Offline Face Recognition System Based on Gabor-Fisher Descriptors and Hidden Markov Models

Zineb Elgarrai[1], Othmane Elmeslouhi[2], Mustapha Kardouchi[3], Hakim Allali[1], Sid-Ahmed Selouani[4]

[1] *FST of Hassan 1st University Settat /LAVETTE Laboratory,*
[2] *FPO of Ibnou Zohr University /LabSIE Laboratory*
[3] *Université de Moncton /Département d'Informatique,*
[4] *Université de Moncton/Département de Gestion de l'Information*

Abstract — **This paper presents a new offline face recognition system. The proposed system is built on one dimensional left-to-right Hidden Markov Models (1D-HMMs). Facial image features are extracted using Gabor wavelets. The dimensionality of these features is reduced using the Fisher's Discriminant Analysis method to keep only the most relevant information. Unlike existing techniques using 1D-HMMs, in classification step, the proposed system employs 1D-HMMs to find the relationship between reduced features components directly without any additional segmentation step of interest regions in the face image. The performance evaluation of the proposed method was performed with AR database and the proposed method showed a high recognition rate for this database.**

Keywords— **Face Recognition, Hidden Markov Models, Gabor wavelets, Fisher's Discriminant Analysis**

I. Introduction

Face recognition from still images and video sequences has become an important part of user authentication and security infrastructure in recent years. Face recognition (FR) system consists of two major tasks: face feature extraction and face classification (Fig 1). Both of them, have an important impact on the performances of recognition method.

Feature extraction approaches are classified usually into two main categories: feature-based methods where features are extracted from local facial features and holistic methods where features are extracted from the whole face image [1] [2] [3].

In this paper we used a local face descriptor based on Gabor features [4] [5]. Gabor features have been widely used in face identification because of their good performances in illumination and facial expressions change. This robustness is due to the fact that Gabor kernels captures salient visual properties such as spatial localization, orientation selectivity and spatial frequency characteristic [6][7]. However, Gabor magnitude features have a very high dimensionality which needs long computational time. To overcome this dimensionality issue, the Gabor features magnitude are projected to a subspace using Fisher's Discriminant Analysis method (also known as Linear Discriminant Analysis (LDA)) to select the most discriminative features representing the most important information [5].

To deal with the classification task, many methods have been proposed in this topic. Some approaches are based on neural network [8], SVM [9] and HMMs [10] [11] [12] [13] [14].

The methods based on HMMs can be classified roughly into three categories: 1D-HMMS [10] [14], pseudo 2D- HMMs [13] and full 2D-HMMs [11] [15]. In this paper we propose a new way based on

1D-HMMs for the face classification step. The proposed approach is somewhat different from the conventional 1D-HMMs methods, by the fact that it works well without the need to previously localize significant facial regions like eyes, nose, mouth, etc. Thereby, the method can be used even if there are no frontal images in the database. The rest of the paper is organized as follows. In Section 2, Gabor-Fisher features are briefly resumed, and, in Section 3, the proposed HMM approach is described. Experimental results are presented in Section 4. Finally, in Section 5, conclusions are drawn.

II. Features Extraction

Gabor wavelets (also called Gabor kernels or filters) have proven themselves to be a powerful tool for facial image feature extraction and recognition. The 2D Gabor wavelet can be represented by two components: a real and an imaginary components [4]:

-*real component:*

$$g(z, \lambda, \theta, \varphi, \sigma, v) = \exp(\frac{\dot{x}^2 + v^2 y'^2}{2\sigma^2})\cos(2\pi\frac{x'}{\lambda} + \varphi) \quad (1)$$

-*imaginary component:*

$$g(z, \lambda, \theta, \varphi, \sigma, v) = \exp(\frac{\dot{x}^2 + v^2 y'^2}{2\sigma^2})\sin(2\pi\frac{x'}{\lambda} + \varphi) \quad (2)$$

where x'= x cos θ + y sin θ and y' = x sin θ + y cos θ and (x , y) is the coordinate of a pixel in the image plan .and λ represents the wavelength of the sinusoidal factor, θ represents the orientation of a Gabor function, φ is the phase offset, σ is the standard deviation of the Gaussian envelope and is the spatial aspect ratio.

The Fig. 2 shows the result of the convolution of a facial image with 40 banks of Gabor kernel (8 orientations and 5 scales). However, for a given face image, even for a small face image of, for example, 64 × 64 pixels, the 40 magnitude responses reside in a 163840 (64 ×64 × 40) dimensional space, which is far too extensive for efficient processing and storage. To overcome this dimensionality issue, we use the fisher's discriminant analysis (also called LDA) dimensionality reduction technique to project the Gabor magnitude feature vectors into a subspace where between-class variations of the projected patterns are maximized while within-class variations are minimized [16]. To avoid singularity issues, when computing the inverse of the within-class scatter matrix, the LDA reduction method is implemented in the Principal Component Analysis PCA subspace as suggested in [17]. The resultant descriptors are called: Gabor-Fisher descriptors. We notice that for the LDA, the best features dimension is the number of
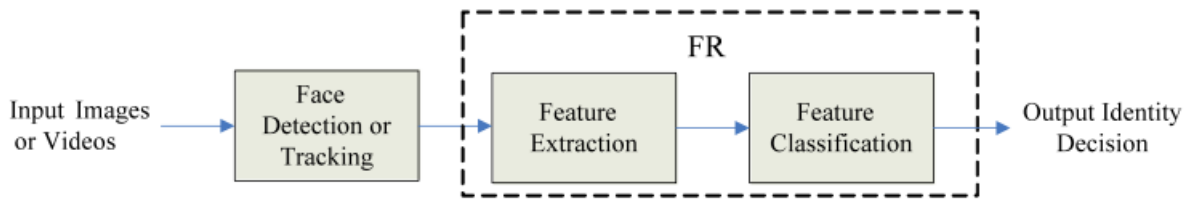
Fig. 1. Face recognition system architecture.

subject minus one [18]. Finally, each Gabor-Fisher feature obtained is sampled to small sequences with a size L. By this way, we avoid any supplementary segmentation of face regions as depicted in previous face recognition methods based on 1D-HMMs.
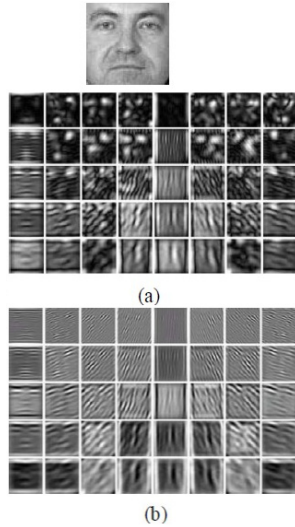

(a)

(b)

Fig. 2. An example of the Gabor magnitude and phase output: (a) the magnitude and (b) the phase output of the filtering operation with a bank of 40 Gabor filters

III. Features Classification

An HMM is a Markov chain with a finite number of unobservable states [25]. Although the Markov states are not directly observable, each state has a probability distribution associated with the set of possible observations. As mentioned above, 1D-HMMS approach is used in the recognition step. Generally, the elements of 1D-HMMs are defined by the triplet $\lambda = \{A, B, \Pi\}$ where, A is an $N \times N$ state transition matrix that gives the state transition probabilities between N states, B is an $k \times N$ emission probability matrix while being in a particular state, and Π is an $1 \times N$ matrix, called initial state probability matrix, and it gives the probability of being in a particular state at the start of the process. To use the proposed system we need three steps:
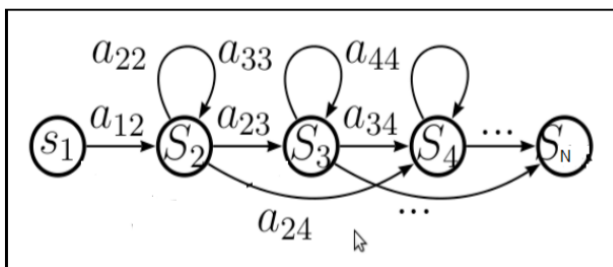
Fig. 3. An example of a Left-Right 1D-HMM with N states

A. System configuration

As mentioned above, we use a 1D left to right Hidden Markov Model (see Fig.3). In this model, the number of states N is equal to the number of classes in dataset minus one divided by L (size of sampled features obtained in the features extraction step). The entry state 1 and the exit state N are non-emitting. The transition matrix A will have N rows and N columns. We assume that each observation probability distribution is represented with single Gaussian distributions. In this case the density of the observation u_t in state i is thus:

$$b_i(u_t) = \frac{1}{\sqrt{(2\pi)^k \det(\Sigma_i)}} \exp(\frac{-1}{2}(u_t - u_i)^t \Sigma_i^{-1}(u_t - u_i))$$

(3)

Where k is the dimension of u_t, and where u_i and Σ_i are the mean vector and covariance matrix, respectively.

Finally, the Gabor-Fisher features will be divided into two sets: training and testing sets.

B. Training step

During this step, the Gabor-Fisher features training set will be injected in the system to estimate the face model parameters for each subject. These parameters can be estimated using an iterative procedure, known as the Baum–Welch Algorithm (BWA) [19].

C. Recognition step

It follows the training step. In this step, the proposed system utilizes the Viterbi algorithm [19] to find the highest likelihood score between features in the testing sets and models obtained in the training step.

IV. Experimental Result

In this section, some experiment results will be given to evaluate the proposed method. The cropped AR face database [20] is used for this purpose. The AR face database contains over 3,200 frontal color face images of 126 subjects (26 different images for each person), including different facial expressions, with various occlusions and under different lighting conditions. Most of the pictures were recorded in two sessions (separated by two weeks). All images were taken by the same camera under tightly controlled conditions of illumination and viewpoint.

For our experiments, like in the work of [20], in the first time, 100 different subjects (50 males and 50 females) were randomly selected from this database, then, all selected images were segmented using an oval-shaped mask and finally all color images are transformed into gray images. Fig. 4 shows some sample images extracted from the obtained face database.
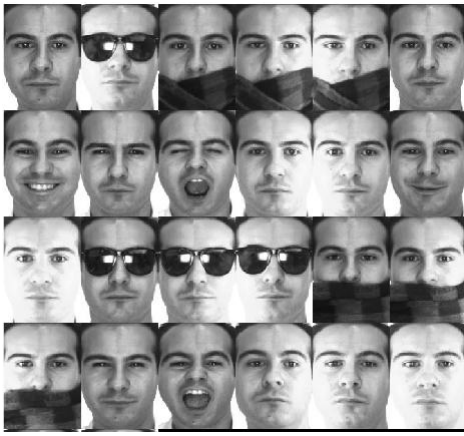
Fig. 4. Sample images from the AR face database

In this experiment part, we use Hidden Markov Models Toolkit (HTK) [21]. The HTK toolkit is builded initially for speech recognition system. We followed the same steps as in [22] to adapt this toolkit to the face recognition domain.

The first set of experiments was performed to discover the optimal number of states per model. To find the best value of this parameter, tables 1 show recognition rates obtained when using different number of states. Obtained results show that the best recognition rate is obtained when the number of states equals 18 for AR database and it shows high sensitivity when moving from 18 states to 19 states.

Furthermore, to prove the effectiveness of the proposed approach we vary sizes of testing and training sets. Table 2 illustrates the recognition rates for various training sizes.

From the above results we can see that for the AR database the recognition rate reaches the best value (i.e 100%) when the size of the training sets is equal to 15.

Finally, we make a comparison between the gabor descriptor with others descriptors like: HOG [23] and LBP [24]. We notice that the size of features of those three descriptors is reduced using Fisher's Discriminant Analysis method. The comparison results are shown in table 3. From this table, we notice that the results obtained by the gabor filters outperform those obtained using HOG and LBP descriptors in the AR datasets.

TABLE I
RECOGNITION RATE FOR DIFFERENT NUMBER OF STATES FOR
AR DATABASE.

Number of states	RECOGNITION RATE (%)
13	66.55
15	79.45
17	96.91
18	100
19	none

TABLE II
RECOGNITION RATE VS NO. OF TRAINING SAMPLES

No. of Training Samples Recognition Rate	No. of Training Samples Recognition Rate
9	94,35%
12	98,14%
14	99,83%
15	100%
9	94,35%

TABLE III
FACE RECOGNITION RATES USING DIFFERENTS DESCRIPTORS
WITH 15 TRAINING SAMPLES FOR AR DATASETS.

Descriptors Recognition Rate	Descriptors Recognition Rate
LBP	94 ,18%
HOG	98 ,12%
Gabor-Fisher	100%

V. CONCLUSION

A new face recognition system based on 1D-HMM was presented in this paper. In the first, the system extract facial features using Gabor filters and reduce their dimensionality using Fisher's Discriminant Analysis method. Secondly the resultant descriptors are resampled and injected in a 1D- HMM to achieve the training and recognition steps. The AR standard database is used to evaluate the proposed system. The recognition rate reaches 100%.

Our future research will be focused on testing the performances of the proposed system in low resolution facial images in order to integrate it in real applications such as surveillance.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Comput. Surv., vol. 35, no. 4, pp. 399–458, Dec. 2003.

[2] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition," Comput. Vis. Image Underst., vol. 101, no. 1, pp.

[3] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2d and 3d face recognition: A survey," Pattern Recogn. Lett., vol. 28, no. 14, pp. 1885–1906, Oct. 2007.

[4] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," Trans. Img. Proc., vol. 11, no. 4, pp. 467–476, Apr. 2002.

[5] V.Struc and N. Pavesic, "The complete gabor-fisher classifier for robust face recognition," EURASIP Advances in Signal Processing, vol. 2010, p. 26, 2010.

[6] U. Park, Y. Tong, and A. K. Jain, "Age-invariant face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 5, pp. 947–954, May 2010.

[7] J. Sung and D. Kim, "Pose-robust facial expression recognition using view-based 2d + 3d AAM," IEEE Transactions on Systems, Man, and Cybernetics, Part A, vol. 38, no. 4, pp. 852–866, 2008.

[8] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," Trans. Neur. Netw., vol. 8, no. 1, Jan. 1997.

[9] G. Guo, S. Z. Li, and K. L. Chan, "Face recognition by support vector machines," in Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on. IEEE, 2000, pp. 196–201.

[10] F. Samaria, "Face recognition using hidden markov models," Phd thesis, Cambridge University, 1994.

[11] H. Otluman and T. Aboulnasr, "Low complexity 2-d hidden markov model for face recognition," in Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on, vol. 5. IEEE, 2000, pp. 33–36.

[12] V. Kohir and U. Desai, "Face recognition using dct-hmm approach," in Workshop on Advances in Facial Image Analysis and Recognition Technology (AFIART) Germany, 1998.

[13] A. V. Nefian, M. H. Hayes, and III, "An embedded hmm-based approach for face detection and recognition," in In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999, pp. 3553–3556.

[14] M. Bicego, U. Castellani, and V. Murino, "Using hidden markov models and wavelets for face recognition," in Proceedings of IEEE International Conference on Image Analysis and Processing (ICIAP03, 2003, pp. 52–56.

[15] H. Othman and T. Aboulnasr, "A simplified second-order HMM with application to face recognition," in International Symposium on Circuits and Systems (ISCAS 2001), 6-9 May 2001, Sydney, Australia, 2001, pp. 161–164.

[16] T. Savic and N. Pavesic, "Personal recognition based on an image of the palmar surface of the hand," Pattern Recogn.,vol. 40, no. 11, pp. 3152–3163, 2007.

[17] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman,"Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 19, pp. 711–720, 1996.

[18] M. Loog, B. van Ginneken, and R. Duin, "Dimensionality reduction by canonical contextual correlation projections," in Computer Vision - ECCV 2004, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, vol. 3021.

[19] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," Proceedings of the IEEE,vol. 77, no. 2, pp. 257–286, 1989.

[20] A. M. Martınez and A. C. Kak, "Pca versus lda," IEEE Trans. Pattern Anal. Mach. Intell., pp. 228–233, 2001.

[21] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey et al.,"The htk book (for htk version 3.4)," Cambridge university engineering department, vol. 2, no. 2, pp. 2–3, 2006.

[22] Z. Elgarrai, O. E. Meslouhi, H. Allali, M. Kardouchi, and S.-A. Selouani, "Face recognition system using gabor features and htk toolkit," in Proceedings of the 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems, ser. SITIS '14, 2014, pp. 32–36.

[23] O. Deniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," Pattern Recogn. Lett., vol. 32, no. 12, pp. 1598–1603, 2011.

[24] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, 2002.

[25] H. Bolivar, M. Martínez, R. González-Crespo, O. Sanjuán. "Complexity analysis of a matchmaker based on hidden Markov model for decentralized grid scheduling", vol. 5, no. 3, pp. 190-197., 2014.

Zineb Elgarrai received her engineer degree in Computer science in 2009 at the National School of computer science and systems analysis (ENSIAS)
School, Rabat, Morocco. In 2012 she joined the LAVETE Laboratory of FST of Hassan 1st University, Settat, Morocco. Her actual main research interests concern Face Recognition and Computer Vision. E-mail: elgarrai@gmail.com

Othmane Elmeslouhi is an Associate Professor in the Mathematics and Computer Science Department of the Polydisciplinary Faculty Ouarzazate of Ibnou University.
He has a PhD from the University of 1st University Settat in computer engineering. His research interests include Computer Vision, Pattern Recognition and Medical Imaging. E-mail: o.elmeslouhi@uiz.ac.ma

Mustapha Kardouchi is a full professor at the Computer Science Departement, Moncton University, Canada. His research interests are primarily in the field of Computer Vision with a focus on Image Recognition and Video Analysis.
E-mail : mustapha.kardouchi@umoncton.ca

Sid Ahmed Selouani is a full professor at the Information Management Departement, Moncton University, Canada. His research interests are Speech Recognition and their applications.
E-mail: sid-ahmed.selouani@umoncton.ca

Hakim Allali was born in Morocco on 1966. He received the Ph.D degree from Claude Bernard Lyon I University (France) in 1993 and the "Docteur d'Etat" degree from Hassan II-Mohamedia University, Casablanca (Morocco) in 1997. He is currently Professor at Faculty of Sciences and Technologies of Hassan 1st University of Settat (Morocco) and director of LAVETE Laboratory.
He is executive manager and founder of IT Learning Campus. His research interests include technology enhanced learning, modeling, image processing, computer networking and GIS.E-mail: hakim-allali@hotmail.fr