Progressive AAM Based Robust Face Alignment

Daehwan Kim, Jaemin Kim, Seongwon Cho, Yongsuk Jang, Sun-Tae Chung, and Boo-Gyoun Kim

***Abstract*—**AAM has been successfully applied to face alignment, but its performance is very sensitive to initial values. In case the initial values are a little far distant from the global optimum values, there exists a pretty good possibility that AAM-based face alignment may converge to a local minimum. In this paper, we propose a progressive AAM-based face alignment algorithm which first finds the feature parameter vector fitting the inner facial feature points of the face and later localize the feature points of the whole face using the first information. The proposed progressive AAM-based face alignment algorithm utilizes the fact that the feature points of the inner part of the face are less variant and less affected by the background surrounding the face than those of the outer part (like the chin contour). The proposed algorithm consists of two stages: modeling and relation derivation stage and fitting stage. Modeling and relation derivation stage first needs to construct two AAM models: the inner face AAM model and the whole face AAM model and then derive relation matrix between the inner face AAM parameter vector and the whole face AAM model parameter vector. In the fitting stage, the proposed algorithm aligns face progressively through two phases. In the first phase, the proposed algorithm will find the feature parameter vector fitting the inner facial AAM model into a new input face image, and then in the second phase it localizes the whole facial feature points of the new input face image based on the whole face AAM model using the initial parameter vector estimated from using the inner feature parameter vector obtained in the first phase and the relation matrix obtained in the first stage. Through experiments, it is verified that the proposed progressive AAM-based face alignment algorithm is more robust with respect to pose, illumination, and face background than the conventional basic AAM-based face alignment algorithm.

***Keywords*—**Face Alignment, AAM, facial feature detection, model matching.

I. INTRODUCTION

THE face alignment, whose objective is to localize facial feature points such as eye-brows, eyes, nose, mouth and contour of chin, etc., is important in face recognition, face based user interface, driver watching, animation, and other many applications needing face modeling [1]. Face shape is not a rigid but a deformable object, which is well known to be more difficult to detect than the rigid object, and thus face alignment has drawn a lot of research attentions [2,3,4,5,6].

AAM (Active Appearance Model) is one of the effective methods in detecting deformable 2D objects. AAM has been proposed first by Edwards et al. in [7], has later been extended by Cootes et al. [8,9], and Stegmann et al. [10]. Afterwards, a lot of research works about improvements of AAM have been reported [2,11,12].

Face alignment using AAM is relatively stable, but is known to be sensitive to initial values [13]. AAM tries to detect (localize) object (feature points) by fitting AAM model parameters so as to minimize the squared error between model texture represented by the parameters and the real object texture. Therefore, AAM fitting is a mathematical optimization problem. However, the search space is not convex even though the cost function is convex, so that finding the global minimum is not guaranteed [14]. That is, if the initial points do not start sufficiently near from the global minimum, in some cases the AAM fitting algorithm may converge to a local minimum so that face alignment using the AAM does not produce precise results. Under various face poses and illumination conditions, the contour of chin tends to be more variant than the inner parts of the face such as eyes, eyebrow, nose, mouth, and etc. Also, chin areas may be easily affected by the background of the face images. Therefore, for many new images with some different poses from the mean shape of the constructed AAM model and with complex background, the initial values for feature points in the contour of chin may not start sufficiently near from the global minimum (the true feature points in the chin contour of a new image). Since AAM fitting is to minimize the square of the whole error (at all facial feature points), large errors of some facial feature points will affect errors in other facial feature points. Thus, localizing relatively stable inner facial feature points can be affected by the feature points in the outer chin area when one tries to localize the whole facial feature points at the same time.

In this paper, we utilize the fact that feature points of the inner face are relatively less variant and less affected by the background surrounding the face than those in the outer part (like the chin contour), and propose a progressive AAM-based face alignment algorithm. The proposed progressive AAM-based face alignment algorithm first constructs two AAM models: the inner face AAM model, and the whole face AAM model. The inner face AAM model is constructed only from the inner facial feature points of the model faces and the whole face AAM model is constructed as usually, that is, constructed from the whole facial feature points of the model faces. In the fitting stage, the proposed algorithm aligns the face progressively through two phases. In the first phase, the feature parameters of the inner parts of a new face image are obtained by fitting the inner face AAM to the new face image. In the second phase, the initial values of the whole feature parameters of the new face image are estimated from the

Daehwan Kim, Jaemin Kim and Seongwon Cho are with A.I. Lab., Hongik University, Seoul, Korea.

Yongsuk Jang, Sun-Tae Chung and Boo-Gyoun Kim are with School of Electronic Engineering, Soongsil University, Seoul, Korea (e-mail: cst@ssu.ac.kr).

obtained feature parameters in the first phase using the least square method. Finally, the whole facial feature points are localized by fitting the whole face AAM to the new face image using the estimated initial values. It is observed and verified through experiments that the proposed progressive AAM-based face alignment is more robust with respect to poses, illumination, and background than the usual basic AAM-based face alignment [8,9,10].

Let us briefly describe some related works with our work. Li et al. [2] extends DAM (Direct Appearance Model) to multi-view face alignment. DAM does not combine the shape and textures like AAM by using the texture to directly predict the shape during the iterations of the parameters updates, which can lead to the improved convergence speed and accuracy. Jiao et al. [3] uses Gabor wavelet features to reinforce the ASM local searching in face alignment. Huang et al. [4] deals with face alignment under variable illumination by employing two forms of relatively lighting-invariant information: edge phase congruency and local intensity normalization. Xin and Ali [5] introduced component shape model in addition to global shape model in order to deal with various poses and expressions. To cope with computational problems, they employ Haar-wavelet features different from 1-D profile texture feature in ASM. Zhang et al. [6] trained local Haar-like feature texture classifiers using massive data sets to achieve robust face feature classifiers under various poses, expressions, and illuminations, and showed accuracy and robustness of their method. Matthews and Baker [11] proposed inverse compositional approach to improve the speed of AAM fitting, and Batur and Hayes [12] proposed adaptive AAM which improves speed and accuracy. Cristinacce et al. [13] suggested a multistage approach to facial feature detection, which detect facial feature points roughly using PRFR method, and next localize the facial feature points more precisely using a version of the AAM search. Main difference between all the above mentioned face alignment research works and our approach is that our proposed algorithm utilizes more robust set of local facial feature points like that of the inner facial feature points and uses the obtained local information for fitting the whole facial feature points.

The rest of the paper is organized as follows. The basic AAM is shortly reviewed in Section 2. In Section 3, our proposed progressive AAM-based face alignment algorithm is presented. Experiment results are given in Section 4. Finally the conclusion is given in Section 5.

II. BASIC AAM (ACTIVE APPEARANCE MODEL)

In this section, we briefly explain the basic AAM using face as an example of AAM application. This explanation is based on [8,9]. As stated in [8,9], AAM consists of two stages: Modeling stage and Fitting stage.

A. AAM Modeling

We assume the number of facial feature points of face is v, and the number of face model images is M. Then, the shape vector X of each face is defined as the vector consisting of v feature points, that is, $X = (x_1, y_1, x_2, y_2, ..., x_v, y_v)^t$ where (x_i, y_i) is a coordinate of i-th facial feature point and 't' means the transpose of vector (and matrix).

Here we also assume that each shape vector X is already normalized by Procrutes analysis [15].

Applying PCA analysis into all shape vectors from M face model images, one can obtain a mean shape S_0, and n characteristic mode vectors S_i $(i = 1, ..., n)$. Then, shape vector X of each model face or a new face can be represented as a linear combination of the mean shape S_0, and n characteristic shape mode vectors S_i $(i = 1, ..., n)$ as follows,

$$X = S_0 + \sum_{i=1}^{n} S_i b_i = S_0 + P_s \vec{b}_s \qquad (1)$$

(where $P_s = [S_1, ..., S_n]$, $\vec{b}_s = (b_1, ..., b_n)^t$).

In order to construct a statistical model of the gray-level texture, each model image is needed to be warped into the mean shape so as to let feature points match the mean shape using Delaunay triangulation algorithm [16]. After that, one samples gray-level information from the shape-normalized model images and normalizes it to remove global lighting effects and then forms a texture vector g. Likewise as in the shape vectors, PCA analysis is applied to all texture vectors from M face model images, and then one can obtain a mean normalized gray-level vector $\bar{T}$ and k characteristic texture mode vectors T_i $(i = 1, ..., k)$.

Likewise as in the shape vectors, texture vector g of each model face or a new face can be represented as a linear combination of the mean normalized gray-level vector T_0 and k characteristic texture mode vectors T_i $(i = 1, ..., k)$ as follows.

$$g = T_0 + \sum_{i=1}^{k} T_i c_i = T_0 + P_g \vec{b}_g \qquad (2)$$

(where $P_g = [T_1, ..., T_k]$, $\vec{b}_g = (c_1, ..., c_k)^t$).

Combining (1) and (2) can be represented using a common parameter vector $\vec{c}$ as follows [8,9].

$$X = S_0 + Q_s \vec{c}$$
$$g = T_0 + Q_g \vec{c} \qquad (3)$$

Modeling shape and texture as in (3) is called (combined) AAM.

B. AAM Fitting

Fitting AAM model (3) into a new image is an optimization problem.

By finding parameter vector $\vec{c}$ minimizing the squared error between the texture of a modeled face as (3) and a new image,

one can decide the face represented by (3) as the face to be aligned in the new image. And, facial feature points represented by $X = S_0 + Q_s\vec{c}$ are considered as the final feature points to be localized. How to find parameter vector $\vec{c}$ minimizing the squared error and how to set the initial parameter vector for fitting process depends on AAM algorithms. Usually, the initial parameter vector is set to match the mean shape.

III. PROGRESSIVE AAM-BASED FACE ALIGNMENT

A. Outline of Progressive AAM-Based Face Alignment

The progressive AAM-based face alignment algorithm proposed in this paper consists of two stages as in the basic AAM: Modeling and relationship derivation stage and Face Fitting stage.

1) Modeling and relation derivation stage
   a) Firstly, construct the inner face AAM model from the inner facial feature points of the model faces, which are more stable and less variant than outer parts of the face.
   b) Secondly, construct the whole face AAM model from the whole facial feature points of the model faces.
   c) Thirdly, derive the relation between the parameter vector of the inner face AAM model and the parameter vector of the whole face AAM model

2) Face Fitting stage
   a) Fit the inner face AAM model into a new incoming face, that is, find the inner face AAM parameter vector fitting the inner face AAM model.
   b) Estimate the initial values for the parameter vector of the whole face AAM fitting by using the inner face AAM parameter vector obtained in 1)-c).
   c) Fitting the whole face AAM model into a new incoming face by using the estimated initial values.

In this paper, we choose 94 facial feature points; 77 for the inner parts of face, and 17 for chin contour.

B. Estimation of the Initial Values for the whole AAM Parameter Vector

As stated in Section 3.1, the most important step in the proposed face alignment algorithm is how to estimate the initial values for the whole face AAM parameter vector from the result of fitting the inner face AAM model.

Using the fact that there exists a relationship between the inner AAM parameter vector and the whole AAM parameter vector, we solve the relationship via an optimization problem setting.

Suppose that the number of the elements of the inner face AAM parameter vector is p, and that the number of the elements of the whole face AAM parameter vector is q. Also suppose that the inner face AAM is modeled as follows.

$$\hat{X} = \hat{S}_0 + \hat{Q}_s\hat{\vec{c}}$$
$$\hat{g} = \hat{T}_0 + \hat{Q}_g\hat{\vec{c}} \tag{4}$$

Also, when each image of total M model face images is modeled as the inner face AAM like (4), let us denote the parameter vector of the inner AAM model as $\hat{\vec{c}}^i$ $(i = 1, ..., M)$

Again, when each image of total M face model images is modeled as the whole face AAM like (3), let us denote the parameter vector of the whole AAM model as $\vec{c}^i$ $(i = 1, ..., M)$.

Now, we know that $\vec{c}^i$ and $\hat{\vec{c}}^i$ has a relationship, and we assume that for $\vec{c}^i$ and $\hat{\vec{c}}^i$, the following linear relationship holds.

$$\vec{c}^i = R^i\hat{\vec{c}}^i \quad (i = 1, ..., M) \tag{5}$$

Moreover, if we additionally assume that R^i is almost constant, that is, $R^i = R$ for all $(i = 1, ..., M)$, then we can find an optimal R^o satisfying (5) for all $(i = 1, ..., M)$ by solving the following (6).

$$R^0 = \underset{R}{\arg\min} \| C_{tot} - RC_{int} \|^2 \tag{6}$$

(where $C_{int} = [\hat{\vec{c}}^1, ..., \hat{\vec{c}}^M]$, $C_{tot} = [\vec{c}^1, ..., \vec{c}^M]$).

Now, after we obtain the parameter vector $\hat{\vec{c}}$ fitting the inner face AAM for a new incoming face image, we can estimate the initial values $\vec{c}_0$ for the parameter vector fitting the whole face AAM model as follows.

$$\vec{c}_0 = R^0\hat{\vec{c}} \tag{7}$$

IV. EXPERIMENTS

A. Experiment Environments

In order to evaluate our proposed progressive AAM-based face alignment algorithm and compare it with the basic AAM based face alignment algorithm, we use 2 face databases in our experiments: domestic face database and IMM face database [17].

The first database, domestic face database consists of 450 images of 90 persons with 5 different poses. Each image is JPEG with a resolution of 640×480. Some face images of our domestic made face database are shown in Fig. 1.

Fig. 1 Some face images of the domestic face database

The second, the IMM face database consists of 240 images of 40 different human faces with 6 different poses or expression or illumination. Each image is JPEG with 640×480, and some face images of IMM database are shown in Fig. 2.

Fig. 2 Some face images of the IMM face database

The main difference between two databases is that the IMM face database has a simple background.

We construct the two AAM models for each database: the inner face AAM and the whole face AAM. The number of the face model images used for training for the domestic database is 65 (13 persons × 5 poses), 14.4% of total face 450 images. For the IMM database, the number of model images is 35 (7 persons × 5 poses), 14.6% of total 240 face images.

The number of the elements of the parameter vector is determined so as to take care of 95% of the energy of the image set: 66 for the inner face AAM parameter vector and 57 for the whole face AAM parameter vector.

B. Experiment Results

We decide the case that the all 94 final fitting feature points are within 3 pixels from the true facial feature points as success in face alignment, and otherwise as failure (refer to Fig. 3 and Fig. 4).

The experiment results for the IMM face database are summarized in Table I and Fig. 3 shows difference in face alignment results due to the basic AAM-based algorithm and our proposed algorithm. 180 face images among the IMM database, which excludes some distracted pose images (1 per each person) and gray-level images, are used for testing.

Fig. 3 Face alignment results, (a) basic AAM-based one (b) the proposed progressive AAM-based one (IMM DB)

Also, the experiment results for the domestic face database are summarized in Table II and Fig. 4 shows difference due to both the basic AAM-based algorithm and our proposed algorithm.

TABLE I
COMPARISON BETWEEN THE BASIC AAM-BASED FACE ALIGNMENT ALGORITHM ANDTHE PROPOSED PROGRESSIVE AAM-BASED FACE ALIGNMENT ALGORITHM (IMM FACE DB)

Algorithm	Num. of test images	Num. of success	Rate of success (%)
Basic AAM	180	168	93.3
Progressive AAM	180	173	96.1

TABLE II
COMPARISON BETWEEN THE BASIC AAM-BASED FACE ALIGNMENT ALGORITHM AND THE PROPOSED PROGRESSIVE AAM-BASED FACE ALIGNMENT ALGORITHM (DOMESTIC FACE DB)

Algorithm	Num. of test images	Num. of success (%)	Rate of success (%)
Basic AAM	450	401	89.1
Progressive AAM	450	439	97.5

Experiment results in Table I and II show that our proposed progressive AAM-based face alignment algorithm performs better especially in the case of the domestic face database case, which have more complex background than the IMM face database, compared with the case of the IMM face database. That is, the basic AAM may work relatively well for the case of frontal view face images with simple background, but its performance may deteriorate even for the frontal view but with complex background, for which our proposed progressive AAM-based face alignment algorithm still works robustly.

Fig. 4 Face alignment results, (a) basic AAM-based one (b) the proposed progressive AAM-based one (Domestic DB)

V. CONCLUSION

The experiments in this paper show that the proposed progressive AAM-based face alignment, which fits the more stable facial feature point set first and fits the less stable ones using the first stable parameters, works effectively.

We are currently testing the proposed progressive AAM-based face alignment under more various illuminations,

poses, and expressions, and preparing to compare it with other prominent face alignment algorithms like [6,11,12].

REFERENCES

[1] S. Z. Li and A. K. Jain, *Handbook of Face Recognition*, Springer, 2004.

[2] S. Z. Li, Y. S. Cheng, H. J. Zhang, Q. S. Cheng "Multi-view face alignment using direct appearance models," Automatic Face and Gesture Recognition, 2002. Proceedings. *Fifth IEEE Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 309 – 314, May 2002.

[3] F. Jiao, S. Z. Li, H-Y. Shum, D. Schuurmans, "Face Alignment Using Statistical Models and Wavelet Features," *Proc. 2003 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Vol.1, pp. I-321 - I-327, June, 2003.

[4] Y. Huang, S. Lin, S.Z. Li, H. Lu, H.-Y. Shum, "Face Alignment Under Variable Illumination," *Proc. Sixth IEEE Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 85-90, May 2004.

[5] S. Xin and H. Ai, "Face Alignment under Various Poses and Expressions," ACII2005, LNCS 3784, pp. 40-47, 2005.

[6] L. Zhang et al., "Robust Face Alignment Based on Local Texture Classifiers," ICIP 2005. *IEEE Int'l Conf. on Image Processing*, Vol. 2, pp.354-357, Sep. 2005.

[7] G. Edwards, C. J. Taylor, and T. F. Cootes, "Interpreting Face Images using Active Appearance Models," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, pp. 300-30, 1998.

[8] T. F. Cootes, G. J. Edwards, C. J. Taylor, H. Burkhardt, and B. Neuman, "Active Appearance Models," in *Proc. Eur. Conf. Computer Vision*, vol. 2, pp. 484–498, 1998.

[9] T. F. Cootes, D. J. Edwards, and S. J. Taylor, "Active Appearance Models," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 23, no. 6, pp. 681–685, Jun. 2001..

[10] M. B. Stegmann, B. K. Ersboll, R. Larsen, "FAME -- A Flexible Appearance Modelling Environment," IEEE Transactions on Medical Imaging, Vol. 22, Iss.10, pp. 1319 – 1331, Oct. 2003.

[11] I. Matthews and S. Baker, "Active Appearance Models Revisited," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 135 – 164, Nov. 2004.

[12] A.U Batur and M.H. Hayes, "Adaptive Active Appearance Models**,"** *IEEE Transactions on Image Processing*, Vol. 14, Issue 11, pp. 1707 – 1721, Nov. 2005.

[13] D. Cristinacce, T. Cootes, and I. Scott, "A Multi-Stage Approach to Facial Feature Detection," *Proc. British Machine Vision Conference 2004*, Vol.1, pp.277-286.

[14] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press

[15] J. C. Gower, "Generalized Procrustes Analysis," Psychometrika, 40:33--51, 1975.

[16] D. T. Lee and B. J. Schachter, "Two Algorithms for Constructing a Delaunay Triangulation," *Int. J. Computer Information Sci*. 9, pp.219-242, 1980.

[17] IMM face database and AAM-API, http://www2.imm.dtu.dk/~aam/