Reliable Face Alignment Using Two-Stage AAM

Sunho Ki, Daehwan Kim, Seongwon Cho, Sun-Tae Chung, Jaemin Kim, Yun-Kwang Hong, Chang Joon Park, Dongmin Kwon, Minhee Kang, Yusung Kim, and Younghan Yoon

***Abstract*—**AAM (active appearance model) has been successfully applied to face and facial feature localization. However, its performance is sensitive to initial parameter values. In this paper, we propose a two-stage AAM for robust face alignment, which first fits an inner face-AAM model to the inner facial feature points of the face and then localizes the whole face and facial features by optimizing the whole face-AAM model parameters. Experiments show that the proposed face alignment method using two-stage AAM is more reliable to the background and the head pose than the standard AAM-based face alignment method.

***Keywords*—**AAM, Face Alignment, Feature Extraction, PCA.

I. INTRODUCTION

PRECISE and reliable facial feature localization is necessary for applications such as face recognition or face modeling [1]. It is well known that precision of detecting facial feature points affects success rate of face recognition [2]. Thus, face alignment, whose objective is to localize facial feature points such as eye-brows, eyes, nose, mouth and contour of chin, etc., is important for achieving success in face recognition and other many applications needing face modeling. Face shape is not a rigid but a deformable object, which is more difficult to detect than the rigid object, and thus face alignment has drawn a lot of research attentions [3, 4, 5, 6, 7].

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 [8], has later been extended by Cootes et al. [9, 10]. Afterwards, a lot of research works about improvements and extensions of AAM have been reported [3, 11, 12].

Face alignment using AAM is relatively stable, but is known to be sensitive to initial values [13]. AAM tries to localize object (feature points) by fitting AAM model parameters so as to minimize the squared error between the 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, the AAM fitting algorithm may converge to a local minimum so that face alignment using the AAM does not produce precise results. 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). Thus, prediction of good initial values which are located sufficiently near from the global minimum is significant for producing precise face alignment.

For reliable AAM-based face alignment, we can utilize the following two observations. First, the inner AAM model fitting is less sensitive to the background and the head pose. That is, we know that under various head poses and illumination conditions, the inner parts of the face such as eyes, eyebrow, nose, mouth, and etc. tends to be less variant and less affected by the background of the face images than the outer parts of the face such as the contour of chin. Second, the dominant parameters (corresponding to large eigenvalues) of the whole face-AAM model are highly correlated with those of the inner face-AAM model. Based on these two observations, we propose two-stage AAM approach in reliable face alignment, where in first stage fits more stable inner parts of a new face, and in the second stage, fits the whole face of the new face but with initial values predicted using the least square method from the estimated inner face-AAM parameters in the first stage. As opposed to the initial parameters of the whole face-AAM in the standard AAM which are usually set by the mean shape, the initial parameters of the whole face-AAM in our proposed two-stage AAM approach may be rendered closer to the global minimum since they are predicted from the more stable inner face-AAM parameters estimated in the first stage inner face-AAM fitting.

It is observed through experiments that the proposed two-stage AAM approach in face alignment is more reliable with respect to head poses, illuminations, and background than the usual standard AAM-based face alignment.

Main difference between all the above mentioned face alignment research works and our approach is that our proposed method utilizes more stable local facial feature points

Manuscript received April 30, 2008. This work was supported in part by the Korean Government (Small and Medium Business, Ministry of Information and Communication) Research Fund, the BK21 Project and Soongsil University Research Fund.

Sunho Ki, Daehwan Kim, Dongmin Kwon, Minhee Kang, Yusung Kim and Younhan Yoon are with the Electrical Information Control Department, Hongik University, Korea (e-mail: Suno_ki@hotmail.com).

Seongwon Cho is with School of Electronic and Electrical Engineering, Hongik University, Korea (e-mail: swcho59@hanmail.net).

Sun-Tae Chung is with Soongsil University, Korea. (e-mail: cst@ssu.ac.kr).

Jaemin Kim is with School of Electronic and Electrical Engineering, Hongik University, Korea (e-mail: jmkim@hongik.ac.kr).

Yun-Kwang Hong and Chang Joon Park are with Game S/W Technology Research Team, Digital Contents Division, ETRI, Korea (e-mail: ykhong@etri.re.kr).

like the inner facial feature points and extends the obtained local information for fitting the whole facial feature points.

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

Fuzzy logic control (FLC) has been widely applied to industries [1][2][3][4][5][6][7]. These various applications require convenient method that construct fuzzy controller. Thus, the automatic construction of rules has been an important research problem [8][9][10][11][12][13]. Expert knowledge is used to construct knowledge base for rule-based system. In other words, in the knowledge acquisition step, an expert express his knowledge in IF-THEN rules. When an expert is not available, rules can be inferred from input-output relation data.

Differently from non-fuzzy system, fuzzy systems need fuzzy sets used in them. To develop such systems, the membership functions for fuzzy sets as well as fuzzy rules should be defined. Fuzzy rule-based control system has an advantage that knowledge is represented in the form of IF-THEN rules, so knowledge can be formulated as linguistic form in comparison to neural network.

In this paper, we present a new method for the automatic construction of rule-base needed for fuzzy control system based on fuzzy neural network and object function when input-output relation data are not available. Also, we propose a neural network and its learning algorithm to fine-tune the parameters used in the control system.

II. Standard AAM (Active Appearance Model)

In this section, we briefly explain the standard AAM using face as an example of AAM application. This explanation is based on [9, 10]. As stated in [9, 10], 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 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 the *i-th* facial feature point and the superscript t means transpose of vector (and matrix). Here we also assume that the each shape vector X is already normalized by Procrustes analysis [15].

Applying PCA analysis into all shape vectors from M face model images, the shape vector X of each face can be represented as a linear combination of the mean shape S_0, and n characteristic shape mode vectors.

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 (intensity) from the shape-normalized image and normalizes it to remove global lighting and then forms a texture vector g. As in the shape vectors, PCA analysis is applied to for all texture vectors from M face model images. The texture vector g of each face can be represented as a linear combination of the mean normalized gray-level vector T_0 and k characteristic texture mode vectors.

In order to reduce the number of parameters, the shape and texture parameters are combined as

$$\begin{aligned} X &= S_0 + Q_s\vec{c} \\ g &= T_0 + Q_g\vec{c} \end{aligned} \tag{1}$$

where $\vec{c}$ is a common parameter vector [9, 10]. Modeling shape and texture as in (1) is called a (combined) AAM.

B. *AAM Fitting*

Fitting AAM model (1) 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 (1) and a new image, one can decide the face represented by (1) 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, that is, $\vec{c}_0 = \vec{0}$.

III. Two-Stage AAM Approach in Face Alignment

In this section, we briefly explain the standard AAM using face as an example of AAM application. This explanation is based on [9, 10]. As stated in [9, 10], AAM consists of two stages: Modeling stage and Fitting stage.

A. *Outline of Face Alignment Using Two-Stage AAM*

The face alignment method using two-stage AAM proposed in this paper consists of two parts as in the standard AAM: Modeling part and Face Fitting part.

1) Modeling part
   Construct two AAM model: the inner face-AAM model, and the whole face-AAM model. The inner face-AAM model is constructed from the inner facial feature points of the model faces, which are more reliable and less variant than outer parts of the face. And, the whole face-AAM model is constructed as usual in standard AAM, that is, from the whole facial feature points of the model faces.

2) Face Fitting part
   a) First, fit the inner face-AAM model into a new incoming face, that is, estimates the inner face-AAM parameters fitting the inner face-AAM model.

b) Second, predict the initial values for the dominant parameters of the whole face-AAM fitting from the estimated values of the inner face-AAM parameters.
c) Third, fitting the whole face-AAM model into a new incoming face by using the predicted initial values.

B. How to Predict the Initial Values for the whole AAM Parameters

As stated in Section III-A, the most important process in the proposed two-stage face alignment method is how to predict the initial values for the whole face-AAM parameters from the estimated values of the inner face-AAM parameters.

For that, we need to seek the relation between the parameters of the inner face-AAM model and the parameters of the whole face-AAM model.

Since the set of the whole facial feature points includes the inner facial feature points, obviously one can argue that the dominant parameters (corresponding to large eigenvalues) of the whole AAM model are highly correlated with those of the inner AAM model, which leads to a relationship between the inner AAM parameters and the whole AAM parameters. Then, we seek the relation in an optimization problem setting as follows.

The inner face-AAM is modeled as follows as the standard AAM (1).

$$\hat{X} = \hat{S}_0 + \hat{Q}_s \hat{\vec{c}}$$
$$\hat{g} = \hat{T}_0 + \hat{Q}_g \hat{\vec{c}} \tag{2}$$

Also, when each image of total M model face images is modeled as the inner face-AAM as the standard AAM (1). Let us denote the parameter vectors 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 as the standard AAM (1), let us denote the parameter vectors 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.

$$\vec{c}^{\,i} = R^i \hat{\vec{c}}^{\,i} \quad (i = 1, ..., M) \tag{3}$$

If we represent $\hat{\vec{c}}^i \ (i = 1, ..., M)$ and $\vec{c}^i \ (i = 1, ..., M)$ as in matrix form C_{inner} and C_{tot} as follows,

$$C_{\text{inner}} = [\hat{\vec{c}}^1, ..., \hat{\vec{c}}^M], \ \ C_{\text{tot}} = [\vec{c}^1, ..., \vec{c}^M] \tag{4}$$

then, equation (3) can be represented as the following equation.

$$C_{\text{tot}} = RC_{\text{inner}} \tag{5}$$

Now, one can obtain the optimal matrix R^0 satisfying equation (5) by solving the following optimization problem (6).

$$R^0 = \arg\min_R \| C_{\text{tot}} - RC_{\text{inner}} \|^2 \tag{6}$$

That is, R^0 is the R which minimizes $\| C_{\text{tot}} - RC_{\text{inner}} \|^2$.

After we obtain the parameters $\hat{\vec{c}}$ of the inner face-AAM fitting for a new incoming face image, we can predict the initial values $\vec{c}_0$ for the parameters of the whole face-AAM fitting as follows.

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

IV. Experiments

A. Experiment Environments

In order to evaluate our proposed face alignment method and compare other research results, we use two face databases in our experiments: IMM face database [17], and a domestic face database. The IMM face database consists of 240 images of 40 different human faces with 6 different poses, expression, or illumination. Each image is JPEG of 640×480 size.

The IMM database includes smile expression image, but those images (1 per each person) are excluded and only 200 face images among the IMM database are used for testing.

The domestic face database consists of 415 images of 83 persons with 5 different poses or illuminations. Each image is JPEG with a resolution of 640×480. Some face images of our domestic face database are shown in Fig. 1.

Fig. 1 Some face images of the domestic face database

The main difference between two databases is that the domestic face database has a more complex background and more variant illumination conditions.

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 in the IMM database is 35 (7 persons 5 poses / illuminations), 17.5% of total 200 face images. For the domestic face database, the number of model images is 150 (30 persons 5 poses / illuminations), 36% of total 415 face images.

We select 58 facial feature points in total for the IMM DB: 45 for inner facial feature points and 13 feature points for chin areas, and select 94 facial feature points in total for the domestic DB: 77 for inner facial feature points and 17 feature points for chin areas. The reason for selecting more facial feature points for the domestic DB is that the domestic DB has more complex background and more variant illumination conditions.

The number of the elements of the parameter vector is determined so as to take care of 95% of the energy of the image data set: 6 for the inner face-AAM parameter vector and 7 for the whole face-AAM parameter vector in the IMM face database, and 47 for the inner face-AAM parameter vector and 44 for the whole face-AAM parameter vector in the domestic face database.

The specifications of PC used for our experiments are as follows: Intel Core2Duo Conro E6600 (4MB L2 Cache, 2.4GHz@3.24GHz) and 2GB (DDR2-880) main memory. We used one CPU for our experiment and used Intel C++ compiler 9.1.

B. Experimental Results

For our experiments, we finish face alignment process when fitting error is less than 0.0001, and we decide the case where the all 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. 2 and Fig. 3). The processing time measures the interval between the time when all trained data and a test image data is loaded into memory and the time when face alignment process is finished.

The experiment results for the IMM face database are summarized in Table I. Fig. 2 shows two face alignment results: the standard AAM-based face alignment method and our proposed two-stage AAM-based face alignment method.

TABLE I
COMPARISON BETWEEN THE STANDARD AAM AND THE PROPOSED METHOD (IMM FACE DB)

Method	Num. of test images	Success rate(%)	Processing time(ms)
Standard AAM	200	92.0	24.3
Proposed method	200	95.5	35.4

(a)

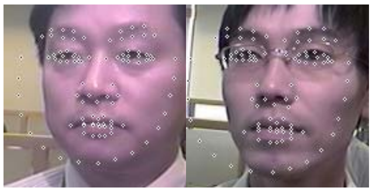
(b)

Fig. 2 Face alignment results, (a) Standard AAM (b) Proposed method (IMM DB)

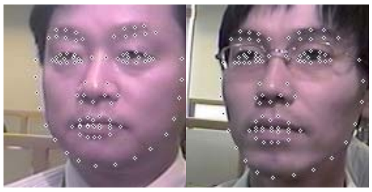
The experiment results for the domestic face database are summarized in Table II. Fig. 3 shows two face alignment results: the standard AAM-based face alignment method and our proposed two-stage AAM-based face alignment method.

TABLE II
COMPARISON BETWEEN THE STANDARD AAM AND OUR PROPOSED METHOD (DOMESTIC DATABASE)

Method	Num. of test images	Success rate(%)	Processing time(ms)
Standard AAM	415	84.6	95.7
Proposed method	415	94.9	175.9



(a)



(b)

Fig. 3 Face alignment results, (a) Standard AAM (b) Proposed method (Domestic DB)

Tables I and II of experiment results show that our proposed two-stage AAM-based face alignment method performs better than the standard AAM-based face alignment. Our proposed method works much better for the domestic face database, which have more complex background than the IMM face database, compared with the IMM face database. That is, the standard AAM may work relatively well for 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 two-stage AAM-based face alignment method still works well.

The reason why processing time for the domestic DB takes longer than that for the IMM DB is that the experiments for the domestic DB have more modeling parameters, and more facial feature points to detect and use more modeling images, all of which affect processing time, than those for the IMM DB.

V. CONCLUSION

The experiments in this paper shows that the proposed face alignment method using two-stage AAM, which first fits more stable parameters, the inner face-AAM model parameters and then fits less stable parameters, the whole face-AAM model parameters using the stable parameters, works effectively. Current implementation of our proposed method incurs more processing time in face alignment even if it does not cause not more than twice as big as the standard AAM. When we can remove redundancy in using two-stage AAM, the processing time will be reduced.

In the future, we will test the proposed face alignment method using two-stage AAM under more various illuminations, poses, and expressions, and compare it with other prominent face alignment method as in [7, 11, 12].

ACKNOWLEDGMENT

We thank Mikkel B. Stegmann for kind permission to use AAM-API Software [17]. This work was supported in part by the Korean Government (Small and Medium Business, Ministry of Information and Communication) Research Fund, the BK21 Project and Soongsil University Research Fund.

REFERENCES

[1] Li, S.Z., Jain, A.K.: Handbook of Face Recognition. Springer (2004).

[2] Shan, S., Chang, Y., Gao, W., Cao, B., Yang, P.: Curse of Mis-alignment in Face Recognition: Problem and a Novel Mis-alignment Learning Solution, Proc. 6th IEEE Conf. on Automatic Face and Gesture Recognition, (2004) 314 – 3203.

[3] Li, S.Z., Cheng, Y. S., Zhang, H. J., Cheng, Q. S.: Multi-View Face Alignment Using Direct Appearance Models. Proc. 5th IEEE Conf. on Automatic Face and Gesture Recognition, (2004) 309-314.

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

[5] Huang, Y., Lin, S., Li, S.Z., Lu, H., Shum, H.-Y.: Face Alignment Under Variable Illumination. Proc. 6th IEEE Int'l Conf. on Automatic Face and Gesture Recognition, (2004) 85-90.

[6] Xin, S., Ai, H.: Face Alignment under Various Poses and Expressions. ACII2005, LNCS 3784, (2005) 40-47.

[7] Zhang, L., Haizhou A., Xin, Huang, C., Tsukiji, S.: Robust Face Alignment Based on Local Texture Classifiers. IEEE Conf. on Image Processing, Vol. 2, (2005) 354-357.

[8] Edwards, G., Taylor, C. J., Cootes, T. F.: Interpreting Face Images using Active Appearance Models. Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, (1998) 300-305.

[9] Cootes, T. F., Edwards, G. J., Taylor, C. J., Burkhardt, H., Neuman, B.: Active appearance models. Proc. Eur. Conf. Computer Vision, Vol. 2, (1998) 484-498.

[10] Cootes, T. F., Edwards, G. J., Taylor, C. J.: Active Appearance Models. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 23, no. 6, (2001) 681-685.

[11] Matthews, I., Baker, S.: Active Appearance Models Revisited. Int. Journal of Computer Vision, Vol. 60, No. 2, (2004) 135 – 164.

[12] Batur, A.U., Hayes, M.H.: Adaptive Active Appearance Models. IEEE Trans. Image Processing, Vol. 14, Issue 11, (2005) 1707- 1721.

[13] Cristinancce, D., Cootes, T. , Scott, I.: A Multi-Stage Approach to Facial Feature Detection. Proc. British Machine Vision Conference, Vol.1, (2004) 277-286.

[14] Boyd. S., Vandenberghe, L.: Convex Optimization. Cambridge University Press.

[15] Gower, J. C.: Generalized Procrustes analysis. Psychometrika, Vol. 40, (1975) 33-51.

[16] Lee, D. T., Lee, D. T., Schachter, B. J.: Two Algorithms for Constructing a Delaunay Triangulation. Int. J. Computer Information Sci. Vol.9, (1980) 219-242.

[17] DTU(Technical university of Denmark) IMM face database,http://www2.imm.dtu.dk/~aam/.