

FACE RECOGNITION COMMITTEE MACHINE

Ho-Man Tang, Michael R. Lyu and Irwin King

Department of Computer Science and Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong
{hmtang, lyu, king}@cse.cuhk.edu.hk

ABSTRACT

Face recognition has been of interest to a growing number of researchers due to its applications on security. Within past years, there are numerous face recognition algorithms proposed by researchers. However, there is no unified framework for the integration. In this paper, we implement different existing well-known algorithms, Eigenface, Fisherface, Elastic Graph Matching (EGM), Support Vector Machine (SVM) and neural network, to give a comprehensive testing under same face databases. Moreover, we present a Face Recognition Committee Machine (FRCM), which is a novel approach for assembling the outputs of various face recognition algorithms to obtain a unified decision with improved accuracy. The machine consists of an ensemble of the above algorithms to cope with various face images. We have tested our system with ORL face database and Yale face database. A comparative experimental result of different algorithms with the committee machine demonstrates that the proposed system achieves improved accuracy over the individual algorithms.

1. INTRODUCTION

Face recognition has raised extensive attentions since 1990. The trend is driven by increasing demands on security applications like access control, authentication and identification. There are numerous algorithms proposed by researchers which claimed to have satisfactory result. However, the algorithms are tested under different frameworks. Therefore, we give a comprehensive comparison of the five well-known algorithms (Eigenface, Fisherface, EGM, SVM and Neural network) on same databases in this paper.

We present a novel Face Recognition Committee Machine consisting of five experts above. It fuses the knowledge acquired by the experts to arrive at a unified decision. Each expert shows various performance on different conditions. By assembling the results of the experts, we can obtain a final decision with better accuracy over individuals.

The rest of the paper is organized as follows. Section 2 gives a brief review of the algorithms. Section 3 describes our FRCM system. Section 4 presents and discusses the experimental results. A conclusion and the future work are given in Section 5.

2. ALGORITHMS REVIEW

Eigenface [1] works by finding eigenvectors and eigenvalues of covariance matrix C from training set images $\{T_1, \dots, T_M\}$

$$C = \frac{1}{M} \sum_{i=1}^M (T_i - \psi)(T_i - \psi)^T, \quad (1)$$

where ψ is the average face. By projecting the images on the face space formed by the eigenvectors, we can compute their Euclidean distance efficiently. The training set image with minimum distance from the test image would be the best match in recognition.

Fisherface [2] is similar to Eigenface but it uses FLD instead of PCA. FLD projects away variation in lighting and facial expression while maintaining discriminability by choosing an optimal projections as follows:

$$W_{\text{opt}} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|}, \quad (2)$$

where S_B and S_W is the between-class and within-class scatter matrix respectively.

Elastic Graph Matching [3] is based on the dynamic link architecture. Each facial feature is extracted by Gabor wavelet transform on the fiducial points as a jet. A face is represented by an image graph G consisting of N nodes of jets. Test image graph G^I is compared to all modal graphs G^M by the cost function:

$$C_{\text{total}}(G^I, G^M) = \lambda S_e(G^I, G^M) - S_v(G^I, G^M), \quad (3)$$

where λ is rigidity coefficient, S_e is edge comparison function and S_v is vertex similarity function. The training set image with minimum cost would be the best match.

Support Vector Machine [4] is based on Structural Risk Minimization principle. For linearly separable data, SVM looks for a separating hyperplane which separates the data with the largest margin. For linearly non-separable data, it maps the data into a high dimensional space $x \in \mathcal{R}^J \mapsto \Phi(x) \in \mathcal{R}^h$ with kernel function $\Phi(x)$ to find the hyperplane[5]. As SVM was originally developed for two-class classification, multi-class classification can be extended by using "one-against-one" or "one-against-all" approaches.

3. FACE RECOGNITION COMMITTEE MACHINE

Committee machine has been widely used in neural networks. A number of researchers have applied it to improve the performance of a neural network [6][7]. The basic idea of a committee machine is to ensemble a mixture of experts and to combine the result of each expert. There are mainly two kinds of committee machines:

1. **Static Structure:** This is generally known as an ensemble method. Input data is not involved in combining the committee experts. Examples includes ensemble averaging and boosting.
2. **Dynamic Structure:** Input is directly involved in the combining mechanism that employs an integrating unit to adjust the weight of each expert according to the input.

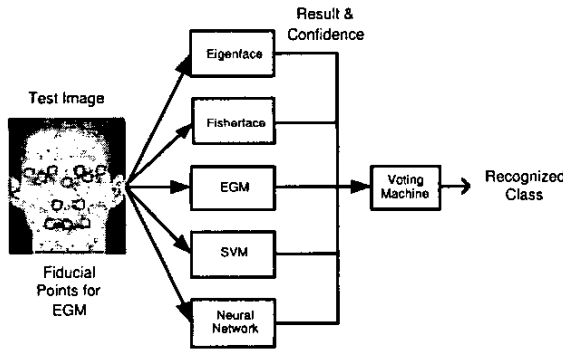


Fig. 1. FRCM System Overview and Fiducial points

Figure 1 provides an overview of our FRCM system and table 1 lists some implementation details for reference. Our proposed FRCM adopts the static structure with five well-known experts in face recognition. All are proven with good classification ability in the literature [8][9]. Among the algorithms, Eigenface, Fisherface are template matching methods and EGM is a graph matching method. Due to the difference in nature of the algorithms, each expert shows various performance on different conditions. By combining the result of the experts, we can arrive at a final result with improved performance.

Table 1. Implementation details of FRCM

Implementation	Detail Description
Committee Machine	Static structure with 5 expert
	Average Performance used as weight function
Eigenface	50 eigenvectors used
EGM	40 Gabor filters (8 orient. & 5 freq.) used
	12 fiducial points selected (Fig 1) manually
SVM	Polynomial kernel function used
	"One-against-One" approach used
Neural Network	Feed forward backpropagation network
	Fisher projection used as feature vector
	40 hidden nodes used
Programming Lang.	Visual C++
Operating System	Windows

Each expert gives its **result** r and **confidence** c for the result to the voting machine. We introduce the use of confidence as a weighted vote for the voting machine to avoid low confidence result of individual expert from affecting the final result. In order to find results and confidences of various algorithms, we adopt different approaches.

- **Eigenface, Fisherface and EGM:** We use K nearest-neighbor classifiers. Five nearest training set projections with the test image projection are chosen for Eigenface and Fisherface, and five training set graphs with the lowest cost are chosen for EGM. The final result for expert i is defined as the class j with the highest votes v in J classes among the five results:

$$r(i) = \arg \max_j (v(j)), \quad (4)$$

where its confidence is defined as the number of votes of the result class divided by K , i.e.,

$$c(i) = \frac{v(r(i))}{K}. \quad (5)$$

- **SVM:** To recognize a test image in J different classes, ${}_J C_2$ (i.e., $\frac{J(J-1)}{2}$) SVMs are constructed. The image is tested against each SVM and the class j with the highest votes in all SVMs is selected as the recognition result $r(i)$. The confidence is defined as the number of votes of the result class divided by $J - 1$:

$$c(i) = \frac{v(r(i))}{J - 1}, \quad (6)$$

where $J - 1$ is the maximum number of vote a class could obtain.

- **Neural network:** We choose a binary vector of size J for the target representation. The target class is set to one and the others are set to zero. The class j with output value closest to 1 is chosen as the result and the output value is chosen as the confidence.

The weights in FRCM are evaluated in our testing for different algorithms under ORL and Yale face database. We take the average accuracy for the algorithms as weights (shown in table 2 and table 3 respectively). The use of weights in the voting machine further reduces the chance for an expert who performs poorly on average from affecting the ensemble result even if it has high confidence on the result. After collecting the results r and confidences c from the five experts, the voting machine assembles the results by calculating the score s of each class as follows:

$$s(j) = \sum_{i=1}^5 w(i) * c(i), \forall j \in r(i). \quad (7)$$

The class with the highest score would be selected as the recognized class of our FRCM. We define the score in such a way that only experts with high performance on average and high confidence on the result would take most significant score in the final decision.

4. EXPERIMENTAL RESULTS

Two sets of experiments are presented to evaluate the performance of FRCM and individual algorithms. We adopt leaving-one-out cross validation method for the experiment. For a given sample of n images in a class, a classifier is trained using $(n - 1)$ images in that class and tested on the remaining single case. The test repeats n times, each time training a classifier with leaving-one-out. Thus, all images are used for training and testing to produce a thorough result.

4.1. The ORL Database of Faces



Fig. 2. Snapshot of ORL database

First experiment is performed on the ORL face database from AT&T Laboratories Cambridge. The images are gray-scale with a resolution of 92×112 pixels. The database contains 400 images, including 40 distinct people, each with

10 images that vary in position, rotation, scale and expression. The images are taken under constant lighting condition. Figure 2 shows a snapshot of 4 individuals.

From the ORL result shown in Table 2, FRCM (98.8%) has improvement in accuracy over the individual algorithms in the testing. We notice that Fisherface and SVM obtain higher accuracy (over 97%) than the others. This is due to the fact that both Fisherface and SVM inherits better classification ability in general cases. We can see the effect of the committee machine in image set 7 that none of the experts has 100% accuracy but FRCM achieves it. The result also demonstrates that with the use of confidence and weight function, poor result from some experts would not affect the ensemble result significantly.

Table 2. ORL Result

Image Set	Eigen	Fisher	EGM	SVM	NN	FRCM
1	92.5%	100.0%	90.0%	95.0%	92.5%	95.0%
2	85.0%	100.0%	72.5%	100.0%	95.0%	100.0%
3	87.5%	100.0%	85.0%	100.0%	95.0%	100.0%
4	90.0%	97.5%	70.0%	100.0%	92.5%	100.0%
5	85.0%	100.0%	82.5%	100.0%	95.0%	100.0%
6	87.5%	97.5%	70.0%	97.5%	92.5%	97.5%
7	82.5%	95.0%	75.0%	95.0%	95.0%	100.0%
8	92.5%	95.0%	80.0%	97.5%	90.0%	97.5%
9	90.0%	100.0%	72.5%	97.5%	90.0%	100.0%
10	85.0%	97.5%	80.0%	95.0%	92.5%	97.5%
Average	87.5%	98.3%	77.8%	97.8%	93.0%	98.8%

4.2. Yale Face Database



Fig. 3. Snapshot of cropped Yale database

Second experiment is performed on Yale face database from Yale University. The images are gray-scale and are cropped to a resolution of 116×136 pixels. The database contains 165 images, including 15 distinct people, each with 11 images that vary in both expression and lighting. A snapshot of 4 individuals in the database is shown in Figure 3. The result of FRCM on Yale database to classify the 15 people under different conditions is given in Table 3.

Table 3. Yale Result

Image Set	Eigen	Fisher	EGM	SVM	NN	FRCM
centerlight	53.3%	93.3%	66.7%	86.7%	73.3%	93.3%
glasses	80.0%	100.0%	53.3%	86.7%	86.7%	100.0%
happy	93.3%	100.0%	80.0%	100.0%	93.3%	100.0%
leftlight	26.7%	26.7%	33.3%	26.7%	26.7%	33.3%
noglasses	100.0%	100.0%	80.0%	100.0%	100.0%	100.0%
normal	86.7%	100.0%	86.7%	100.0%	93.3%	100.0%
rightlight	26.7%	40.0%	40.0%	13.3%	26.7%	33.3%
sad	86.7%	93.3%	93.3%	100.0%	93.3%	100.0%
sleepy	86.7%	100.0%	73.3%	100.0%	100.0%	100.0%
surprised	86.7%	66.7%	33.3%	73.3%	66.7%	86.67%
wink	100.0%	100.0%	66.7%	93.3%	93.3%	100.0%
Average	75.2%	83.6%	64.2%	80.0%	77.6%	86.1%
No light	85.9%	94.8%	70.4%	93.3%	88.9%	97.8%

From the Yale result, FRCM (86.1%) also outperforms all the individuals on average. The main reason for some non-satisfactory result is due to the fact that Yale database contains variations in left and right lighting (4th and 7th column in Fig. 3). The accuracy for both leftlight and rightlight in FRCM is 33.0% only. For algorithms taking the whole image as input like Eigenface, the accuracy would drop significantly because the lighting would greatly affect the pixel values. We notice that EGM works relatively better in the light testings than other algorithms. This is due to the use of Gabor wavelet transformation of fiducial points in EGM rather than in the whole image. Without the lighting variations, FRCM achieves 97.8% accuracy, which is comparable to the ORL result (98.8%).

5. CONCLUSION AND FUTURE WORK

In this paper, we perform a comprehensive experiment on five well-known face recognition algorithms to compare the accuracy of the algorithms under the same framework. We conclude that Fisherface and SVM are the best classifiers among them. Both achieves over 93% accuracy in general cases. However, none of them has high accuracy under lighting variation in Yale test.

Moreover, we propose a Face Recognition Committee Machine. We introduce the use of confidence on experts' results and weight function on the committee machine which can reduce the chance for poor result of certain expert from affecting the ensemble result. The success has been demonstrated on the result of ORL and Yale test. It achieves 98.8% and 97.8%(without lighting variation) accuracy respectively which outperforms all other individual.

In the Yale test, we notice that FRCM does not perform satisfactorily on rightlight and leftlight testing. The reason for this is due to the lack of an expert in the committee machine which can accurately recognize a face under various lighting condition. Our future work will focus on including an expert for lighting variation like Illumination Cone [10]

in order to make further improvement.

6. ACKNOWLEDGMENTS

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK 4360/02E & 4351/02E).

7. REFERENCES

- [1] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1991, pp. 586-591.
- [2] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, July 1997, vol. 19, pp. 711-720.
- [3] M. Lades, J.C. Vorbruggen, C. von der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," in *IEEE Trans. on Computers*, March 1993, vol. 42, pp. 300-311.
- [4] Vladimir Vapnik, *Statistical Learning Theory*, John Wiley and Sons, Inc., 1998.
- [5] J. C. Burges Christopher, "A tutorial on support vector machines for pattern recognition," in *Data Mining and Knowledge Discovery*, 1998, vol. 2, pp. 121-167.
- [6] Zhang B.T. and Joung J.G., "Building optimal committees of genetic programs," in *Parallel Problem Solving from Nature - PPSN VI*, 2000, pp. 231-240.
- [7] Min Su and M. Basu, "Gating improves neural network performance," in *Proc. IEEE Conf. on IJCNN '01*, 2001, vol. 3, pp. 2159-2164.
- [8] Jun Zhang, Yan Yong, and M. Lades, "Face recognition: eigenface, elastic matching, and neural nets," in *Proc. IEEE*, 1997, vol. 85, pp. 1423-1435.
- [9] Guo Guodong, S.Z. Li, and Chan Kapluk, "Face recognition by support vector machines," in *Proc. Fourth IEEE International Conf. on Automatic Face and Gesture Recognition 2000*, 2000, pp. 196-201.
- [10] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, June 2001, vol. 23, pp. 643-660.