Probabilistic Active Appearance Models

Y. Zhang[1] Y. Benhamza[2] K. Idrissi[1] C. Garcia[1]

[1] Universite de Lyon, CNRS

INSA-Lyon, LIRIS, UMR CNRS 5205, F-69621, France

[2] Laboratoire LARATIC, INTTIC Oran Algeria

{Yuyao.Zhang, Khalid.Idrissi, Christophe.Garcia}@insa-lyon.fr

ybenhamza@ito.dz

Abstract

Active Appearance Models (AAM) allow to synthesize novel images by linear combinations in a statistic model which is built by applying Principal Component Analysis (PCA) on a training dataset. The model of PCA essentially assumes that the observed variables are gathered in an uncorrelated way into vectors, but most of the time it is not the case. In this paper, we present a probabilistic method based on the analysis of the correlation between observed variables, which can better follow the distribution of data, in order to improve Principal Component Analysis used for feature extraction in Active Appearance Models. The proposed approach, called Probabilistic Active Appearance Models *(PAAM) is proven to be more adaptive to the variable distribution. As a consequence, images of faces illuminated from different directions can be more robustly be synthesized by PAAM. The experimental results show that the proposed method provides higher accuracy than classical Active Appearance Model for face alignment in a point-to-point error sense.*

Mots clefs

Active Appearance Model, PCA, Probability

1 Introduction

Face recognition has been a well investigated topic in image processing and computer vision. In the last decade, large efforts have been done in searching for a face recognition system that is capable of working with "real-world" faces. Among these efforts, Active Appearance Model, first proposed in [1], is a non-linear, generative, and parametric model of a certain visual phenomenon [2], have been proved to be a good statistic tool to model shape and appearance variation of objects. In particular, AAM is extensively applied to the detection, tracking and analysis/synthesis of face images.

AAM uses Principal Component Analysis (PCA) to model shape and appearance variations across pose, expression, illumination and identity. However, the latent assumption of PCA does not always hold true. For example, to model pose changes with 2D models is a non linear problem and a common alternative method which can be preferred like kernel PCA, first introduced by B. Scholkopf et al. [3], which is a non-linear model to account for different pose changes and also expression, illumination variance of frontal faces. In [4], Romdhani et al. use KPCA with point distribution models to model faces with large variation of out-of-plane motions. However, KPCA is only applied on shapes, and is not embedded into texture model. Moreover, it suffered from a rough approach of the reconstruction problem. In [5],B. Moghaddam et al. introduced an unsupervised technique called Probabilistic Appearance Models, for visual learning which is based on density estimation in high-dimensional spaces using an eigenspace decomposition. A probabilistic function is used to reconstruct the image which achieve the highest likelihood of the original image. R. Hamdan et al. presented a low complexity approximation of Probabilistic Appearance Models in [6].To estimate the correlations between samples in training set, a Karhunen-Loeve transform is applied to reduce the dimension of training images. More recently, Miller et al. [7] have learned an entropy measure to align images with respect to the distribution of the data.

In this paper, a novel statistic approach is proposed to build the shape, texture and appearance models of AAM. In order to model the distributions of the face features, a probabilistic covariance matrix, which is proved to be more adaptive than the standard covariance matrix, is constructed and used in the Active Appearance Model. Our method outperforms standard AAM especially in the case of complex illumination and non frontal pose of faces.

The organization of this paper is as follows. Section 2 presents the way of constructing the probabilistic covariance matrix. Section 3 presents experimental results assessing the performance of the proposed algorithm in comparison with AAM. Section 4 concludes the paper.

2 Distribution based principal component analysis

A standard Active Appearance Model explains novel images by linear combination of statistic models which are built by applying Principle Component Analysis on training data. Therefore, PCA is not designed to extract non-linear features from the shape and texture of the non frontal or non uniformly illuminated faces. In general, both illumination and pose variations remain to be difficult to handle in face recognition.

We consider that a distribution based principal component analysis method is more appropriate for handling the multiple variations which are caused by the changes of light source. The following subsections aim at presenting the original AAM method, and the proposed Probabilistic Active Appearance Model method.

2.1 Active Appearance Model

Active Appearance Model is an algorithm which allows to generate a synthetic image as close as possible to a particular target image by making use of constraints of the appearance models. An appearance model is combined by two linear subspaces, one for the shape and one for the texture which are both learnt from a labelled set of training images [1].

Interpreting a novel image is an optimization problem in which the method minimizes the difference between a new image and one synthesized by the appearance model. The difference vector δI can be defined :

$$\delta I = I_{\mathrm{i}} - I_{\mathrm{m}} \tag{1}$$

Where I_{i} is the vector of grey-level values of the pixels in the image, I_{m} is the vector of grey-level values reconstructed from the current model parameters.

This method proceeds in three steps :

I) Principal Component Analysis (PCA) is applied respectively on the shape training database and a shape-free texture training database. PCA created the statistical shape and texture model as the follows :

$$s = \bar{s} + Q_{\mathrm{s}} b_{\mathrm{s}} \tag{2}$$

$$t = \bar{t} + Q_{\mathrm{t}} b_{\mathrm{t}} \tag{3}$$

Where s ans t are the vectors of shape and texture from one face. Q_{s} represents eigenvectors extracted from shape database while b_{s} is a vector controlling the shape variance. Q_{t} represents eigenvectors extracted from texture database , and b_{t} is a vector controlling the texture variance.

Another PCA is then applied on the samples of vector b, which is combined to b_{s} so as to construct the appearance parameter c :

$$b = Q * c \tag{4}$$

Where Q is the matrix of PCA eigenvectors, c is a vector controlling both b_{s} and b_{t} at the same time.

II)A experiment matrix in which each control parameter c is disturbed with known values and the residuals of each displacement in each image is measured to build a relationship between the parameter variation and the image differences δI. This relationship is :

$$\delta c = R * \delta I \tag{5}$$

Where R is the experiment matrix, δc and δI represent the parameter variation and the image differences respectively.

III) The fitting procedure in which by varying the model parameters c, the magnitude of the difference vector $\Delta = (\delta I)^2$ is minimized to locate the best match between model and image .

2.2 Proposed approach

Latent assumptions of PCA. PCA is perhaps one of the oldest and best-known methods in multivariate analysis and data mining. It was first introduced by Pearson [8], who used it in a biological framework. Now it is mostly used as a tool in exploratory data analysis and for making predictive models. Let $X = [x_1 x_2 \cdots x_n]$ be a matrix $X^{d \times n}$, where each column x_i is a data sample, N is the number of training samples, and D is the number of observed variables (coordinates of points or pixel values). The principal components can be derived from several criteria. Two criteria - minimal reconstruction error and maximal preserved variance - are described in the following two subsections. A third criterion corresponding to distance preservation (metric multidimensional scaling) is also described.

The limitation of PCA is caused by several latent assumptions of the data. First, the PCA model essentially assumes that the observed variables in vector x_i follow a Gaussian distribution. Second, all the observed variables are randomly gathered into vector x_i, that means these elements are uncorrelated (no linear dependencies behind them). Third, the observed variables must be centered.

Probability covariance matrix. In order to construct a more adaptive model, the statistical properties of the data should agree with the PCA latent assumptions. However, one of the three assumptions is frequently ignored. When the observed variables contained in the vector are related, instead of using covariance matrix $C_{xx} = E\left\{xx^T\right\}$, a probabilistic covariance matrix is more appropriate for describing the data. The probabilistic covariance matrix can be defined as [9] :

$$Cov(k,l) = \sum_{i=1}^{d} \sum_{j=1}^{d} (k_i - \mu_k)(l_j - \mu_l) P(k_i) P(l_j) \tag{6}$$

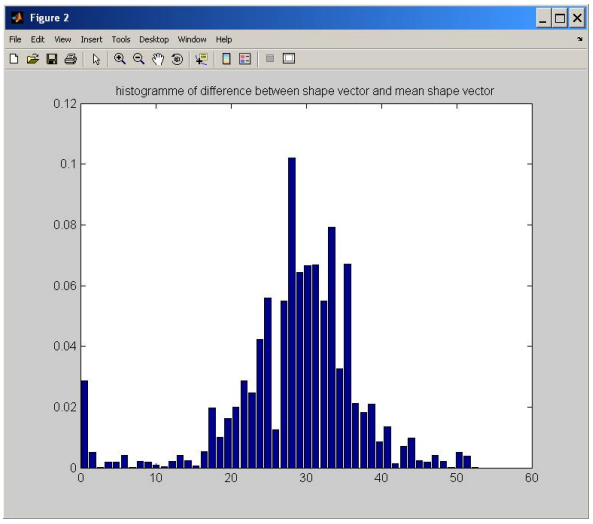

Figure 1 – *Histogram of Euclidean distances between elements of shape vector*

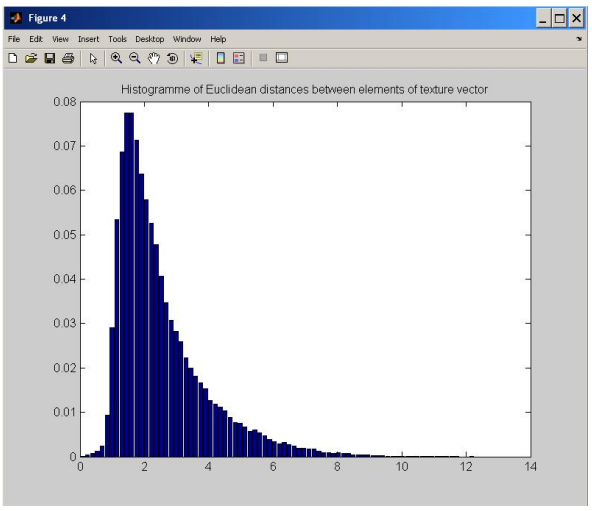

Figure 2 – *Histogram of Euclidean distances between elements of texture vector*

where k_i and l_j represent elements of dimension d from pairwise scalars of matrix X, $P(k_i)$ and $P(l_j)$ represent to the probability which is computed by the distribution function of element k_i and l_j respectively. μ_k and μ_l represent to the average of the elements in scalar k_i and l_i respectively.

The data we use in this paper to build Active Appearance Model are images from CMU PIE database [10]. We built histograms of the Euclidean distances between elements contained in each vector to study the dependencies of observed variables. As illustrated in Figure 1 and Figure 2, the observed variables are not uncorrelated, for both shape vectors and texture vectors, the Euclidian distances between each observed variable follows a Gaussian distribution.

According to the dependency analysis of observed variables of our data, $P(k_i)$ in Eqn.(1) is defined as Gaussian distribution function :

$$P(k_i) = \frac{1}{2\pi\sigma} e^{-\frac{|k_i - \mu_k|}{2\sigma k^2}} \tag{7}$$

Where μ_k represent to the average of the elements in scalar k_i,while σ_k^2 represent to the variance between the elements in scalar k_i.

3 Experimental Results

Accurate alignment of face images is a very important step in application such as facial expression analysis and face recognition. This section highlights that the proposed model provides higher accuracy benefits than classical Active Appearance Model for face alignment in both point-to-point error and pixel-to-pixel error sense.

We tested the proposed method on the CMU Pose, Illumination, and Expression (PIE) database of human faces [11]. For the experiments on the variation of illumination, the training database is built from a subset of the CMU database as shown in Figure 3. The test set is built with the images of the persons shown in Figure 4. We manually labelled 1200 images with the size of $640 * 486$ pixels. To establish the models 58 landmarks were placed on each face image : 8 points for the mouth, 11 points for the nose, 16 points for both eyes, 10 points for both eyebrows, and 13 points for the chin. The warped images have approximately 7325 pixels inside the facial mask.

Figure 3 – *Persons in the training database*

Figure 4 – *Persons in the test database*

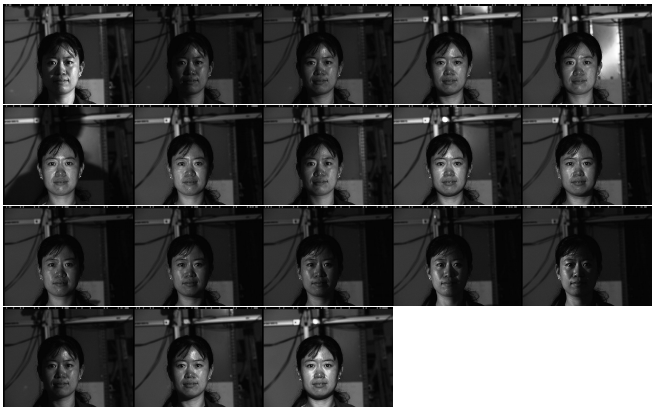

Figure 5 – *Examples of images of the same person under different illumination conditions (CMU PIE)*

The illumination related model is built using all the frontal faces which were captured by camera number 27 from the condition which is named "illumination" in the CMU PIE database.

As shown in figure 5, the training set containing 320 images from 16 persons, each person being represented by 20 frontal face images under 20 different illumination conditions. The test set is built in the same way, containing 200 images from 10 persons.

Illustrated by Figure 6, the first 4 rows are the test result on the known images (the images from the training set). The following three rows are test results on unknown images (faces not contained in the training set). The images in the left column are synthesized by the proposed method, compared with fitting result of classical AAM in the centre column and the images in the right column are the original images. One can notice that much better results have been obtained by the proposed method.

To evaluate the performance of the proposed method, the manually annotated landmarks are considered as the ground truth shape information. For each image the landmarks automatically labelled by the Active Appearance Models are compared with the ground truth landmarks. A distance measure, $D(x_{gt}, x)$, gives an interpretation of the fit between two shapes, the ground truth x_{gt} and the actual shape x. Point-to-point error E_{pt} is defined as the Euclidean distance between each corresponding landmark in (8).

$$E_{pt} = \frac{1}{n} \sum \sqrt{(x_i - x_{gt,i})^2 + (y_i - y_{gt,i})^2} \qquad (8)$$

where x_i and y_i) are the coordinates of the landmarks automatically labelled, $x_{gt,i}$ and $y_{gt,i}$ are the manually annotated landmarks which considered as the ground truth.

Interpreting a novel image is an optimization problem in which the method minimizes the error between the grey level of the pixels contained in a new image and the grey level of the pixels synthesized by the appearance model. Therefore, we can also evaluate our results with the pixel-to-pixel error E_{pix} which can be defined as (9).

$$E_{pix} = |\delta I|^2 = |I_i - I_m|^2 \qquad (9)$$

Where I_i is the vector of grey-level values of the pixels in the image, I_m is the vector of grey-level values of the pixels reconstructed with the current model parameters.

To evaluate the proposed method (PAAM), Eqn.(10) and (11) is used to compute the gain (defined as G) of the reduction of error, between PAAMs and standard AAMs. In order to highlight the degree of error reduction, G is computed into percentage.

$$G(pt) = \frac{E_{pt}(PAAM) - E_{pt}(AAM)}{E_{pt}(AAM)}\% \qquad (10)$$

$$G(pix) = \frac{E_{pix}(PAAM) - E_{pix}(AAM)}{E_{pix}(AAM)}\% \qquad (11)$$

The gain percentages in reducing the error (point-to-point error and pixel-to-pixel error) are presented in Table 1, computed by Eqn.(8)-(11).

Table 1 – *Gain percentage between AAM using PCA and Probabilistic PCA for the illumination problem (Eqn.(10) and (11))*

	$G(pt)$	$G(pix)$
Training database	32.4%	28.0%
Test database	27.9%	25.6%

Figure 7 and Figure 8 show error curves obtained from the "Standard AAM" (asterisk curve) and from the "Probabilistic AAM" (round curve). Each curve in figure 7 are the mean point-to-point error made by the fitting test on the 16 known face images (faces contained in the training set) under 20 different illuminations. Illuminations from numbers 2 to 5 are effected by the light source from the left side of the face, illuminations from 13 to 18 are effected by the light source from the right side of the face. While the rest illuminations (from number 6 to 12 and number 19 to 21) are less complicated and lighted by light source in front of

the face. As the curves give a consistent result, from illuminations number 6 to 12, both E_{pt} and E_{pix} are less made in the experiment.

This error curve depicts the robustness of using probabilistic AAM. We can see that with the proposed method, some errors are still made, but they are not as strong as with the standard method where errors are made upon when lighting source is on one side.

In figures 9 and 10 same experiment is conducted on the 10 unknown faces to create error curves. We can remark that for both method, errors are smaller with known faces than with unknown faces, which is a logical outcome of AAM.

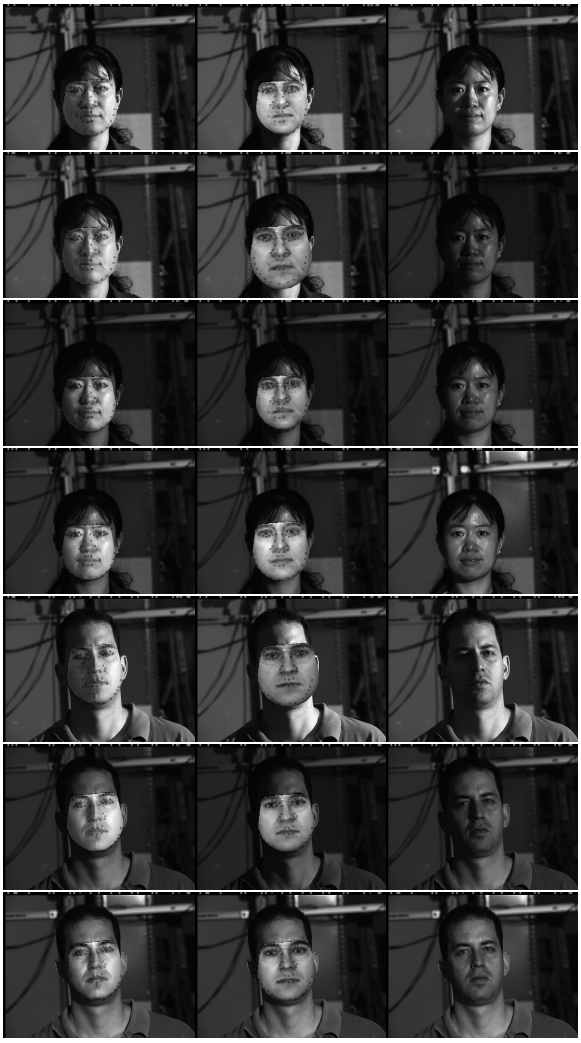

Figure 6 – *Visual comparison of fitting results between AAM using PCA and Probabilistic PCA for the illumination problem (Left column : PAAM results, Middle column : AAM results, Right column : Original images)*

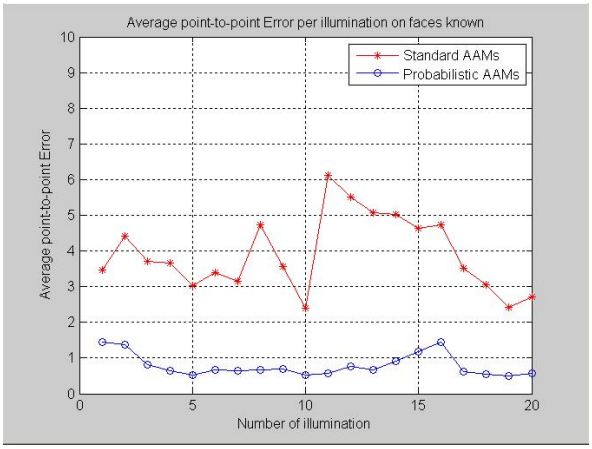

Figure 7 – *Average point-to-point error per illumination on faces known*

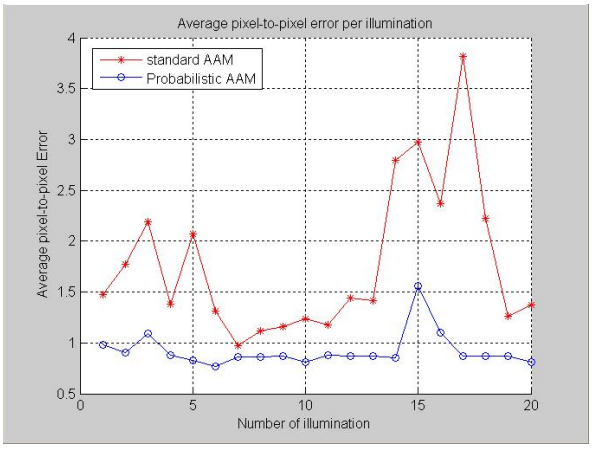

Figure 8 – *Average pixel-to-pixel error per illumination on faces known*

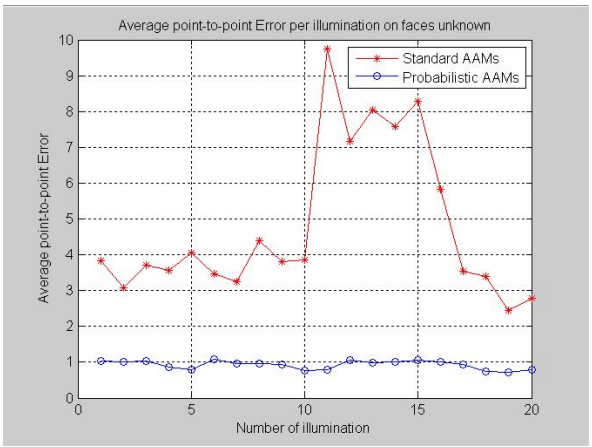

Figure 9 – *Average point-to-point error per illumination on faces unknown*

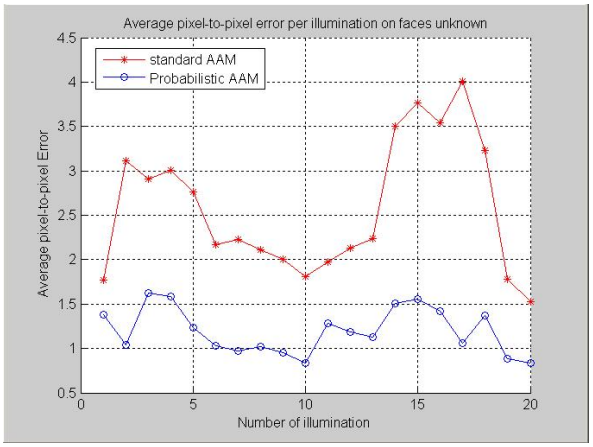

Figure 10 – *Average pixel-to-pixel error per illumination on faces unknown*

Conclusion

In this study, we have proposed a probabilistic method for the AAM fitting algorithm that is particularly robust to the illumination changes of face images. We have shown that the model built by the proposed method is much less sensitive to illumination variations. With this novel method, the fitting procedure can accurately synthesize the semi-bright-semi-dark faces affected by the illumination conditions. Next step of this work is to generalize the proposed method to complex pose variation and to the combined pose and illumination variation.

Références

[1] G.J Edwards T.F. Cootes et C.J. Taylor. Active appearance models. *Computer Vision Lecture Notes in Computer Science*, 1407/1998(484-498), 1998.

[2] Iain Matthews et Simon Baker. Active appearance models revisited. *INTERNATIONAL JOURNAL OF COMPUTER VISION*, 60(135–164), 2003.

[3] A. Smola B. Scholkopf et K.-R. Muller. Kernel principal component analysis. *Advances in Kernel Methods Support Vector Learning*, (327-352), 1999.

[4] Shaogang Gong Sami Romdhani et Alexandra Psarroura. A multi-view nonlinear active shape model using kernel pca. *Proceeding of the 10th British Machine Vision Conference*, (483-492), 1999.

[5] Baback Moghaddam et Alex Pentland. Probabilistic visual learning for object detection. *The 5th International Conference on Computer Vision, ICCV'95*, (687), June 1995.

[6] Laurent Thoraval Raouf Hamdan, Fabrice Heitz. A low complexity approximation of probabilistic appearance models. *Pattern Recognition*, Vol. 36, Issue 5(1107-1118), 2003.

[7] V. Jain G. Huang et E. Learned-Miller. Unsupervised joint alignment of complex images. *In IEEE 11th International Conference on Computer Vision, ICCV*, (1-8), 2007.

[8] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, (559-572), 1901.

[9] Michael E. Tipping et Chris M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society*, 61(611-622), 1999.

[10] Terence Sim, Simon Baker, et Maan Bsat. The cmu pose, illumination, and expression (pie) database of human faces. Rapport technique CMU-RI-TR-01-02, Robotics Institute, Pittsburgh, PA, January 2001.

[11] Baker et M. Bsat. The cmu pose, illumination, and expression (pie) database. *Proceedings. Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, (46-51), 2002.