

In this experiment, both methods locate facial landmarks with the same initial positions which is aided by face detection [16] along with eye location [17]. In the training phase, to decide how many pixels on each side of the landmark should be selected to make the SVMBASM perform best regarding the average error, different numbers of pixels on either side of the landmark are chosen to construct different local structures and SVM classifiers for each landmark for original ASM and SVMBASM, respectively. Fig. 5 plots the average error with different numbers of pixels on each side against the number of iteration. From Fig. 5(a) and 5(b), it seems that the original ASM needs fewer iterations to converge. However, neither method converges because the average error for both methods begin to increases after the so called "convergence", which is at the position with the minima average error. As shown in Fig. 5(a), 5(b), 5(c), although neither method converges when the number of pixels on each side takes the value of 3, 4 and 5, the average error of SVMBASM is remarkably lower than that of original ASM. When this number takes the value of 6, 7 and 8, as shown in Fig. 5(d), 5(e), 5(f), the original ASM does not converge as before, the average error of SVMBASM, on the contrary, continuously decreases. In addition, we can also see that the average error of SVMBASM does not clearly change after 25 iterations. This indicates that the SVMBASM converges after 25 iterations.

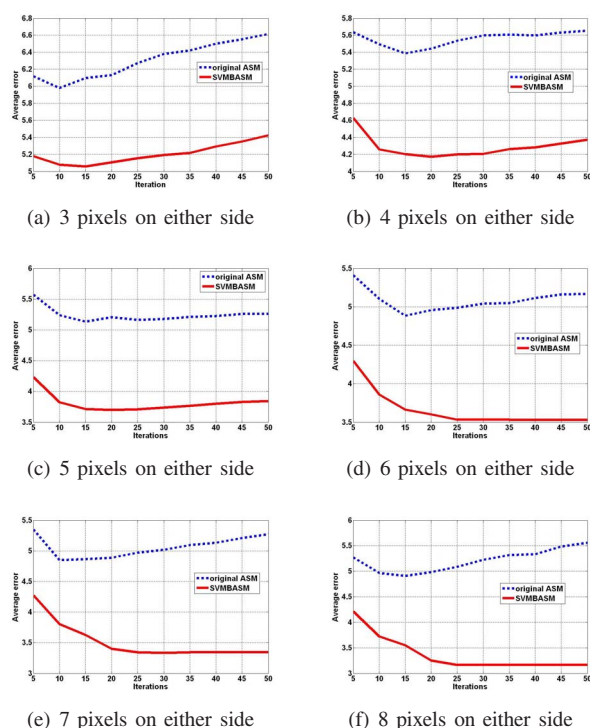


Fig. 5. Average error against the number of iterations with different numbers of pixels on either side of the landmark

The average frequency of convergence is also used to compare the convergence performance of the two methods. Unlike its definition reported in [8], we set the minimum of the



average error of the two methods as the criterion to determine the convergence. A method converges if its average error is less than such minimum after a fixed number of iterations. The average frequency of convergence for both methods is plotted in Fig. 6. It is clearly that the average frequency of convergence for the original ASM is very low in all cases. This is because:

- 1) As shown in Fig. 5, the average error of original ASM in all cases does not decrease continuously, which makes its most fitting trials disable to converge;
- 2) Average error of SVMBASM, always less than that of the original ASM, is selected as the criteria of the convergence;

From Fig. 6 we can also see that the average frequency of convergence for SVMBASM is not satisfying when the number of pixels on each side takes the value of 3, 4 and 5. It is because that the sub-profiles used to train classifiers in these cases contain less discriminative information, which makes it difficult to efficiently distinguish them from each other. On the other side, as shown in Fig. 5, most of the trials converge if this number takes the values of 6, 7 or 8. In fact, we find that the SVMBASM performs best when it takes the value of 12. Nevertheless, in this case, the algorithm needs more time for fitting without obvious improvement in fitting accuracy. Thus a tradeoff between accuracy and computation time is considered. In the following experiments, this number takes the value of 8 all the time. Figure 6 also indicates that the average frequency of the convergence for the SVMASM is quite higher than that of the original ASM in all cases. This further proves that SVMBASM performs better. Some location

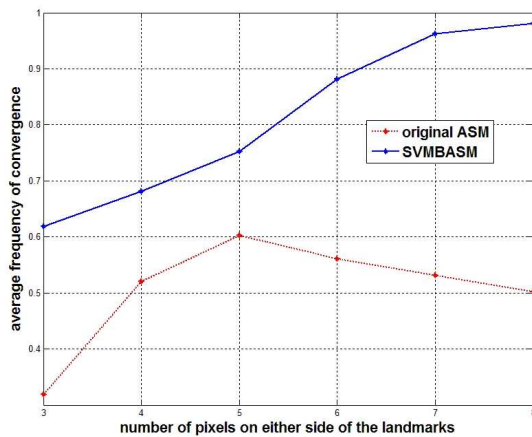
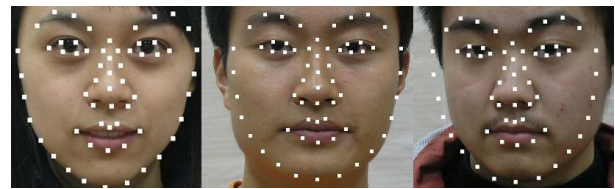


Fig. 6. Average frequency of convergence

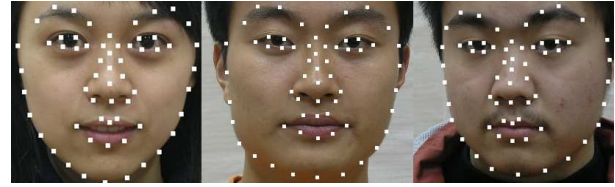
results of SVMBASM and original ASM are included in Fig. 7.

### C. The second set of experiments

In this experiment, it is going to compare the robustness of two methods to the changes of the initial coordinates of the landmarks under affine transform. Given a test image,



(a) Results of the ASM



(b) Results of the SVMBASM

Fig. 7. Comparative results of both methods

the ground truth coordinates of landmarks were systematically perturbed to generate the initial pose parameters to start the algorithms with. Both algorithms were performed on a large number of such trials and the average frequency of convergence was computed by measuring how often the algorithms converge after 25 iterations. We generate trials for perturbations of varying magnitude and plot the frequency of convergence against the magnitude of the perturbation. For each test image, we aligned the mean shape  $\bar{a}$  to the ground truth shape and then obtain the starting pose parameters  $s, \theta, X_c, Y_c$ . Some turbulences were introduced by systematically displacing the ground truth shape about  $\pm 40$  pixels distance in  $x$  and  $y$  directions, scaling its size by a factor of between  $[-0.3 \sim +0.3]$ , and rotating between  $[-30^\circ \sim +30^\circ]$ . Changes on the positions, scales and rotations were carried on separately in order to obtain the different initial positions of the landmarks for the fitting process to be done later on. The average frequency of convergence against the different displacements of pose parameters are plotted in Fig. 8. In the experiments, if the average frequency of the convergence is larger than 90%, the relative algorithm is convergent. From Fig. 8(a) and 8(b), it can be seen that the SVMBASM displays good convergence with up to 20 pixels(15% of the face width) displacement in  $x$  direction and 15 pixels displacement in  $y$  direction. The performance in  $y$  direction is not good as in  $x$  direction due to the affection caused by the hair. When some landmarks of the starting shape model lie on the hair, the accuracy of finding task for these landmarks will suffer from the affection of hair. Fig. 8(c) shows that the SVMBASM converges on the condition that the rotation displacements of the starting shape model lies between  $[-15^\circ \sim +15^\circ]$ . Fig. 8(d) tells us that the SVMBASM can converge as long as the displacements of the scale of the starting shape model varies from  $[-0.15 \sim 0.25]$ , from which we can also see that the SVMBASM tends to converge when the scale of the starting shape model large than 1. In this experiment, if the value of the average frequency of convergence, 90%, is chosen as the criteria to determine

the convergence of the algorithms, the original ASM is not convergent in any case. Four location results of both methods

Science Foundation of China (No:60675023, No: 60602012) and China 863 project(No:2007AA01Z164).

## REFERENCES

- [1] G. C. Feng and P. C. Yuen, "Multi-cues eye detection on gray intensity image," *Pattern Recognition*, vol. 34, no. 5, pp. 1033–1046, 2001.
- [2] S. H. Leung, S. L. Wang, and W. H. Lau, "Lip image segmentation using fuzzy clustering incorporating an elliptic shape function," *Image Processing, IEEE Transactions on*, vol. 13, no. 1, pp. 51–62, 2004.
- [3] Z. H. Zhou and X. Geng, "Projection functions for eye detection," *Pattern Recognition*, vol. 37, no. 5, pp. 1049–1056, 2004.
- [4] K. F. Lai and R. T. Chin, "Deformable contours: modeling and extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 11, pp. 1084–1090, 1995.
- [5] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
- [7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
- [8] I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135–164, 2004.
- [9] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 8, pp. 924–933, 2002.
- [10] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for computer vision," *World Wide Web Publication*, February, 2001.
- [11] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 2000.
- [12] M. E. Mavrouforakis and S. Theodoridis, "A geometric approach to support vector machine (svm) classification," *Neural Networks, IEEE Transactions on*, vol. 17, no. 3, pp. 671–682, 2006.
- [13] S. Zheng, J. Liu, and J. W. Tian, "A new efficient svm-based edge detection method," *Pattern Recognition Letters*, vol. 25, no. 10, pp. 1143–1154, 2004.
- [14] J. Ramirez, P. Yelamos, J. M. Gorriz, and J. C. Segura, "Svm-based speech endpoint detection using contextual speech features," *Electronics Letters*, vol. 42, no. 7, pp. 65–66, 2006.
- [15] C. Du, J. Yang, Q. Wu, T. Zhang, H. Wang, L. Chen, and Z. Wu, "Extended fitting methods of active shape model for the location of facial feature points," *LECTURE NOTES IN COMPUTER SCIENCE*, vol. 4338, p. 610, 2006.
- [16] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [17] T. Kawaguchi and M. Rizon, "Iris detection using intensity and edge information," *Pattern Recognition*, vol. 36, no. 2, pp. 549–562, 2003.

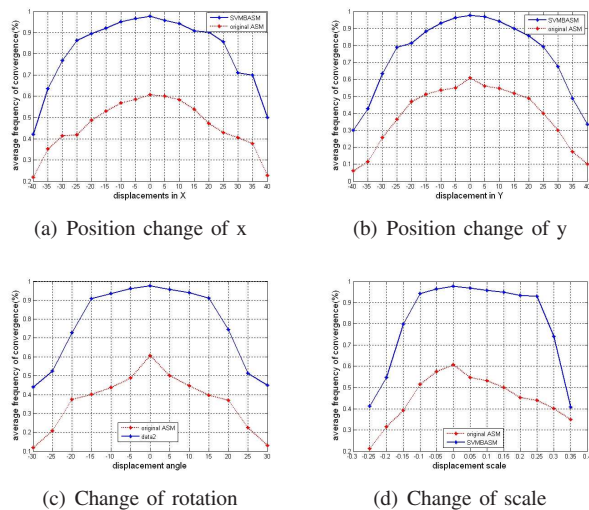


Fig. 8. Average errors against the number of iterations with the different numbers of pixels on either side of the landmark

by perturbing the pose parameters in  $X_c$ ,  $Y_c$ ,  $\theta$  and  $s$  are given in Fig. 9 from left to right. They again tell us that SVM-BASM outperforms original ASM.

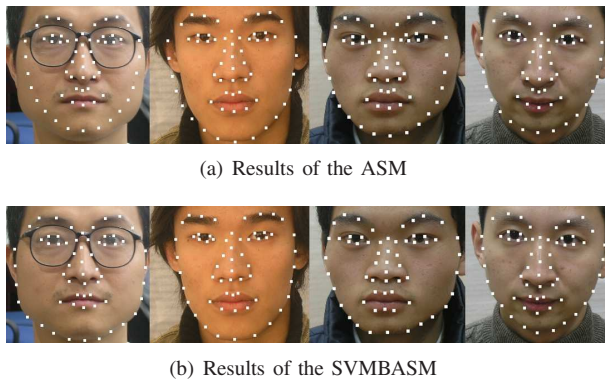


Fig. 9. Comparative results by changing the pose parameters

## VI. CONCLUSION

We have presented a new method called SVM-BASM for the location of facial landmarks, where the task of finding new position for facial landmark is converted into a small sample size classification problem. We have assessed the location performance of the proposed method on our SJTU face database. Two sets of experiments show that that the advocated method outperforms the original ASM in terms of the average error as well as the average frequency of convergence.

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions. This research has been supported by the National Natural



2008  
**CIIT**

Proceedings of  
**2008 IEEE 8<sup>th</sup> International Conference  
on Computer and Information Technology**

**8-11 July 2008 ■ Sydney, Australia**

[Hub Page](#)

[Table of Contents](#)

[Author Index](#)

[Search](#)

[Support](#)

[Install Software](#)

**Editors:**

**Qiang Wu, Xiangjian He, Quang Vinh Nguyen, Wenjing Jia and Maolin Huang**

Proceedings of the 2008 IEEE 8th International Conference on Computer and Information Technology. IEEE Catalog Number CFP08355-CDR, ISBN 978-1-4244-2358-3. Library of Congress 2008902859. Copyright © 2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. Photo: Hamilton Lund. For technical support please contact Causal Productions (info@causalproductions.com).

*See attached 70c and highlighted items.*

*2008000787 70c*



# **Proceedings**

**2008 IEEE 8th International Conference  
on Computer and Information Technology**

## **CIT 2008**

8-11 July 2008 • Sydney, Australia

### **Editors**

**Qiang Wu, Xiangjian He, Quang Vinh Nguyen, Wenjing Jia and Maolin Huang**

### **Organized by**

**CIT Organizing Committee**

### **In cooperation with**

**IEEE Computer Society**

**IEEE Technical Committee on Scalable Computing**

**University of Technology, Sydney**

**Research Institute for Information and Communication Technology,**

**Korea University, Korea**

**BK 21 Information Technology Division, Korea University, Korea**

**ARC Research Network in Enterprise Information Infrastructure (EII), Australia**

**Federation of Chinese Scholars in Australia (FOCSA), Australia**

**Australian Chinese ICT Professional Society, Australia**

## MESSAGE FROM CHAIRS

On behalf of the Organizing Committee and the Technical Program Committee, it is our great pleasure to welcome you to Sydney and the IEEE 8th International Conference on Computer and Information Technology 2008 (CIT2008).

As an IEEE CS fully sponsored conference, International Conference on Computer and Information Technology (CIT) has become a major platform for researchers and industry practitioners from different fields of computer and information technology. Each year, CIT attendees appreciate and benefit from multidisciplinary exchanges in computer and information technology. CIT has attracted many high quality research papers which highlight foundational work that strives to push beyond limits of existing computer technologies, including experimental efforts, innovative systems, and investigations that identify weaknesses in existing IT services.

The topics of CIT2008 include:

- Database and Data Mining
- High Performance Computing
- Multimedia and Computer Graphics
- Ubiquitous and Sensor Networks
- Information Systems
- Service Oriented Computing
- Internet and Web Applications
- Networking and Mobile Computing
- Image Processing, Computer Vision and Video Surveillance
- VLSI and Computer Systems
- Software Engineering

This year, the number of papers submitted to CIT2008 has overwhelmed those of previous editions. CIT2008 received more than 550 submissions coming from various countries and selected about 150 (27%) as CIT2008 regular papers to be included in the CIT2008 proceedings. The revision process for all papers has been rigorous and thorough, including peer-reviewing from at least two independent and qualified reviewers, and reviewing from the local organization committee. The accepted papers for this volume of proceedings are expected to be of very high quality. We would like to express our sincere gratitude to all the reviewers, members of the Program Committee and the organization committee who have provided CIT2008 with such a qualified technical program. We also would like to express our special appreciation to the following organizers of various invited sessions of CIT2008, Dr. Liang Wang, Dr. Ming Li, Dr. Xin Geng, Dr. Jiali Cui, Dr. Joan Lu, Dr. Fahim Akhter and Dr. Huwida Said.

We deeply thank all the authors, attendees, sponsors, organizers and all others who have been involved in organizing this event at any level for their invaluable support. We hope that you will enjoy the conference in all its aspects so that we will be able to count on your interest and support also in the future.

Xiangjian (Sean) He  
General Chair, CIT2008

Qiang Wu  
Program Chair, CIT2008