

LOCATING FACIAL FEATURES BY ROBUST ACTIVE SHAPE MODEL

Jiani Hu, Yu Li, Weihong Deng, Jun Guo, Weiran Xu

Beijing University of Posts and Telecommunications, Beijing, China
jnhu @bupt.edu.cn, liyu.sice.bupt@gmail.com, { whdeng, guojun, xuweiran }@bupt.edu.cn

Abstract

Active shape model statistically represents a shape by a set of well-defined landmark points and can model object variations using principal component analysis. However, the shape generated by standard active shape model is unsmooth when the test sample has a large variation compared with the training images. In this paper, we introduce a robust active shape model for facial feature location. First, a color information and 2-dimension based local feature model is presented to characterize salient facial features, such as the eyes and the mouth. Then, a regularized principal component analysis based shape model is proposed to construct a smooth global shape. We evaluate our approach on a challenging dataset containing 2,000 well-labeled facial images with a large range of variations in pose, lighting and expression. Experimental results demonstrate the efficiency and effectiveness of the proposed approach.

Keywords: facial feature; active shape model; local profile;

1 Introduction

The goal of facial feature location is to match a 2D face shape with a facial image. The shape is made up of a set of important points, including face contour, eye centers, nose, mouth, and etc. Precise localization of the facial feature plays an important role in facial analysis. Its applications range from face recognition [1], face tracking, facial expression analysis, 3D face modeling, face cartoon animation [2], to face search and annotation [3]. However, the variety of human faces, expressions, facial hair, glasses, poses, and lighting contribute to the difficulty of automatic and accurate location of facial features.

Active Shape Model (ASM) has been shown to be a powerful tool to aid the facial feature location. Classical ASM firstly build a point distribution model for each feature, which captures the local shape variants and gradient distributions of a set of landmark points. Then, a shape model is proposed to generate a globally reliable face shape by the set

of local points using principal component analysis [4].

In order to construct a robust shape model, many improved ASMs have been developed. The improvements mainly lie in two aspects: 1) training refined local feature models and 2) using statistic framework to estimate the shape parameters. For local feature description models, two non-linear ASMs are proposed to characterize distributions of intensity [5] and to incorporate a reduced set of differential invariant features as local image descriptors [6]. S. Zhang proposes a cascade structure based ASM to give an accurate and robust description of landmarks [7]. Recently, a new feature model is introduced which is robust when illumination changes dramatically [8]. For shape parameter estimation, a Bayesian tangent shape model is proposed to estimate parameters by Bayesian inference [9]. Coughlan and Ferreira [10] introduce a Markov Random Field based shape model to capture both the local image structure and the shape prior. AdaBoosted histogram classifiers are included to model local features and optimize shape parameter [11]. Huang et al [8] use a nonlinear constrained Gaussian process latent variable model to represent global shape prior.

However, most of the above algorithms are inadequate to localize exact feature points when the illumination, pose, and deformation are changing dramatically. The reason might lie in following: 1) local features could not be represented accurately by gray-level information and 1D profile, 2) the searching process of each point by Mahalanobis distance might be unfaithful, and 3) principle component analysis (PCA) based global shape model couldn't ensure the smoothness of a new shape.

In this paper, we utilize the color information and 2D profile to characterize local facial features. This can enhance the localization of geometric features that represent the external boundaries in face images. Then we propose a robust regularized principle component analysis based shape model, which can generate smooth new shapes. Our experimental results show that the proposed

approach outperforms the standard ASM technique for facial feature extraction.

The rest of this paper is organized as follows. Section 2 discusses the classic ASM approach for facial feature location. Section 3 demonstrates the proposed algorithm. Section 4 presents the experimental results and conclusions are given in Section 5.

2 Active shape model

Given an image I , the task of facial feature location is to find the position of a number of landmarks on the 2D image. A landmark represents a distinguishable point in the face image, for example, centers of eyes, nose, mouth, and etc. 68 landmarks used in this paper are shown in Figure 1. Given a set of training faces with manually labeled landmarks, standard ASM approach first learns local profile models of every landmark and a global shape model. Then it can predict the feature points of a new face image by the processing of model searching.

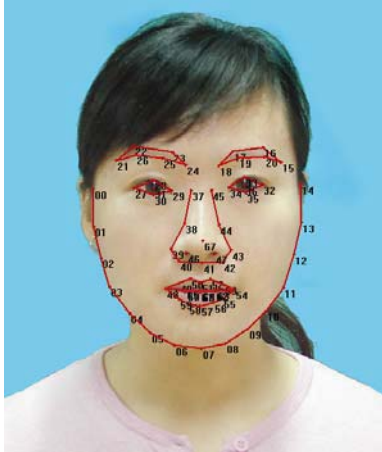


Figure 1. A face with 68 positioned landmarks

2.1 Local profile models

The profile models describes local image features around each landmark, which are modeled as the first derivative of the sample profiles perpendicular to the landmark contour. To form the profile vector \mathbf{g} at a certain landmark, standard ASM algorithm samples image intensities along a one-dimensional line, which is a vector at the landmark orthogonal to a shape edge. Figure 2 shows a profile vector. The profile vector is formed as follows: 1) set each element of the profile vector to the gray level (0...255), 2) replace each profile element by the intensity gradient, which is done by replacing the profile element at each position i with the difference between it and the element at $i-1$, 3) divide each element of the resulting vector by the sum of the absolute values of all vector elements.

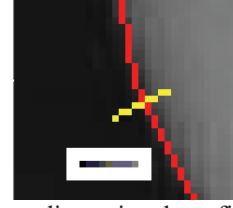


Figure 2. A one-dimensional profile. (The yellow line shows the location of a 1D profile, which is orthogonal to the face shape boundary demonstrated by the red line. The inset shows the image intensity along the profile.)

During training, local profile models are built for each landmark by creating a mean profile $\bar{\mathbf{g}}$ and a covariance matrix \mathbf{S}_g of all training profiles at that landmark. If the length of the profile is 7, $\bar{\mathbf{g}}$ will have 7 elements, and \mathbf{S}_g will be a 7*7 matrix. If there are 68 landmarks, there will be 68 separate $\bar{\mathbf{g}}$'s and \mathbf{S}_g 's.

In the process of searching, several search profiles are formed by sampling the input image in the neighborhood of the landmark. Each search profile is centered at small displacements perpendicular to the landmark contour. The distance between a search profile \mathbf{g}_s and the model mean profile $\bar{\mathbf{g}}_i$ is calculated using the Mahalanobis distance

$$f_i(\mathbf{g}_s) = (\mathbf{g}_s - \bar{\mathbf{g}}_i)^T \mathbf{S}_i^{-1} (\mathbf{g}_s - \bar{\mathbf{g}}_i) \quad (1)$$

One of the search profiles, which has the lowest distance, is regarded as the position of the landmark. This process is repeated to search every landmark before handing control to the shape model.

2.2 The shape model

The task of the shape model is to organize the separate feature points suggested by the profile models to an allowable face shape. The shape model assumes that a shape can be represented as a linear combination of the training shapes:

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \Phi \mathbf{b} \quad (2)$$

where $\hat{\mathbf{x}}$ is the generated shape vector (all the x- followed by all the y-coordinates, i.e. $\hat{\mathbf{x}} = \{x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n\}^T$). $\bar{\mathbf{x}}$ is the mean shape, which is the average of the aligned training shapes \mathbf{x}_i and defined as

$$\bar{\mathbf{x}} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{x}_i \quad (3)$$

where n_s is the number of shapes in the training set. Φ is made up of the eigenvectors of the covariance matrix Σ_s of the training shape points

$$\Sigma_s = \frac{1}{n_s - 1} \sum_{i=1}^{n_s} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (4)$$

After applying a standard principal components analysis, the eigenvalues of Σ_s are ordered descending ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2n}$, n denotes the number of landmarks in the training set). The number of eigenvectors to be chosen depends on empirical testing.

2.3 ASM based facial features location

To summaries, the ASM based facial feature points location algorithm is to find a proper parameter \mathbf{b} to generate the shape of test image from the training set, which has two parts. The one is training, and the other is searching. In the training process, every image is aligned to the mean shape. And the local profile covariance matrixes \mathbf{S}_g 's and shape covariance matrix Φ are calculated respectively. In the searching process, the algorithm is shown as follows:

- 1) Initialize current shape of the test image according to the mean shape of the training set.
- 2) Search for the optimal landmarks according to the local profile models.
- 3) Update the shape parameters under the constraints of statistical shape model using equation (2).
- 4) Repeat 2) and 3) until the change of the model is under the range of the given threshold.

3 Robust active shape model

In this section, we introduce a Robust Active Shape Model (RASM) to improve the generalization ability of classical ASM. The improvements are as follows: 1) local image models using 2D profile and color information are applied to characterize salient facial features, and 2) regularized shape model is proposed to construct a smooth global shape. Figure 3 shows 2D and color information based profile.

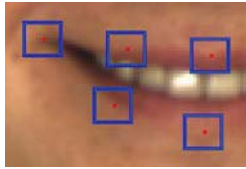


Figure 3. Some 2D profiles. (Each 2D profile is the area inside the blue box, red points represent current landmarks)

3.1 Local image models using 2D profile and color information

As stated in section 2, the optimal feature points are limited along the orthogonal line of the shape edge. Moreover, the standard ASM only use gray-level image to model the local structure. In order to construct robust local structure models, we extend

the profile from 1D to 2D and apply color-level information.

We sample a square region around the landmark to build local profile models, because a 2D profile can captures more information of the landmark. The covariance matrix of a set of 2D profiles is formed by treating each 2D profile matrix as a long vector, and calculating the covariance of the vector. Moreover, we assume that the three color channels, i.e. Red, Green, and Blue, are statistically independent and the normalized first derivative of the color values in 2D profile for each individual channel satisfies a multivariate Gaussian distribution. Then, a weighted sum of multivariate Gaussian model is proposed to model the local structure of the feature points in training set. The best matching of a search profile in RGB color space, \mathbf{g}_{rgb} , to the reference model is carried out by:

$$f_i(\mathbf{g}_{rgb}) = \sum_{i \in r, g, b} w_i \cdot (\mathbf{g}_i - \bar{\mathbf{c}}_i)^T \Sigma_i^{-1} (\mathbf{g}_i - \bar{\mathbf{c}}_i) \quad (5)$$

where $\bar{\mathbf{c}}_i$ is the mean profile of the i -th color, w_i is the weighting factor for the i -th color component. In this paper, we assume three color channels have the same contribution. Thus, the weight w_i is 1/3.

3.2 Regularized shape model

Standard ASM using principal components analysis to generate a new shape, but the number of eigenvectors to be chosen depends on empirical testing. We propose a Regularized Shape Model (RSM), which is robust when test image varying dramatically from the training set. The essence of regularized shape model is applying regularized principal components analysis to acquire a robust eigenvectors matrix Φ in equation (2).

The covariance matrix Σ_s of the training shape points can be factorized into following form using the principal components analysis (PCA),

$$\Sigma_s = \frac{1}{n_s - 1} \sum_{i=1}^{n_s} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T = \Phi \Lambda \Phi^T \quad (6)$$

$$\text{with } \Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_{2n}\}$$

where $\Phi \in R^{2n \times 2n}$ is an orthogonal eigenvector matrix and $\Lambda \in R^{2n \times 2n}$ is a diagonal eigenvalue matrix with diagonal elements in decreasing order $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2n}$ (n denotes the number of landmarks in the training set). To deal with the principal component selecting problem of Σ_s , the RSM performs a form of regularization by adding a multiple of identity matrix to Λ , as $\Lambda = \Lambda + \sigma \mathbf{I}$, for some $\sigma > 0$. The determination of σ is not a trivial problem: if $\sigma \rightarrow \infty$, we will lose the information of Σ_s , and a small value of σ may not be sufficient effective. Cross-validation can be applied to estimate the optimal σ , but this solution is computationally expensive and the optimal value

varies in different training data. The RSM circumvents this problem by taking into account both the spectral energy and the magnitude requirement. The eigenvalue spectral of Σ_s provides a good criterion for estimating the

optimal σ . The RSM determines the optimal σ^* as

$$\sigma^* = \lambda_m, \quad (7)$$

$$\text{with } J(m) = \min_m \left\{ m \left| \sum_{i=1}^m \lambda_i / \sum_{i=1}^{2n} \lambda_i \geq E(\theta) \right. \right\}$$

where $J(m)$ determines the value of m according to the proportion of energy captured in the first m eigenvectors. In the paper, we empirically set the threshold $E(\theta) = 0.95$ for all the experiments. It is

evident that with a settled $E(\theta)$, σ^* is self-adaptive

to different training data according to the eigenvalue spectral of Σ_s .

4 Experimental results

To evaluate the proposed algorithm, we collect 2,000 images from two databases, including AR database [12] and FRGC database [13]. We randomly select 1,000 images for training and the rest 1,000 images for testing.

For performance evaluation, we adopt average error, which is defined as the distance between the manually labeled shape and the resulting shape of our RASM:

$$AE = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{n} \sum_{j=1}^n \text{EDist}(P_{ij} - L_{ij}) \right) \quad (8)$$

where N is the total number of the test images, n is the number of the landmark points in the training set (for our case, $n=68$), L_{ij} is the j -th landmark point in the manually labeled shape of the i -th test image manually labeled, P_{ij} is the j -th landmark point in the resulting shape of ASM for the i -th test image. The function EDist is the Euclidean distance between the two points.

Table 1. Performance comparison

Method	Average Error	Improvement to the baseline
standard ASM (baseline)	3.76	-----
2D profile and color image based ASM	3.60	4.26%
regularized shape model on gray image	3.55	5.59%
RASM	3.47	7.71%

As Table 1 showing, the average error for standard ASM is about 3.76 pixels. After applying 2D profile and color information based local profile models on standard ASM, the average error is reduced to 3.60 pixels. The error is reduced to 3.55 pixels after applying regularized shape model on gray-level based 1D profile. With the above two schemes applied the error is reduced to 3.47 pixels per point. The overall improvement of our strategies to the standard ASM by our performance system is about 7.71%. Figure 4 demonstrates the comparison result of standard ASM and RASM.



Figure 4. Result of the standard ASM and RASM. (a) Shapes generated from standard ASM. (b) Shapes generated from RASM.

5 Conclusions

In order to solve the facial features location problem, we propose a Robust Active Shape Model. The main contributions of the new model are as follows:

- 1) Color information and 2D profiles are utilized to characterize local facial features. This enhances the localization of geometric features that represent the external boundaries in face images.
- 2) To avoid generating an unreasonable global shape, regularized principal component analysis based active shape model is applied. It can improve the generalization ability of standard

ASM, which is robust when test image varying dramatically from the training set.

The proposed approach is compared with the standard ASM. Experimental results demonstrate that our method is robust to changes in illuminations, different expressions, and obstacles.

Our future work will be focused on the accuracy of the salient features localization and some more deliberate fused local texture models for the landmarks such as eye centers, mouth corners. Global shape model with appearance is also to be exploited.

Acknowledgements

This work was supported by The Fundamental Research Funds for the Central Universities, ZTE Corporation Scientific Fund, The 111 project under Grant No.B08004, and National Natural Science Foundation of China under Grant No. 60905017.

References

- [1] L. Wiskott, J.M.Fellous, N.Kr"uger, and C.von der Malsburg. Face recognition by elastic bunch graph matching. *In Intelligent Biometric Techniques in Fingerprint and Face Recognition*, pp. 355-396, 1999.
- [2] Hong Chen, Ying-Qing Xu, Heung-Yeung Shum, Song-Chun Zhu, and Nan-Ning Zheng. Example-based facial sketch generation with non-parametric sampling. *In ICCV01*, pp. 433-438, 2001.
- [3] Jingyu Cui, Fang Wen, Rong Xiao, and Xiaou Tang. Easyalbum: An interactive photo annotation system based on face clustering and re-ranking. *In proceedings of the SIGCHI conference on Human factors in computing systems*, pp.7, 2007.
- [4] T. F. Cootes, C. J. Taylor, and J. Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, 61:38-59, 1995.
- [5] B.V. Ginneken, A.F. Frangi, J.J. Staal, B. M. T. H. Romeny, and M. A. ViergeverB. Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21(8), pp.924-933, 2002.
- [6] F. Sukno, S. Ordas, C. Butakoff, S. Cruz, and A.F. Frangi. Active shape models with invariant optimal features IOF-ASMs. *In proceedings of AVBPA*, pp.365-375, 2005.
- [7] S. Zhang, L.F. Wu, Y. Wang, Cascade MR-ASM for Locating Facial Feature Points. *In proceedings of 2nd International Conference on Biometrics*, pp.683-691, 2007.
- [8] Y. Huang, Q. Liu, D. Metaxas. A Component Based Deformable Model for Generalized Face Alignment. *In proceedings of eleventh IEEE International Conference on Computer Vision*, pp.1-8, 2007.
- [9] Y. Zhou, L. Gu and H.J. Zhang. Bayesian tangent shape model: Estimating shape and pose parameters via Bayesian inference. *In IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 109-118, 2003.
- [10] J. Coughlan and S. Ferreira. Finding deformable shapes using loopy belief propagation. *In The Seventh European Conference on Computer Vision*, pp. 453-468, 2002.
- [11] Y.Z. Li and W. Ito. Shape parameter optimization for Adaboosted active shape model. *In proceedings of IEEE International Conference on Computer Vision*, pp. 259-265, 2005.
- [12] A.M. Martinez and R. Benavente. The ar face database. CVC Technical Report #24, 1998.
- [13] P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer. Overview of the Face Recognition Grand Challenge. *In proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 947-954, 2005.