Face Gender Classification Based on Active Appearance Model and Fuzzy k-Nearest Neighbors

Len Bui, Dat Tran, Xu Huang and Girija Chetty
Faculty of Information Sciences and Engineering
University of Canberra, ACT 2601, Australia

Abstract - *We present a novel method for face gender classification problem. This method employs the powerful Active Appearance Model (AAM) method for modeling human faces and extracting feature vectors, and the simple but powerful Fuzzy k-Nearest Neighbors (Fuzzy k-NN) method for classification. Experiments for the proposed approach have been conducted on FERET data set and the results show that the proposed method could improve the classification rates.*

Keywords: Face gender classification, Active Appearance Model, Fuzzy k-Nearest Neighbors

1. Introduction

Human faces contain a lot of important biometrics information. One of them is about sexual category. Therefore, advances in face gender classification would be meaningful to the development of biometrics authentication systems. There are two main approaches to solving the face gender classification problem.

Most of the early research on gender classification focused on geometry based methods. These methods were based on metric features such as face width, face length, mouth size, and eye size for classification. Burton *et al.* [2] used 73 points extracted from a database containing 179 frontal facial images. The reported accuracy rate was 85%. Fellous *et al.* [5] used 22 normalized distances computed from a data set with 109 images. The accuracy rate was reached up 90%. However, facial feature detection and measurement techniques are not sufficiently reliable, and such geometric properties alone are inadequate for gender classification because rich information contained in the facial texture or appearance is discarded. For those reasons, the early research techniques proposed were not effective.

Recently, most of researchers have been interested in the appearance-based approaches. Gollomb *et al.* [6] developed a neural network called SEXNET to identify gender. The experiment was conducted on the dataset of 90 images of young adult faces (45 males and 45 females). The reported accuracy rate was 91.9%. Cottrell and Metcalfe [4] also used neural networks for face emotion and gender classification. The report said that it achieved the perfect results. Brunelli and Poggio [1] applied HyperBF networks for gender classification. The experiments conducted on a data set of 168 images (21 males and 21 females) achieved an accuracy rate of 79%. Gutta *et al.* [8] used a hybrid classifier based on neural networks and inductive decision trees with experiments conducted on FERET dataset. The best average accuracy rate was 96%. Moghaddam and Ming-Hsuan [10] developed gender classifiers using SVMs on FERET database. They reported that a Gaussian kernel SVM was able to achieve a 3.4% error rate. Currently, there are a lot of studies [7, 17-19] using this method.

Although the recent experimental results are good, most of studies still face a problem that is how to get good feature vectors. Since Matthew Turk and Alex Pentland [15] used Principal Component Analysis to deal the face recognition problem, it has become the major mathematical tool to extract feature vectors. To deal with the problem of face variability, some researchers follow model-based methods. Using face models, face image interpretation can be formulated as a matching problem: given a face image to interpret, the face can be located and labelled by adjusting the model parameters such that it generates an image which the best approximation of the real thing. One of the outstanding face models is Active Appearance Model (AAM) invented by Cootes [3]. Fitting an AAM to a face image consists of minimizing the error between the input image and the closest model instance; i.e., solving a nonlinear optimization problem. However, most of the fitting algorithms use the error functions or objective functions without considering the important information on faces. For instance, eyes contain more information than heads. Therefore, we propose a novel method for model fitting. For classification task, we combine k-Nearest Neighbor, the simple but powerful classification method, and Fuzzy Logic [14], the method to deal with uncertain things in classification. The remaining sections of our paper will discuss the implementation of our face gender classification system, related theory, and experiments. Section 2 gives details of AAM for human faces. Section 3 discusses how to use Fuzzy k-Nearest Neighbors for face gender classification. In Section 4, we will describe the implementation and experiments. Finally, we will conclude in Section 5.

2. Active Appearance Models for Human Faces

2.1 Model Construction

Our face model is built using a mathematical method which is called PCA as mentioned above. It was first applied in face classification by Sirovich and Kirby and then Matthew Turk and Alex Pentland [15]. Now it has become the standard method in this field.

Training data $\{x^{(i)}, i = 1 \ldots N\}$

Algorithm 1: Principal Component Analysis

Step 1: Compute the mean of data

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x^{(i)} \quad (1)$$

Step 2: Compute the covariance matrix of data

$$A = \frac{1}{N-1}\sum_{i=1}^{N}\left(x^{(i)} - \bar{x}\right)\left(x^{(i)} - \bar{x}\right)^T \quad (2)$$

Step 3: Compute the eigenvectors ϕ_i an eigenvalues λ_i of matrix A

Finally, Statistical model built from data set is $x = \bar{x} + M.p$, where $M = (\varphi_1\, \varphi_2 \ldots \phi_k)$ and model paramters $p = (p_1, p_2, \ldots, p_k)$.

2.2 Face Model

AAM is a parametric model which was first developed by Cootes to model deformable objects. In addition, it has been successfully applied in modeling human faces. AAM combine statistical shape and texture models. Currently, there are a few AAMs. In the paper, we only discuss independent AAMs. As the name states, independent AAMs have shape and appearance separately. The shape of an independent AAM is defined by a set of points $S = (x_1, y_1, x_2, y_2, \ldots, x_v, y_v)^T$. The texture of the model is defined by a bitmap image $A(x)$. A sample including shape and texture is shown in Fig. 1.

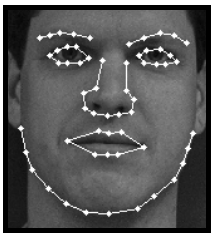

Fig. 1. A sample including shape and texture

Suppose that we have a training dataset including shape data and appearance data $\{(S^{(i)}, A^{(i)}),\ i = 1, \ldots, N\}$. Using PCA that we described in Algorithm 1, we have the shape model *S* and texture model *A* as follows

$$S = S_0 + \sum_{i=1}^{i=N_S} p_i S_i, \qquad A = A_0 + \sum_{i=1}^{i=N_A} \lambda_i A_i \quad (3)$$

2.3 Model Fitting

Fitting an AAM to an image is a non-linear optimization problem to minimize a squared error measure between desired output and the model's outputs. Given model $y = T(x, \theta)$ with θ are model parameters and training data set $\{(x^{(i)}, t^{(i)}), i = 1, \ldots, N\}$. The sum of squared error E is the objective function respect to θ. The problem is to find the optimal θ^* that minimizes E.

$$E(\theta) = \sum_{i=1}^{N}\left(t^{(i)} - y^{(i)}\right)^2 = \sum_{i=1}^{N}\left(t^{(i)} - T\left(x^{(i)}, \theta\right)\right)^2 \quad (4)$$

or

$$E(\theta) = r^T(\theta) r(\theta)$$

$$r(\theta) = \left(t^{(1)} - y^{(1)}, \ldots, t^{(N)} - y^{(N)}\right) \quad (5)$$

Algorithm 2: Gauss-Newton algorithm

Initialize θ

Repeat until convergence

Step 1a: Determine gradient $g = J^T r$

Step 1b: Determine direction vector

$$d = -\frac{1}{2}H^{-1} = -\frac{1}{2}(J^T J)^{-1}$$

Step 2: Set step size $\mu = 1$

Step 3: Update parameters

$$\theta_{next} = \theta_{current} - \frac{1}{2}(J^T J)^{-1} J^T r$$

where J and H are the Jacobian and Hessian matrices of the function r, respectively. Descent direction *d* in gradient-based method is shown in Fig. 2.

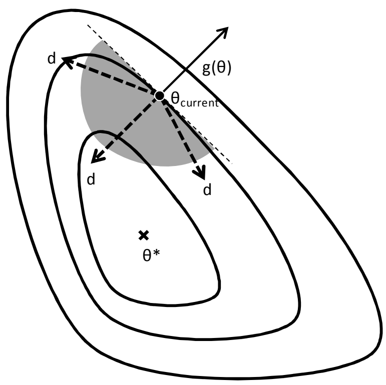

Fig. 2. Descent direction *d* in gradient-based methods

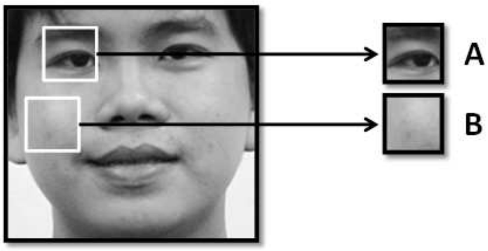

Fig. 3. Two regions A: eye and B: cheek

Suppose that we have an observed image $I(x)$, we need to find parameters p and λ to fit the model to the image. It means that we have to minimize an error function. Most of studies [3, 9] use the error function $E = \sum_x (A_0(x) \quad I(T(x,p)))^2$. As we mentioned before, this formula considers all pixels of A_0 or I have the same weight toward the overall squared error. In fact, each region of face image has different important degree; for example, region A is more significant than region B in Fig. 3. Therefore, we propose a different error function $E = W(x)\sum_x (A_0(x) - I(T(x,p)))^2$ where W is weighting vector. It means that each pixel $I(x)$ has each weight $W(x)$. The value of $W(x)$ is assigned to the variance of each pixel.

Algorithm 3: Fitting algorithm

Initialize the parameters p of model T

Repeat until convergence

Step 1: Warp I with $T \quad I(T(x,p))$ (6)

Step 2: Compute the squared error between the images

$$W(x)\left(A_0(x) - I(T(x,p))\right)^2 \tag{7}$$

Step 3: Warp the gradient ∇I with T

$$\nabla I(T(x,p)) \tag{8}$$

Step 4: Compute the Jacobian

$$\frac{\partial T}{\partial p} \tag{9}$$

Step 5: Compute the steepest descent textures

$$\nabla I \frac{\partial T}{\partial p} \tag{10}$$

Step 6: Compute the Hessian matrix H

$$H = \sum_x \left[\nabla I \frac{\partial T}{\partial p}\right]^T W \left[\nabla I \frac{\partial T}{\partial p}\right] \tag{11}$$

Step 7: Compute the steepest descent parameter updates

$$\sum_x \left[\nabla I \frac{\partial T}{\partial p}\right]^T W(x)\left[A_0(x) - I(T(x;p))\right] \tag{12}$$

Step 8: Compute Δp

$$\Delta p = H^{-1} \sum_x \left[\nabla I \frac{\partial T}{\partial p}\right]^T W(x)\left[A_0(x) - I(T(x,p))\right] \tag{13}$$

Step 9: Update the parameters p

$$p = p + \Delta p \tag{14}$$

3. Fuzzy k-Nearest Neighbors

Fuzzy *k*-Nearest Neighbor methods are simple and effective for classification. Suppose that we have a labeled data set $D = \{(x^{(i)}, \theta^{(i)}),\ i = 1, \dots, N\}$ where $\theta^{(i)} \in \{w_1, w_2, \dots, w_C\}$, the class label for $x^{(i)}$. Given a new point x, we have to determine the label for x.

Algorithm 4: Conventional *k*-NN algorithm

Step 1: Finding k nearest neighbors of x

Step 2: The decision rule to choose a label for x

$$\theta(x) \leftarrow \{\text{majority class label of k nearest neighbors}\} \tag{15}$$

Fig. 4 shows an example of *k*-nearest neighbors of *x* and distance metric *d*.

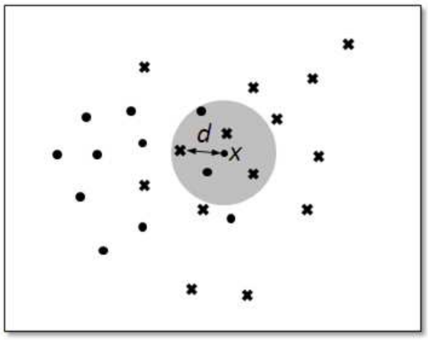

Fig. 4. An example of *k* nearest neighbors of *x* and distance metric *d*

Obviously, the decision rule of the above algorithm is quite simple and rough. Therefore, we decide to use the distance metric to smooth the decision rule. We define the membership function $MF_{x_i}(x)$ for each x_i. The values of these functions are computed from the distance from x to x_i.

Algorithm 5: Fuzzy k-NN algorithm

Step 1: Finding k nearest neighbors of x, $N_k(x)$

Step 2: Compute the value of membership function of x for each class label w_i

$$MF_{w_i}(x) = \frac{1}{k}\sum_{x_i \in N_k(x), \theta(x_i) = w_i} MF_{x_i}(x) \tag{16}$$

Step 3: The decision rule to choose a label for x

$$\theta(x) = \underset{w_i}{argmax}\left(MF_{w_i}(x)\right) \tag{17}$$

4. Implementation and Experiments

First of all, we choose the public database IMM [11-12] to build AAM model. The IMM face database comprises 240 still images of 40 different human faces at 6 different poses, all without glasses. The gender distribution is 7 females and 33 males. The facial structures were manually annotated using 58 landmarks: eyebrows, eyes, nose, mouth and jaw. A total of seven point paths were used; three closed and four open. To improve the fitting algorithm, we design pyramid-AAM models.

To evaluate our proposed approach, we choose FERET face database to test our method. The FERET database was collected at George Mason University between August 1993 and July 1996. It contains 1564 sets of images for 14,126 images that include 1199 individuals and 365 duplicate sets of images. We design face gender classification system including two main modules of extraction and classification. In the first module, face regions are identified and extracted from the background of the input image by using the well-known algorithm developed by Viola and Jones [13, 16]. The information of this step is the prior knowledge to initialize parameters p in Algorithm 2 because it provides rough estimates of the location and scale of each detected face. Next, the system will run the algorithm 2 to fit AAM model to face image. Finally, the parameters of the best fitting model instance are used for classification by using Algorithm 5 that I described before.

From FERET database, we constructed two data sets for training and testing. The size of each data set is 200 including 100 males and 100 females that are randomly chosen from FERET database.

After training the classifier by registering all people in training data set, we conducted three experiments for gender classification.

Experiment 1: We only use the distance metrics $L_1(x,y) = \sum_{i=1}^{n} |x_i - y_i|$ and $L_2(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$; the value of k is set to 1, 3, 5, 7 and 9. In addition, we use the conventional AAM fitting algorithm [3, 9]. According to the experiment, the good results can be reached when the value of k is 5. There is no significant change in error rates we increase the value of k.

Table 1. Experiment 1 results

Distance metric	*Error rate (in %)*		
	Male	Female	Overall
L_1	9.2	11.3	10.3
L_2	7.4	10.1	8.8

Experiment 2: We use the proposed AAM fitting algorithm in this experiment. According to the results, there is a significant improvement in classification rates.

Experiment 3: We continue to use the proposed AAM fitting algorithm and two membership functions, Triangle and Gaussian functions (which are shown in Fig. 5) for the distance metric L_2 with the value of k to be 5. According to the experiment, it gives the better results than Experiment 2.

Table 2. Experiment 2 results

Distance metric	*Error rate (in %)*		
	Male	Female	Overall
L_1	7.7	9.8	8.8
L_2	5.3	7.5	6.4

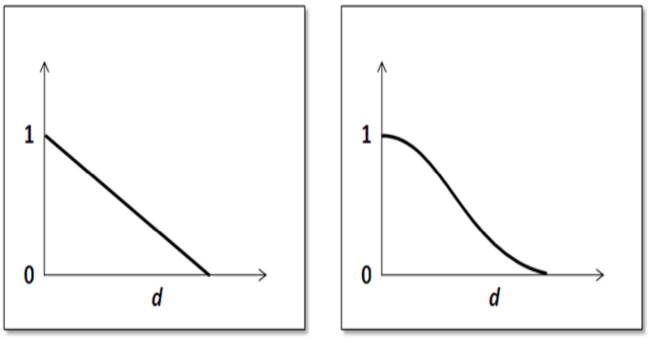

Fig. 5. Triangle shape function and Gaussian shape function of distance *d*

Table 3. Experiment 3 results

Membership function	*Error rate (in %)*		
	Male	Female	Overall
Triangle shape	6.3	8.7	7.5
Gaussian shape	4.2	5.1	4.7

5. Conclusions

In summary, we have proposed a new approach for face gender classification. The first contribution of this paper is to propose a novel method to fit an AAM model, a generative powerful model that is widely applied in many different fields, to an input image. The second contribution of this paper is to suggest using Fuzzy decision rule for classification. We compared our method against conventional fitting method. The results from our method outperformed significantly.

A future direction of research for face gender classification is to combine Support Vector Machine, the most dominant method for pattern classification, with Fuzzy Logic and conduct experiments on large databases.

6. References

1 Brunelli, R., and Poggio, T.: 'Hyperbf networks for gender classification', in Editor (Ed.)^(Eds.): 'Book Hyperbf networks for gender classification' (1992, edn.), pp. 311–314

2 Burton, A.M., Bruce, V., and Dench, N.: ‘What's the difference between men and women? Evidence from facial measurement’, Perception, 1993, 22, (2), pp. 153-176

3 Cootes, T., Edwards, G., and Taylor, C.: ‘Active appearance models’, Computer Vision—ECCV’98, 1998, pp. 484

4 Cottrell, G., and Metcalfe, J.: ‘EMPATH: Face, emotion, and gender recognition using holons’, in Editor (Ed.)^(Eds.): ‘Book EMPATH: Face, emotion, and gender recognition using holons’ (Morgan Kaufmann Publishers Inc., 1990, edn.), pp. 571

5 Fellous, J.: ‘Gender discrimination and prediction on the basis of facial metric information’, Vision research, 1997, 37, (14), pp. 1961-1973

6 Golomb, B., Lawrence, D., and Sejnowski, T.: ‘Sexnet: A neural network identifies sex from human faces’, Advances in neural information processing systems, 1991, 3, pp. 572-577

7 Graf, A., and Wichmann, F.: ‘Gender Classification of Human Faces’ (2010), pp. 1-18

8 Gutta, S.: ‘Gender and Ethnic Classification of Face Images’, in Editor (Ed.)^(Eds.): ‘Book Gender and Ethnic Classification of Face Images’ (1998, edn.), pp. 194-194

9 Matthews, I., and Baker, S.: ‘Active appearance models revisited’, International Journal of Computer Vision, 2004, 60, (2), pp. 135-164

10 Moghaddam, B., and Ming-Hsuan, Y.: ‘Gender classification with support vector machines’, in Editor (Ed.)^(Eds.): ‘Book Gender classification with support vector machines’ (2000, edn.), pp. 306-311

11 Nordstrøm, M., Larsen, M., Sierakowski, J., and Stegmann, M.: ‘The IMM face database’, environment, 2003, 22, (10), pp. 1319-1331

12 Stegmann, M.B., Ersboll, B.K., and Larsen, R.: ‘FAME-a flexible appearance modeling environment’, Medical Imaging, IEEE Transactions on, 2003, 22, (10), pp. 1319-1331

13 Thai Hoang, L., and Len Tien, B.: ‘A hybrid approach of AdaBoost and Artificial Neural Network for detecting human faces’, in Editor (Ed.)^(Eds.): ‘Book A hybrid approach of AdaBoost and Artificial Neural Network for detecting human faces’ (2008, edn.), pp. 79-85

14 Tran, D., Ma, W., Sharma, D., Bui, L., and Le, T.: ‘A Generic Framework for Soft Subspace Pattern Recognition ’, in Er, M.J., and Zhou, Y. (Eds.): ‘Theory and Novel Applications of Machine Learning’ (IN-TECH 2009)

15 Turk, M., and Pentland, A.: ‘Eigenfaces for recognition’, Journal of cognitive neuroscience, 1991, 3, (1), pp. 71-86

16 Viola, P., and Jones, M.J.: ‘Robust real-time face detection’, International Journal of Computer Vision, 2004, 57, (2), pp. 137-154

17 Xia, H., Ugail, H., and Palmer, I.: ‘Gender Classification Based on 3D Face Geometry Features Using SVM’, in Editor (Ed.)^(Eds.): ‘Book Gender Classification Based on 3D Face Geometry Features Using SVM’ (2009, edn.), pp. 114-118

18 Yiding, W.: ‘Gender Classification Based on Enhanced PCA-SIFT Facial Features’, in Editor (Ed.)^(Eds.): ‘Book Gender Classification Based on Enhanced PCA-SIFT Facial Features’ (2009, edn.), pp. 1262-1265

19 Zhen-Hua, W., and Zhi-Chun, M.: ‘Gender classification using selected independent-features based on Genetic Algorithm’, in Editor (Ed.)^(Eds.): ‘Book Gender classification using selected independent-features based on Genetic Algorithm’ (2009, edn.), pp. 394-398.