Locality-constrained Active Appearance Model

Xiaowei Zhao[1,2], Shiguang Shan[1], Xiujuan Chai[1], and Xilin Chen[1]

[1]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
mathzxw2002@gmail.com,{sgshan,chaixiujuan,xlchen}@ict.ac.cn

Abstract. Although the conventional Active Appearance Model (AAM) has achieved some success for face alignment, it still suffers from the generalization problem when be applied to unseen subjects and images. In this paper, a novel Locality-constraint AAM (LC-AAM) algorithm is proposed to tackle the generalization problem of AAM. Theoretically, the proposed LC-AAM is a fast approximation for a sparsity-regularized AAM problem, where sparse representation is exploited for non-linear face modeling. Specifically, for an input image, its K-nearest neighbors are selected as the shape and appearance bases, which are adaptively fitted to the input image by solving a constrained AAM-like fitting problem. Essentially, the effectiveness of our LC-AAM algorithm comes from learning a strong localized shape and appearance prior for the input facial image through exploiting its K-similar patterns. To validate the effectiveness of our algorithm, comprehensive experiments are conducted on two publicly available face databases. Experimental results demonstrate that our method greatly outperforms the original AAM method and its variants. In addition, our method is better than the state-of-the-art face alignment methods and generalizes well to unseen subjects and images.

1 Introduction

Face alignment is the process of moving and deforming a face model to match with the input facial image, which plays an important role in many computer vision problems, such as face recognition, facial expression analysis, face tracking in video, etc. Among sizable literatures on face alignment, the Active Shape Model (ASM) [1] and Active Appearance Model (AAM) [2] are the early popular approaches which attempt to fit the facial image with a statistical generative model.

As an extension of ASM, the power of AAM stems from statistically modeling the shape and appearance variations simultaneously through principal component analysis (PCA) on a set of labeled data. During the fitting procedure, the AAM is aligned by finding the model parameters that minimizing the distance between the observed and the synthesized facial appearance. As indicted by work [3], the AAM performs well if trained to work with a limited number of known subjects (e.g., person-specific Active Appearance Model). However, the alignment performance of AAM degrades quickly when it is trained on a large data

set and fitted to images that were not seen during the training procedure. We assert that this is mostly due to the fact that a single PCA model cannot well capture the non-linear appearance variation of a large training set, which contains larger variations (e.g., pose, expression, lighting, etc) than a small image set. Specifically, the learnt PCA model just preserves the statistically significant features of a training set, ignoring some subtle variations of images.

To tackle the generalization problem of AAM, two main kinds of approaches are proposed: discriminative and generative. The first approach learns a discriminative fitting function, which establishes a mapping between the facial appearance and the correct alignment. For example, Boosted Appearance Model (BAM) [4] exploits Haar-like features to model the appearance of faces, and learns a GentleBoost classifier to distinguish between correct and incorrect alignment. Then, the correct alignment can be obtained by maximizing the classification score. However, there is no guarantee that moving along the gradient of the learnt score function will always improve the alignment [5]. In order to overcome this limitation, Boosted Ranking Model (BRM) [5] proposes to learn a ranking function which is concave within the neighborhood of the correct alignment. Given a pair of images warped from different landmarks, the learnt GentleBoost-based ranking classifier can inform which alignment is better. Furthermore, a non-linear discriminative fitting approach is proposed by Saragih et al. [6], which learns a non-linear multivariate regressor through Boosting and directly predicts the update parameters in each iteration of AAM.

For generative approaches, there are many efforts to handle the multi-model distribution of shapes and appearances. Typically, Gaussian Mixture Model (GMM) is exploited to represent the non-linear shape and appearance variations [7, 8]. Specifically, in literature [8], the multi-model appearance variations are captured by a mixture of probabilistic PCA (MPPCA) [9], noted as MPPCA-AAM. Then, the AAM alignment problem is formulated as a maximum likelihood problem, which can be easily solved by EM-algorithm. Besides the GMM methods, manifold learning technique can also be utilized to learn a non-linear prior for shapes and appearances [10]. However, the variations of facial shape and appearance are too complex to be characterized by a parametric distribution model.

In recent years, sparse representation is also utilized to model the complex non-linear distribution of facial shapes instead of PCA [11]. In their method, a novel Sparse Shape Composition model (SSC) is proposed to adaptively approximate the input shape by a sparse linear combination of training shapes. The effectiveness of SSC is validated on some real world medical applications. However, due to the high dimension (e.g., 40*50) of facial appearance, it is impractical to directly apply sparse representation to non-linear appearance modeling in AAM, which results in a computation-cost problem. Moreover, the multiple iteration times (at least 10 times) in the AAM fitting procedure also increases the computation cost. Fortunately, as empirically pointed by Yu et al., the results of sparse representation tend to be local: nonzero coefficients are often assigned to bases nearby to the encoded data [12]. So, the sparse representation-based

method can be fast approximated by a K-nearest-neighbor (K-NN) search and then solving a constrained least square fitting problem [13].

In this paper, inspired by the empirical observations of Yu et al., we first reformulate the original Active Appearance Model as a sparse representation problem and then approximate it through introducing locality constraint, briefly called Locality-constrained Active Appearance Model (LC-AAM). Specifically, for an input image, we first find its K-nearest neighbors as the face bases and then adaptively fit to it by solving a constrained AAM-like fitting problem. The effect of locality constraint is twofold: (1) Learning the shape and appearance prior for the input facial image through exploring its K-nearest neighbors, which have similar patterns (e.g., pose, expression, subject, etc.). The person-specific AAM is just a special case of our LC-AAM, which requires that the images come from one specific person. (2) The global non-linear facial shape and appearance model is approximated by many localized linear models, which are specific to the input images and result in a faster convergence. To demonstrate the effectiveness of our method, comprehensive experiments are conducted by comparing our LC-AAM method with conventional AAM-based methods (e.g., the popular Inverse Compositional AAM (IC-AAM) [2], Simultaneously Inverse Compositional AAM (SIC-AAM) [3], and MPPCA-AAM [8]) and the state-of-the-art face alignment methods on two publicly available face databases, which contain multiple subjects and cover large variations in pose, expression and lighting, etc. Experimental results demonstrate that our proposed LC-AAM algorithm convincingly outperforms the conventional AAM-based methods. In addition, our method generalizes well to unseen subjects and images and is better than the state-of-the-art face alignment methods on the above-mentioned evaluation sets.

The remaining part of this paper is organized as follows. Section 2 gives a brief review of Active Appearance Model. Section 3 reformulates the original AAM problem as a sparsity-regularized AAM problem, which is future approximated by our Locality-constrained AAM algorithm. The formulation and implementation details are also presented in this section. Section 4 reports the experimental results and also the comparisons with the state-of-the-art methods. Section 5 concludes the paper.

2 Review of Active Appearance Model

In this section, we will first introduce the shape and appearance modeling of the conventional AAM, then describe the AAM-based fitting algorithm. Some notations used in the following sections are also defined in this section.

2.1 Shape and Appearance Modeling

The AAM simultaneously characterizes the intrinsic variations of shape and appearance as linear combination of basis models of variation.

The shape s of an AAM is represented by a set of 2D facial landmarks: $s = (x_1, y_1, x_2, y_2, \ldots, x_n, y_n)^{\mathrm{T}}$. The AAM allows a linear shape variation, which

means that the shape instance s can be expressed as follows:

$$s = s_0 + \sum_{i=1}^{M} \alpha_i s_i \tag{1}$$

where s_0 is the mean shape, s_i is the ith shape basis, and the coefficients $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_M]^{\mathrm{T}}$ are the shape parameters to control the variation of s. The first four shape bases are designed to represent the global similar transformations, such as global translations in x and y direction, rotation, and scaling. The shape model of AAM is statistically learnt from a training set with annotated facial landmarks. Specifically, the training shapes are first geometrically aligned using the *Procrustes* analysis. Then eigenanalysis is applied to the aligned training shapes to obtain the mean shape and shape bases.

To learn the appearance model, each facial image I should be warped into a "shape-normalized" frame, which is usually defined as the mean shape s_0. The warping function, noted as $W(\boldsymbol{u}; \alpha)$, is usually defined as a piecewise-affine warp from the mean shape s_0 to the shape of facial image. Here, α are the shape parameters of facial image I, and $\boldsymbol{u} = (x, y)^{\mathrm{T}}$ denote a set of pixels lie inside the mean shape s_0. So, the "shape-normalized" appearance of facial image I is defined as $I(W(\boldsymbol{u}; \alpha))$. Then the AAM appearance model is computed by applying PCA to the collected "shape-normalized" appearances. Similarly to shape modeling, the appearance instance is generated using a linear combination of L appearance bases:

$$A(\boldsymbol{u}) = A_0(\boldsymbol{u}) + \sum_{i=1}^{L} \beta_i A_i(\boldsymbol{u}) \tag{2}$$

where $A_0(\boldsymbol{u})$ is the mean appearance, $A_i(\boldsymbol{u})$ is the ith appearance basis, and the coefficients $\beta = [\beta_1, \beta_2, \ldots, \beta_L]^{\mathrm{T}}$ are the appearance parameters to control the variation of $A(\boldsymbol{u})$.

In the early literatures [14], the shape and appearance parameters are usually concatenated and a second level PCA is applied to these concatenated parameters to form a more compact representation. However, in this study, the shape and appearance model are taken into account separately.

2.2 Model Fitting

The fitting of AAM is to estimate the optimal shape and appearance parameters to minimize the discrepancy between the synthesized image $A(\boldsymbol{u}) = A_0(\boldsymbol{u}) + \sum_{i=1}^{L} \beta_i A_i(\boldsymbol{u})$ and the observed facial image $I(W(\boldsymbol{u}; \alpha))$. Here, I is the test image, and $I(W(\boldsymbol{u}; \alpha))$ represents the warped "shape-normalized" facial image. Specifically, the AAM fitting algorithm is usually formulated as follows:

$$\min_{\alpha,\beta} \sum_{\boldsymbol{u} \in s_0} [A_0(\boldsymbol{u}) + \sum_{i=1}^{L} \beta_i A_i(\boldsymbol{u}) - I(W(\boldsymbol{u}; \alpha))]^2 \tag{3}$$

where the sum is performed over all of the pixels $\boldsymbol{u}$ in the base mesh s_0. The goal of AAM fitting is to minimize the expression in Eq.(3) simultaneously with respect to the shape parameters α and appearance parameters β.

Traditionally, Eq.(3) is solved by gradient decent method, which is computationally expensive. Recently, the computation cost of AAM fitting is greatly saved by the popular Inverse Compositional (IC) algorithm (i.e., IC-AAM) [2], in which the roles of the image and model are reversed and some time-consuming steps in AAM fitting are pre-computed and remain fixed in the iteration process. However, the IC-AAM only works well in the case that faces exhibit small amounts of variations.

2.3 Limitations of Active Appearance Model

Briefly, there are two main limitations for the conventional AAM, which are caused by the essentially non-linear variations of facial shape and appearance in a large data set.

Firstly, it is impossible to capture the complex non-linear shape and appearance variations of a large image set by just a single PCA model. Specifically, the learnt PCA model usually extracts the statistically significant features of the training set, ignoring some local details which are necessary for a more accurate alignment.

Secondly, in the AAM fitting procedure, the relationship between the appearance error (i.e., $\sum_{\boldsymbol{u}\in s_0}[A_0(\boldsymbol{u}) + \sum_{i=1}^{L}\beta_i A_i(\boldsymbol{u}) - I(W(\boldsymbol{u};\alpha))]$) and the shape parameter update (i.e., α) is assumed to be *close* to linear around the optimum, which is only right in the case that the face appearance variation is very small. However, it is hard to initialize the model parameters close enough to the correct alignment without any shape and appearance priors for the test image.

3 Locality-constrained Active Appearance Model

In this section, to address the limitations of conventional AAM, it is reformulated as a sparsity-regularized AAM problem, where sparse representation is exploited to model the non-linear face representation of AAM. Subsequently, the sparsity-regularized AAM is further approximated by our proposed Locality-constrained AAM algorithm for a fast implementation.

3.1 Sparsity-regularized Active Appearance Model

In this study, it is assumed that the shape and appearance of faces can be *sparsely* represented by faces of a large training set. Here, an explicit parametric model (e.g., PCA) is no longer needed. The aligned training shapes and the corresponding appearances are directly used as the shape and appearance bases. In this way, the original AAM is reformulated as the following sparsity-regularized

AAM problem:

$$\min_{\alpha,\beta}\{\sum_{\boldsymbol{u}\in s_0}[\sum_{i=1}^{N}\beta_i A_i(\boldsymbol{u}) - I(W(\boldsymbol{u};\alpha))]^2 + \lambda_1\|\alpha\|_{l^1} + \lambda_2\|\beta\|_{l^1}\} \tag{4}$$

where N is the number of faces in the training set, s_0 is the mean shape of all training faces, $A_i(\boldsymbol{u})$ is the "shape-normalized" appearance obtained by warped the facial image to the mean shape s_0, β_i is the ith appearance model parameter, α are the shape parameters, $I(W(\boldsymbol{u};\alpha))$ is the observed "shape-normalized" appearance of the input facial image, and λ_1, λ_2 are the regularization coefficients, which add sparse constraints to the shape and appearance parameters α and β respectively.

During the fitting procedure of sparsity-regularized AAM, the "analysis by synthesize" strategy is used to iteratively approximate the correct alignment. In each iteration, the "shape-normalized" facial appearance is first warped according to the current shape parameters α. With the fixed α and the observed appearance $I(W(\boldsymbol{u};\alpha))$, the appearance parameters β are calculated by solving a l_1 constrained optimization problem, Eq.(4). Then a new appearance is synthesized by the calculated appearance parameters β. To calculate the updated shape parameters, another l_1 constrained optimization problem (i.e., Eq.(4) with β and $\sum_{i=1}^{N}\beta_i A_i(\boldsymbol{u})$ fixed), which aims to minimize the discrepancy between the observed appearance and the synthesized appearance. The correct alignment can be obtained until convergence or a maximum iteration number is reached.

However, as pointed out in section 1, the sparsity-regularized AAM is very time-consuming due to the high dimension l_1 optimization problem and multiple iterations in the fitting procedure.

3.2 Approximated Sparsity by Locality Constraint

Inspired by the theory analysis and empirical observations by Yu et al. [12], the sparse coefficients α and β in Eq.(4) tend to be local and images with larger coefficients are more similar to the input image. So, to speed up the fitting procedure of sparsity-regularized AAM, locality constraint is introduced by us for a fast approximation.

Specifically, given an input image I to be aligned, the sparsity constraint in Eq.(4) is replaced with our locality constraint, which is as follows:

$$\min_{\alpha,\beta}\{\sum_{\boldsymbol{u}\in s_0}[\sum_{i=1}^{N}\beta_i A_i(\boldsymbol{u}) - I(W(\boldsymbol{u};\alpha))]^2 + \lambda_1\|\mathbf{d}\odot\alpha\|^2 + \lambda_2\|\mathbf{d}\odot\beta\|^2\} \tag{5}$$

where N is the number of faces in the training set, $A_i(\boldsymbol{u})$ is the appearance basis obtained by warping the train image to the mean shape s_0, β_i is the ith appearance model parameter, α are the shape parameters, and λ_1, λ_2 are the regularization coefficients. Specifically, $\mathbf{d} = [d_1, d_2, \ldots, d_N]$ are the distances between the input image I and the appearance bases $A(\boldsymbol{u}) = [A_1(\boldsymbol{u}), A_2(\boldsymbol{u}), \ldots, A_N(\boldsymbol{u})]$,

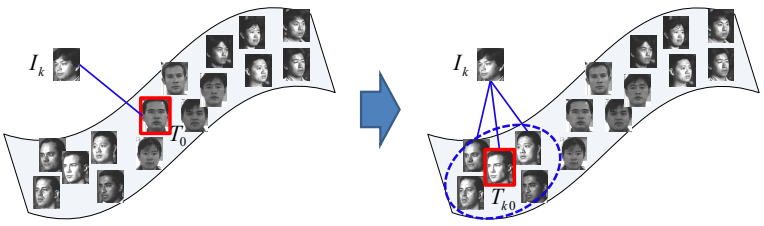


Fig. 1. Original AAM (*left*) vs. Locality-constrained AAM (*right*). T_0 is the mean template the whole training set, T_{k0} is the mean template of the K-nearest neighbors of input image I_k.

which add larger weights to the appearance bases nearer to I. Symbol $\odot$ denotes the element-wise multiplication.

In practice, Eq.(5) is fast implemented by directly setting the coefficients of the distant faces to be zero, i.e., selecting the K-nearest neighbors of I as the shape and appearance bases, which is formulated as:

$$\min_{\alpha,\beta} \sum_{\boldsymbol{u}\in\bar{s}_0} [\sum_{i=1}^{K} \beta_i A_i(\boldsymbol{u}) - I(W(\boldsymbol{u};\alpha))]^2 \tag{6}$$

where K is the number of selected nearest neighbors. It is important to note that $\bar{s}_0$ is the mean shape of the selected K-nearest neighbors.

Specifically, the optimization of Eq.(6) is similar to the conventional AAM-based fitting, which can be solved by gradient decent algorithm or the popular IC-AAM algorithm. Different from the conventional-AAM based methods, the appearance model used in our LC-AAM algorithm is localized linear, where the appearance variations are small and the relationship between the appearance error and the shape parameters is close to linear.

Figure 1 intuitively interprets why our LC-AAM algorithm works better than the original AAM. Specifically, in original AAM, the shape and appearance variations are characterized with a single Gaussian (PCA) model on the whole training set. However, the learnt PCA model cannot well characterize the non-linear feature of the whole training set. In comparison, our LC-AAM algorithm learns many localized linear models, which are specific for the input image, to approximate the global non-linear face model. Essentially, a stronger prior is learnt by exploring the K-nearest neighbors to constrain the shape and appearance subspace of input image. In addition, as illustrated in Figure 1, the input image I_k is closer to the mean template T_{k0} in LC-AAM, which is usually used to initialize the facial shape parameters, than the mean template T_0 in original AAM. So, it is much easier to converge to the correct alignment.

In our implementation, the neighbor samples are determined according to the measurement of the appearance similarity. Specifically, the face region is nor-

malized by the automatically detected two eye centers [15, 16] and the Euclidean distance is adopted to characterize the similarity of two faces. In addition, two kinds of feature descriptors are exploited to compute the similarity of faces, e.g., Histogram of Oriented Gradients (HOG) [17] feature and gray intensity value of image.

4 Experiments

In this section, we evaluate the effectiveness of our LC-AAM algorithm[1] through comparing it with the conventional AAM-based methods and other state-of-the-art face alignment methods.

4.1 Databases and Evaluation Metric

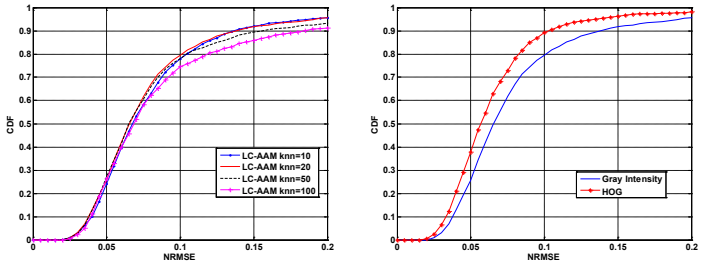
Our evaluation experiments are conducted on two publicly available face data sets. Set 1 is randomly collected from CMU-PIE [18], FRGCv1[19], FERET [20], CAS-PEAL [21], and PubFig [22] face databases, which covers large variations in pose, expression, lighting and image conditions, etc. Totally, there are more than 7000 images in set 1 and 1500 images are randomly selected for testing, noted as T1. Set 2 is collected from Cohn-Kanade face database [23], which includes 486 image sequences (8796 static images) in nearly frontal view from 97 subjects. Each sequence begins with a neutral expression and proceeds to a peak expression. To demonstrate that the generalization capability of our method, 80 subjects are randomly collected for training, and the other remaining 17 subjects (consisting of 1519 images, noted as T2) for testing.

For evaluation metric, the normalized root-mean-squared error (NRMSE) relative to the ground truth is adopted as the error measure for the face alignment. The NRMSE is given as a percentage, computed by dividing the root mean squared error by the distance between the two eye centers. The cumulative distribution function (CDF) of NRMSE is used to evaluate the performance of face alignment algorithm.

4.2 Experimental Results

Comparisons with respect to Different Number of Neighbors To show how our method is affected by the number of nearest neighbors. We conduct our experiments on T1 using K=10, 20, 50, 100 respectively. Moreover, the gray intensity is used to calculate the similarity of face images. From the experimental results, as shown in Figure 2(a), it is observed that K=20 is better than the other parameters. In addition, when a larger K is selected, the alignment results become worse. So, in the following experiments, the number of nearest neighbors is set to 20.

[1] The matlab code of our LC-AAM algorithm can be found at http://vipl.ict.ac.cn/members/xwzhao



(a) Effect of neighbor number. (b) Effect of feature descriptors.

Fig. 2. Performance variation of LC-AAM with different parameter configuration.

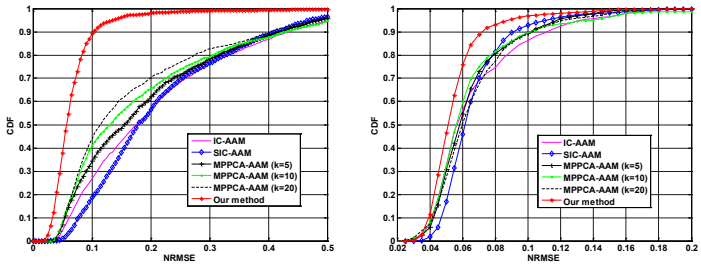
Comparisons with Different Feature Descriptors In this subsection, two kinds of feature descriptors are exploited to compute the similarity of image pairs: HOG feature and gray intensity. Specifically, for HOG feature, the block size, the number of orientation bins are set to 7×7 and 9 separately. The number of nearest neighbors is set to 20 in the comparison experiments.

The comparison experiments are conducted on T1. Experimental results, shown in Figure 2(b), demonstrate that the HOG feature is more suitable to calculate the similarity of faces than the gray intensity feature. The reason is that the HOG descriptor characterizes the boundary features of faces, eliminates some noise contained in gray pixel values.

Comparisons with AAM and its Variants In this subsection, we compare our algorithm with conventional AAM-based algorithms (e.g., IC-AAM, SIC-AAM, and MPPCA-AAM, etc.) on the above-mentioned two data sets. Specifically, the dimensionality of the shape and appearance models for IC-AAM and SIC-AAM is chosen by retaining 90% of the variance in the eigenvalues. For MPPCA-AAM, the number of Gaussian components is set to 5, 10, and 20 respectively. For our LC-AAM, the number of nearest neighbors is set to 20 and the HOG feature is exploited to calculate the similarity of image pairs. Moreover, the iteration number in the fitting procedure of all these methods is set to 25.

The comparison results are shown in Figures 3(a) and Figure 3(b), which demonstrate that our method greatly outperforms the conventional AAM-based methods. Specifically, in T1, which contains more pose variations than T2, our method achieves at least 40% higher than other methods when NRMSE is less than 0.1.

Comparisons with state-of-the-art Methods In this subsection, we compare our algorithm with the state-of-the-art face alignment algorithms, such as MPPCA-ASM [24], STASM [25], and CLM-CMU [26]. Specifically, the MPPCA-ASM method is a kind of ASM-based approach, which exploits the mixture of probabilistic principal component analysis (MPPCA) to model the shape variations. In our previous study, the MPPCA-ASM algorithm performs well when



(a) Comparison results on T1. (b) Comparison results on T2.

Fig. 3. Comparisons with original AAM and its variant methods.

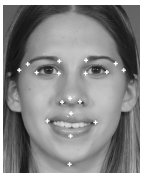


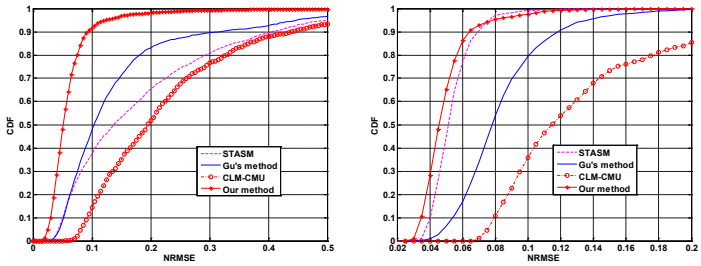
Fig. 4. Common facial landmarks of the evaluated methods.

the Gaussian component of MPPCA is set to 5, which corresponds to 5 different pose intervals in yaw direction. So, the number of Gaussian components is set to 5 in the following experiments. In our LC-AAM, the number of nearest neighbors is set to 20 and the HOG feature is adopted to calculate similarity of image pairs.

To perform a fair comparison, only the common 18 facial landmarks of these methods are used for performance evaluation, as shown in Figure 4. The comparison results are shown in Figures 5(a) and Figure 5(b) respectively. It is observed from the comparison results that our method greatly outperforms the other methods on T1. On T2, our method is comparable with STASM and is better than other methods. It is also interesting to note that the STASM algorithm performs well on near-frontal facial images (T2) but fails to work on images with large pose variation (T1). However, our method achieves consistently good alignment performance on these two test sets. Figure 6 shows some visualized example images of our method on challenging images.

5 Conclusion and Future Work

Instead of characterizing the facial shape and appearance distributions with a single PCA model, the original Active Appearance Model is reformulated as a sparsity-regularized problem. Then, based on the theory and empirical results of Yu et al., the sparse representation problem is further approximated by our Locality-constrained Active Appearance Model algorithm. For an input facial image, a strong shape and appearance prior is learnt by exploiting its K-similar



(a) Comparison results on T1. (b) Comparison results on T2.

Fig. 5. Comparisons with state-of-the-art methods.

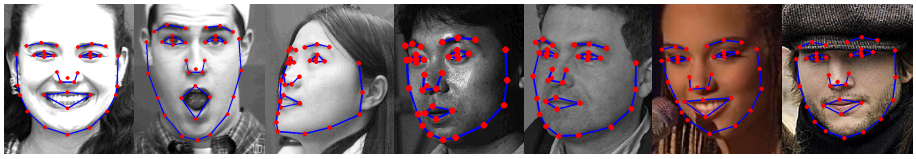


Fig. 6. Visualized example images of our method.

patterns. Essentially, the effectiveness of our LC-AAM stems from learning many localized linear face model instead of a global non-linear face model. By comparison, the localized face model is more compact and much easier to converge to the correct alignment. The effectiveness of our method is validated through comprehensively comparisons with the original AAM method, the variant methods of AAM, and the state-of-the-art face alignment methods on two publicly available face databases. In addition, our proposed method generalizes well to unseen subjects and images than the original AAM and its variants.

Acknowledgement. This work is partially supported by Natural Science Foundation of China (NSFC) under contract Nos. 61025010, 61173065, and 61001193; and the FiDiPro program of Tekes.

References

1. Cootes, T., Taylor, C., Cooper, D., et al.: Active shape models: Their training and application. CVIU **61** (1995) 38–59
2. Matthews, I., Baker, S.: Active appearance models revisited. IJCV **60** (2004) 135–164
3. Gross, R., Matthews, I., Baker, S.: Generic vs. person specific active appearance models. IVC **23** (2005) 1080–1093
4. Liu, X.: Generic face alignment using boosted appearance model. In: CVPR. (2007) 1–8
5. Wu, H., Liu, X.: Face alignment via boosted ranking model. In: CVPR. (2008) 1–8

6. Saragih, J., Goecke, R.: A nonlinear discriminative approach to AAM fitting. In: ICCV. (2007) 1–8
7. Cootes, T.F., Taylor, C.J.: A mixture model for representing shape variation. IVC **17** (1999) 567–573
8. van der Maaten, L., Hendriks, E.: Capturing appearance variation in active appearance models. In: CVPRWS. (2010) 34–41
9. Tipping, M.E., Bishop, C.M.: Mixtures of probabilistic principal component analyzers. Neural Computation **11** (1999) 443–482
10. Etyngier, P., Sgonne, F., Keriven, R.: Shape priors using manifold learning techniques. In: ICCV. (2007)
11. Zhang, S., Zhan, Y., Dewan, M., et al.: Towards robust and effective shape modeling: Sparse shape composition. Medical Image Analysis **16** (2012) 265–277
12. Yu, K., Zhang, T., Gong, Y.: Nonlinear learning using local coordinate coding. In: NIPS. (2009) 2223–2231
13. Wang, J., Yang, J., Yu, K., et al.: Locality-constrained linear coding for image classification. In: CVPR. (2010) 3360–3367
14. Edwards, G.J., Taylor, C.J., Cootes, T.: Interpreting face images using active appearance models. In: FG. (1998) 300–305
15. Zhao, X., Chai, X., Niu, Z., Heng, C., Shan, S.: Context constrained facial landmark localization based on discontinuous haar-like feature. In: FG. (2011) 673–678
16. Zhao, X., Chai, X., Niu, Z., Heng, C., Shan, S.: Context modeling for facial landmark detection based on non-adjacent rectangle (NAR) haar-like feature. Image and Vision Computing **30** (2012) 136–146
17. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR. (2005) 886–893
18. Sim, T., Baker, S., Bsat, M.: The cmu pose, illumination, and expression (PIE) database. In: FG. (2002)
19. Phillips, P., Flynn, P., Scruggs, T., et al.: Overview of the face recognition grand challenge. In: CVPR. (2005) 947–954
20. Phillips, P., Wechslerb, H., Huangb, J., et al.: The FERET database and evaluation procedure for face-recognition algorithms. IVC **16** (1998) 295–306
21. Gao, W., Cao, B., Shan, S., et al.: The CAS-PEAL large-scale chinese face database and baseline evaluations. SMC, Part A: Systems and Humans **38** (2008) 149–161
22. Kumar, N., Berg, A., Belhumeur, P., et al.: Attribute and simile classifiers for face verification. In: ICCV. (2009) 365–372
23. Tian, Y.I., Kanade, T., Cohn, J.F.: Recognizing action units for facial expression analysis. IEEE T-PAMI **23** (2001) 97–115
24. Gu, L., Kanade, T.: A generative shape regularization model for robust face alignment. In: ECCV:Part I. Lecture Notes in Computer Science. Springer-Verlag Berlin, Heidelberg (2008) 413–426
25. Milborrow, S., Nicolls, F.: Locating facial features with an extended active shape model. In: ECCV:Part IV. Lecture Notes in Computer Science. Springer-Verlag Berlin, Heidelberg (2008) 504–513
26. Saragih, J., Lucey, S., Cohn, J.: Deformable model fitting by regularized landmark mean-shifts. IJCV **91** (2011) 200–215