Gender Classification of Face with Moment Descriptors

Wen-Shiung Chen, Po-Yi Lee
Dept. of Electrical Engineering
National Chi Nan University
Puli, Nantou, Taiwan
e-mail: {wschen,s97323529}@ncnu.edu.tw

Lili Hsieh
Dept. of Information Management
Hsiuping University of Science and Technology
Taichung, Taiwan
e-mail: lily@mail.hust.edu.tw

***Abstract*—Moments may effectively describe geometrical properties of an object and thus are widely used in image analysis and pattern recognition, except gender classification. In this work, we propose a modification of eigenmoment and apply to classify sex from facial features. We investigate and compare four moment descriptors with nearest neighbor and SVM in three different feature regions on face. The results show that the features used in gender classification can be effectively represented by moment descriptors.**

Keywords-Biometrics; Moment; Face Recognition; Gender Classification.

I. INTRODUCTION

Gender classification [1][2], detecting someone's sex from a human face, is used for personal authentication and multimedia interaction. Gender classification is really a difficult problem since there is no clear definition or rule on how to distinguish between males and female faces. Many different kinds of feature extraction methods have been applied in an attempt to solve the gender classification problem [3]. Literature shows that shape reveals a better effectiveness on classifying gender than texture and color [3]. Although moments were applied in computer vision for a long time [4]-[6], they have not been utilized so far in gender classification. In this paper, we address four different moment descriptors for extracting the features from facial images.

Most works on face recognition and gender classification use the face regions without hair. However, Lapedriza, *et al.* referred to the features of face regions by considering the external region (i.e., with hair) and the internal region (i.e., without hair) [7]. Makinen and Raisamo also used two different face regions and discussed their influences on performance [8]. We thought hair is worthy to be considered since sometimes it is difficult to determine someone's gender either when a man has long hair or a woman has short hair. Three different face regions will be considered in this work, namely, the external feature region, the general feature region and the internal feature region. The external feature region is a facial image with as much hair as possible. Contrarily, the other regions contain a face or faces without hair. The general feature region contains the forehead, the cheek and the chin. The internal feature region is a smaller region between the eyebrow and the lip. This work discusses the effectiveness of different regions on gender classification based on moments.

II. THE PROPOSED METHOD

A. System Overview

The proposed gender classification system consists mainly of three modules: pre-processing, feature extraction, and classification, as shown in Figure 1. First, the pre-processing module employs face detector to detect and localize the positions where faces are (i.e., face zone) from the input image. A cropped region, containing a face or faces, is the output. It performs four major tasks including feature region selection, face detection, image resize and histogram enhancement for the input face images. Next, the feature extraction module adopts moment descriptors to generate the feature vectors. Finally, the classification module employs nearest neighbor and the support vector machines to recognize the face images by comparing the feature vectors with what enrolled in databases.

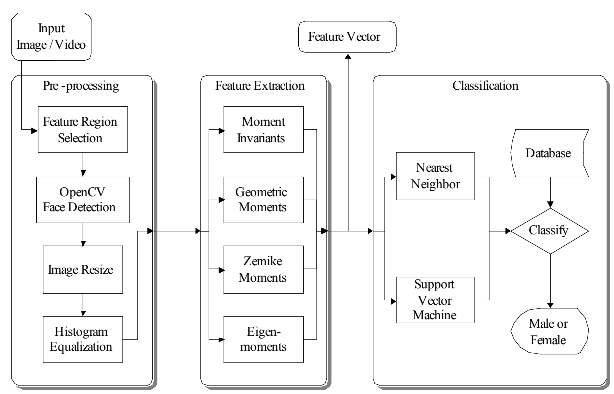

Figure 1. Diagram of the proposed face gender classification system.

B. Pre-processing

Three different face regions, external feature region, general feature region and internal feature region, shown in Figure 2, are considered. First, we extract a face image scaled to the size of 24×24 pixels, and called general feature region. The minimum face size in OpenCV detector [9] is set to 24×24 pixels and the sub-image is scaled after scanning by multiplying the current sub-image size by 1.25 times (i.e., 1st scan sub-image size: 24×24 pixels, 2nd: 30×30 pixels, and so on).

We also extract the larger detected face area from the same images, similarly. We enlarge the width on both sides by 10% (20% in total), the top by 40% and the bottom by 12%, of the image since this region usually contains hair in addition to face, but the remaining area of the image as little as possible, for example, the background. This region is called external feature region. In some cases the region could not be grown as much as intended since image borders were encountered, thus this kind of images were removed from the data. Other cases when images were considered of bad quality and were removed from the data included images in which objects such as a hat, a hand, or some other object, existed in front of the hair or the face. We extract the regions by decreasing the detected face area by 10% on both sides (20% in total), 20% on the top but unchanged to the bottom of the image since the bottom of default detected face area reach lip for many images. This feature region is called internal feature region. It is smaller than general feature region, so no image is removed.

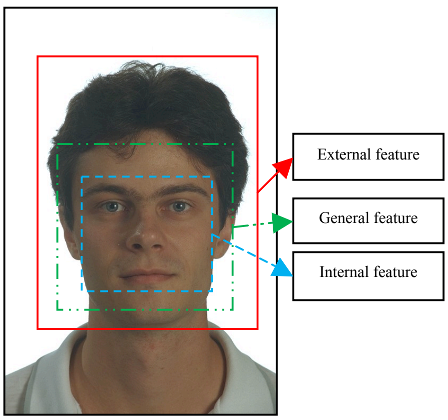

Figure 2. Three different feature regions.

We used 32×40 image size (instead of 24×24) when external feature region is considered since it is closer to the image size. Similarly, the image with internal feature region is resized to 20×20 and transformed to gray-level, and then followed by histogram equalization. Moreover, Makinen and Raisamo achieved the best results without automatic alignment [8]. Manual alignment is a possible way to obtain better results but difficult to implement for a large database. Hence, we do not normalize any image in experiments. Finally, the histogram equalization is performed once before feature extraction.

C. Feature Extraction

The geometric moment (GM) has the form of the projection of $f(x, y)$ $\boldsymbol{f(x,y)}$ onto the monomials $x^p y^q$ $\boldsymbol{x^p y^q}$, where $\boldsymbol{f(x,y)}$ is intensity of an image. Unfortunately, the basis set $\{x^p y^q\}$ is non-orthogonal, so that these moments are not orthogonal. Yap and Paramesran proposed an improvement of GMs by solving an eigenvalue problem in the moment space, called eigenmoment (EM) [10]. So far, eigenmoment is not applied on pattern recognition, especially gender classification, thus we propose to apply it in gender classification in this paper. Assume that our signal dataset is composed of m n-D column vectors s_i $\boldsymbol{S_i}$, where i = 1, 2, …, m. Let the $n\times m$ $\boldsymbol{n\times m}$ matrix $\mathbf{S}$ represent the signal dataset, a transformation matrix $\mathbf{X}$, constructed by GMs, converts the signal into the moment space. The dataset in moment space is represented as $\mathbf{S}^T\mathbf{X}$.

The original dataset, extracted by GMs, is information redundant since the kernel of GM is not orthogonal. Now the concept of principal component analysis (PCA) [11] helps to solve this problem. The goal is now to find a unity vector $\boldsymbol{w}$ $\boldsymbol{w}$ such that the squared sum of the dataset's projection onto this direction is maximal. The squared sum of the total projection of $\mathbf{S}^T\mathbf{X}$ onto $\boldsymbol{w}$ is a function of $\boldsymbol{w}$, denoted by:

$$\boldsymbol{L}(\boldsymbol{w}) = (\mathbf{S}^T\mathbf{X}\boldsymbol{w})^T(\mathbf{S}^T\mathbf{X}\boldsymbol{w}) = \boldsymbol{w}^T\mathbf{X}^T\mathbf{S}\mathbf{S}^T\mathbf{X}\boldsymbol{w}. \tag{1}$$

To maximize $\boldsymbol{L}(\boldsymbol{w})$ under the constraint $||\boldsymbol{w}||=1$, solving

$$\max_w L(w) = \max_w \frac{w^T A w}{w^T w} \tag{2}$$

where $\boldsymbol{A = X^T SS^T X}$ $\boldsymbol{A}=\mathbf{X}^T\boldsymbol{SS}^T\mathbf{X}$ is a real symmetric matrix (or covariance matrix). Equation (2) is called Rayleigh quotient. We use the Lagrange multiplier to form a new objective function:

$$\tilde{L}(w) = \boldsymbol{w}^T\boldsymbol{A}\boldsymbol{w}+\lambda(1-\boldsymbol{w}^T\boldsymbol{w}) \tag{3}$$

where λ is a Lagrange multiplier. The stationary points of $\boldsymbol{\tilde{L}(w)}$ $\tilde{L}(w)$ occur at $\nabla_w \tilde{L}(w) = 0$. Then, we obtain

$$L(w) = \frac{w^T A w}{w^T w} = \lambda \frac{w^T w}{w^T w} = \lambda\,. \tag{4}$$

Therefore, the eigenvectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, …, $\boldsymbol{w}_m$ of $\mathbf{A}$ are the critical points of the Raleigh quotient and their corresponding eigenvalues λ_1, λ_2,…λ_m $\boldsymbol{\lambda_1, \lambda_2, \cdots,}$ are the stationary values of $\boldsymbol{L}(\boldsymbol{w})$. Based on principal components analysis, we can arrange the eigenvalues of $\mathbf{A}$ into a descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. Then, the maximal value of $\boldsymbol{L}(\boldsymbol{w})$ is λ_1 occurring at $\boldsymbol{w} = \boldsymbol{w}_1$, while its minimum is λ_m occurring at $\boldsymbol{w} = \boldsymbol{w}_m$. Since $\mathbf{A}$ is a symmetric matrix, its eigenvalues are positive and the eigenvectors are orthogonal to one another.

The original signal projected onto moment space via GMs can be represented as $\mathbf{S}^{\mathrm{T}}\mathbf{X}$. However, the kernel of GMs is not orthogonal, it makes information redundant. Then we use PCA to find a vector $\boldsymbol{w}$ and the projection $\mathbf{S}^{\mathrm{T}}\mathbf{X}$ onto $\boldsymbol{w}$ can be represented as (1). Notice $\boldsymbol{w}$ is an orthonormal basis obtained by solving (2). Transformation with orthogonal moment kernels into the moment space is equivalent to the projection of the signal onto orthogonal basis. Thus, we improve the information redundancy problem and still keep the low computational complexity of GMs. Furthermore, the eigenvalues of $\mathbf{A}$ close to zero means low representation of dataset in the moment space. We can discard these features

 ISBN: 978-1-61208-283-7

to compact information. In the experiments, we discuss its influence as discarding 53, 27 and zero features. The concept of the proposed method is illustrated in Figure 3.

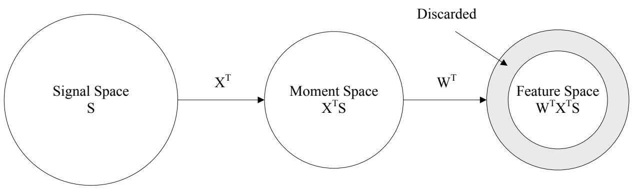

Figure 3. Conceptual chart of eigenmoments [10].

The features of the tested images are extracted by moment descriptors. These features form a high dimension feature vector as input to a classifier. TABLE I summarizes the number of features in our experiments.

TABLE I. # OF FEATURES IN THE EXPERIMENTS

	GMs	Hu moment invariants	Zernike moment	Eigen moment
Order	10	--	10	10
# of features	63	7	36	63

D. Classification

In this paper, we adopt *k*-nearest neighbor (*k*-NN) [11] and support vector machine (SVM) [12] as the classifier. We used LIBSVM [11] to train SVM with RBF kernel and predict classification rates for four moment descriptors. The parameters *Γ* (gamma) and C (cost) will be automatically determined by LIBSVM in experiments.

III. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed gender classification system, we implemented and tested the proposed schemes on two different databases. The first is a simple database collected by ourselves and containing 100 (50 males and 50 females) face images which are frontal with different expressions. The second is the well-known FERET database. In our experiments, only frontal images (FA and FB classes) are used. These classes contain 2,722 images. Duplicated images of the same person are removed, so only one image per person is kept. Next, we pick out 900 (450 males and 450 females) images from the rest. In order to compare fairly, this procedure is the same as [13]. Figure 4 depicts the flowchart of the procedure.

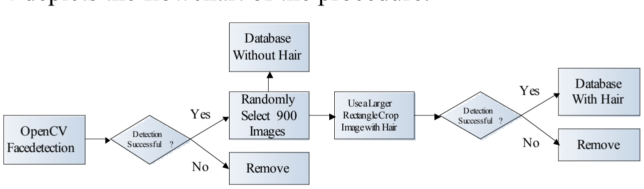

Figure 4. Flowchart of selecting tested images.

We take 900 images for general feature. As described in pre-processing, we increase the detected area for external feature region. In some cases the area could not be grown as much as intended since image borders came across. We removed the images from the data. Zernike moments need to extract a square region transformed into a unit circle disk. Thus, a larger detected face area is used to extract a square region. Those images with crossing borders are removed. TABLE II summarizes the number of experimented images.

TABLE II. IMAGES OF DATABASE IN EXPERIMENTS

Region	Internal feature	General feature	External feature	External feature for ZM
size	20 × 20	24 × 24	32 × 40	40 × 40
Simple database	100	100	100	100
FERET database	900	900	862	600

We implement four moment descriptors and two classifiers for each feature region. The training and testing strategy is five-fold cross validation. The test sets are not overlapped with their respective training sets. TABLEs III and IV summarize the experimental results on simple database and FERET database, respectively.

TABLE III. THE RESULTS ON SIMPLE DATABASE

Simple database	Classification rate (%)					
	External feature 32 × 40		*General feature 24 × 24*		*Internal feature 20 × 20*	
Methods	NN	SVM	NN	SVM	NN	SVM
MI (7)	92	96	82	89	66	75
ZM (36)	90	93	91	95	77	85
GM (63)	93	98	91	97	77	86
EM (10)	93	100	89	98	77	87
EM (36)	82	97	82	95	68	75
EM (63)	67	91	62	90	58	67

TABLE IV. THE RESULTS ON FERET DATABASE

FERET database	Classification rate (%)					
	External feature 32 × 40		*General feature 24 × 24*		*Internal feature 20 × 20*	
Method	NN	SVM	NN	SVM	NN	SVM
MI (7)	71.2	77.8	69.8	78.9	61.8	67.8
ZM (36)	82.4	82.5	81.1	83.4	75.3	81.0
GM (63)	76.0	82.4	76.2	80.9	64.7	78.6
EM (10)	75.0	82.8	78.7	79.9	69.1	74.0
EM (36)	78.2	83.3	80.6	84.8	71.6	78.8
EM (63)	76.9	82.7	81.3	84.8	71.2	80.0

The results show that the performances of eigenmoment are not as well as expected when the features are not discarded. Simple database is a small dataset of only 100 images. The eigenvalues close to zero mean a bad representation of the dataset in the moment space. Accordingly, we can discard these features to compact information. The eigenvalues explicate the proportion of distribution of dataset obtained by using PCA theory. Figure 5 shows the proportion of accumulative eigenvalues in which the accumulation of first 10 eigenvalues exceeds 99.5%, so the performance is improved significantly by discarding the other features.

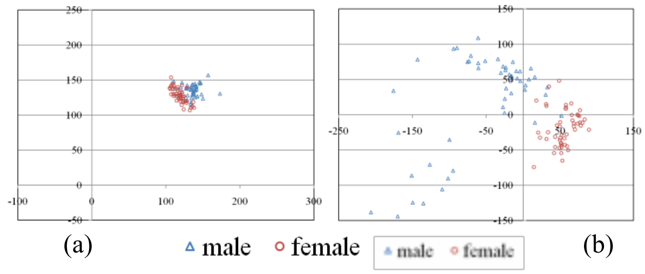

Figure 5. The distribution of first two features. (a) Geometrical moments. (b) Eigenmoments.

In Figure 5(a), we set M_{20} to *x*-axis and M_{02} to *y*-axis. In Figure 5(b), *x*-axis and *y*-axis represent the first and second data distribution eigenvalues of eigenmoment, respectively. It is easy to explain that the performances of eigenmoment overcome those of GMs. The performances get worse in the experiments on FERET database since the images are influenced by all kinds of variations. It is worth to notice that the performances of eigenmoment are not as well as expected again. This results from the variation of number of images. Simple database only contains 100 images while FERET database contains 900 images in our experiment. It is not enough to describe 900 images if we still use 10 features, so we increase to 36 and 63 features. The results become better when we increase the number of features. However, the improvement is not obvious between 36 and 63 features.

One the other hand, we test three feature regions in the experiments. According to the results, the performances of external feature have the best classification rate in simple database. However, to FERET database, the influence of background makes the performances of external feature worse. Overall, the general feature is better for moment descriptors on gender classification problem. The results in [13] are summarized in TABLE V. Comparing our results with those in [13], moment descriptors still have good enough classification rate which is above average level in FERET database.

TABLE V. EXPERIMENTAL RESULTS ON FERET DATABASE WITHOUT NORMALIZTION IN [12]

Methods	Classification rate (%)		
	Without hair (24 × 24)	*With hair (32 × 40)*	*Average classification rate*
Neural network	83.89	90.07	86.98
SVM	84.44	72.85	78.65
Threshold Adaboost	82.22	83.44	82.83
LUT Adaboost	80.56	87.42	83.99
Mean Adaboost	76.67	87.42	82.05
LBP + SVM	75.56	72.19	73.88
Average rate	80.56	82.23	81.40

IV. CONCLUSION

In this paper, we present a gender classification system based on moment descriptors. The results show that the features used in gender classification can be effectively represented by moment descriptors. The classification rate up to 100 % can be achieved when we test 10 features using eigenmoment with SVM on simple database.

REFERENCES

[1] B. Moghaddam and M. Yang, "Learning gender with support faces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 707-711, May 2002.

[2] W.-S. Chen, W.-J. Chang, L. Hsieh, Z.-Y. Lin, "Component-based gender classification based on hair and facial geometry features, IJCCI, 2012, pp. 626-630.

[3] W. Zhao, R. Chellappa, A. Rosenfeld and P. J. Phillips, "Face recognition: A literature survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399-458, December, 2003.

[4] M. K. Hu, "Visual pattern recognition by moment invariants," IEEE Transactions on Information Theory, vol. 8, no. 2, pp. 179-187, February 1962.

[5] H. Shu et al., "Moment-based approaches in imaging, Part 1: Basic features," IEEE Engineering Medicine & Biology Megazine, 2007.

[6] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 489-497, May 1990.

[7] A. Lapedriza, et. al., "Gender recognition in non-controlled environment," IEEE International Conference on Pattern Recognition, vol. 3, pp. 834-837, 2006.

[8] E. Makinen and R. Raisamo, "Evaluation of gender classification methods with automatically detected and aligned Faces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 3, pp. 541-547, March 2008.

[9] "OpenCV 1.0, Open Source Computer Vision Library," http://www.intel.com/technology/computing/opencv/.

[10] P. T. Yap and R. Paramesran, "Eigenmoments," Pattern Recognition, vol. 40, no. 4, pp. 1234-1244, April 2007.

[11] R. Jang, "Data clustering and pattern recognition," http://mirlab.org/jang/books/dcpr/index.asp.

[12] C. C. Chang and C. J. Lin, LIBSVM: A library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm (retrieved May 2011).

[13] E. Makinen and R. Raisamo, "An experimental comparison of gender classification methods," Pattern Recognition Letters, vol. 29, no. 10, pp. 1544-1556, July 2008.

 ISBN: 978-1-61208-283-7