A Face Processing System Based on Committee Machine: The Approach and Experimental Results

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

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

Abstract. In this paper, we propose a heterogeneous committee machine for face processing including face detection and recognition. Our proposed system consists of two components, Face Detection Committee Machine (FDCM) and Face Recognition Committee Machine (FRCM), which employs three and five well-known state-of-the-art approaches respectively. We engage different methodologies to solve the face detection and face recognition problems. We provide a rigorous architecture set-up and experimentation protocol to demonstrate the improved performance of FDCM and FRCM over the individual experts.

1 Introduction

In recent years, the committee machine, an ensemble of estimators, has proven to give more accurate results than the use of a single predictor. The basic idea is to train a committee of estimators and combine the individual predictions to achieve improved generalization performance. Different approaches are proposed by researchers within the last ten years such as ensemble averaging, bagging, gating network and hierarchical mixtures of experts [1]. Recently, researchers have applied the committee machine in face processing, Gutta et al. used an ensemble of Radial Basis Function (RBF) network and decision tree in the face processing problem [2]. Huang et al. formulated an ensemble of neural networks for pose invariant face recognition [3].

Previous researchers applied homogeneous experts (neural networks or RBF) trained by different training data sets to arrive at a union decision. However, no one has yet focused on heterogeneous experts on face detection and recognition problems in the current status. We propose the engagement of committee machine with heterogeneous experts in this paper. This is the first effort to employ different state-of-the-art algorithms as heterogeneous experts on committee machines in face detection and recognition. We include neural networks (NN), Sparse Network of Winnows (SNoW) algorithm, and SVM in face detection to validate any possible face images. In face recognition, we investigate Eigenface, Fisherface, EGM, SVM and NN to classify a face image. All the algorithms
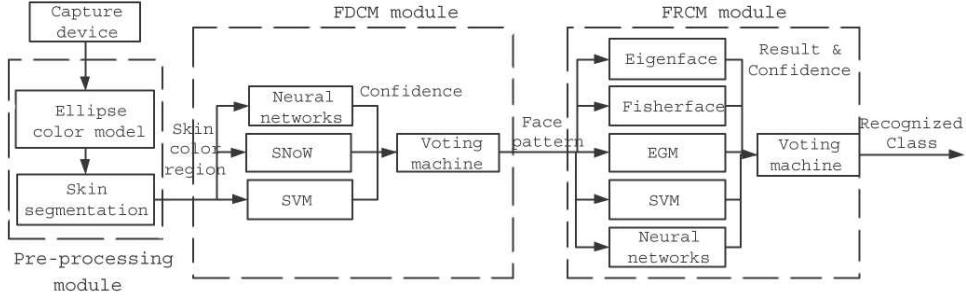
Fig. 1. The system architecture

used are well-known in the field. The results based on the committee machine approach are satisfactory.

2   Background

Within these few years, numerous face detection and recognition methods were proposed by researchers. Turk and Pentland applied Principal Component Analysis (PCA) to face detection [4]. In [5], Rowley et al. used neural networks to learn face and non-face patterns for face detection. Support Vector Machine (SVM) was studied by [6] and demonstrated for the success in detecting frontal faces. Roth et al. proposed Sparse Network of Winnows algorithm [7], which applied the primitive feature space for the learning process. Among the face recognition methods, Eigenface [8] is the most popular one due to its effectiveness. It made use of PCA to find a feature space for projection of face images. A similar approach, Fisherface [9], was proposed later which makes use of Fisher's Linear Discriminant (FLD) instead of PCA. Apart from template matching approaches, Elastic Graph Matching (EGM) [10] was proposed to take into account the human facial features by extracting the features with Gabor wavelet transform. Recently, SVM [11] has gained a wider acceptance in face recognition and were proven with impressive result.

3   A face processing system

We propose a face processing system consisting of three main modules: 1) Pre-processing, 2) **Face Detection** and 3) **Face Recognition**. The system architecture is shown in Fig. 1. We employ the color model to reduce the search space for face finding in the Face Detection module. The detected face is then passed to the Face Recognition module for further recognition.
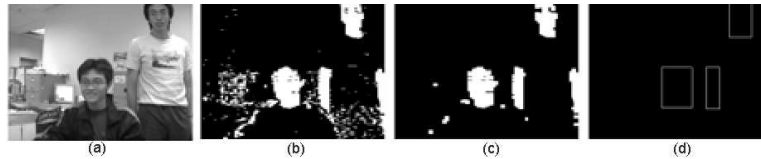
Fig. 2. The image for (*a*) original image, (*b*) binary skin mask, (*c*) binary skin mask after morphological operation and (*d*) face candidates are found
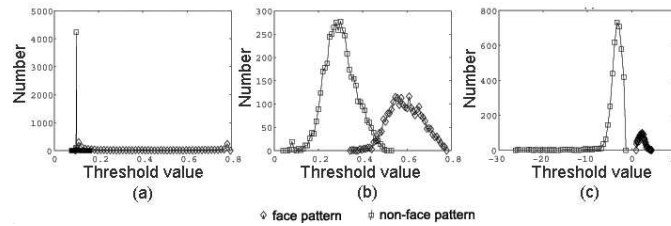

Fig. 3. The distribution of confident value of the training data from *(a)* NN, *(b)* SNoW and *(c)* SVM

3.1 Pre-processing

Firstly, image data from the capture device or video file are transformed from RGB color model to YCrCb color model. In many studies, human skin color is not a uniform distribution around the whole color space. Most of human skin tone can be represented by an elliptical equation in the YCrCb color space [12]. We use the ellipse color model to locate all the flesh color to form a binary mask. Morphological operation is then applied to the binary mask to form a new binary mask. This step can reduce noisy components especially when taking image data from the web camera. The skin segmentation is performed to the image by applying the new binary mask to find out face candidates regions from the image. The pre-processing steps are shown in Fig. 2. This module uses the color information to reduce the search space for finding face candidates. New face candidate regions will be passed into the FDCM.

3.2 Face Detection Committee Machine (FDCM)

The FDCM works according to the confidence value T_i of each expert i. However, the confidence value T_i from each of the experts cannot be used directly. Figure 3 shows that the distribution of the confidence value of each expert varies in a large range. Thus, we need to normalize the value T_i by using the statistical information obtained from the training data:

$$\alpha_i = (T_i - \mu_i)/\sigma_i, \tag{1}$$

where T_i is the confidence value from expert i, μ_i is the mean value of training face pattern data from expert i, and σ_i is the standard derivation of training data from expert i.

One of the reasons why we need to normalize the confidence value is that they are not a uniform function. Another reason is that not all the experts have a fixed range of confidence value e.g., [-1,1] or [0,1]. Using statistical approach to model the problem, the information of confidence value from experts can be preserved. The output value of the committee machine can then be calculated using equation:

$$\beta = \sum_i w_i * (\alpha_i + \sigma_i * \delta_i),\qquad(2)$$

where δ_i is the criteria factor for expert i and w_i is the weight of the expert i. The data is classified as face when the value of β is larger than 0 and non-face pattern when the value is smaller than or equal to 0.

3.3 Face Recognition Committee Machine (FRCM)

Our proposed FRCM adopts the static structure with five well-known experts. As shown in Fig. 1, input image is sent into the five experts for recognition. Apart from using result of each expert, 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. Due to different nature of the experts, we adopt different approaches to find the results and the associated confidence.

– **Eigenface**, **Fisherface** and **EGM**: We employ K nearest-neighbor classifiers, where five nearest training set images with the test image are chosen. The final result for expert i is defined as the class j with the highest votes v in J classes among the five results:

$$r_i = \arg\max_j(v(j)),\qquad(3)$$

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

$$c_i = \frac{v(r_i)}{K}.\qquad(4)$$

– **SVM**: As SVM was originally developed for two-class classification, multi-class classification can be extended by using "one-against-one" approach. To recognize a test image in J different classes, $_JC_2$ (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 like Equation 4 with $J-1$ (the maximum number of votes a class) instead of K.

– **NN**: 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.

Table 1. CBCL face database

Table 3. Experimetal results on images from the CBCL testing set

	Training Set	Testing Set
Face Pattern	2429	472
Non-face Pattern	4548	23573

Table 2. CBCL results

	True Positive	False Positive
NN	71.4%	15.2%
SNoW	71.6%	15.1%
SVM	81.2%	13.2%
FDCM	**84.1%**	**11.4%**

	False Alarm Rate			
Detection Rate	**NN**	**SNoW**	**SVM**	**FDCM**
10%	0.56%	0.41%	0.05%	0.02%
20%	1.37%	1.09%	0.16%	0.07%
30%	2.54%	1.67%	0.44%	0.14%
40%	4.11%	2.92%	0.83%	0.41%
50%	6.32%	4.91%	1.60%	0.77%
60%	9.47%	8.47%	3.07%	1.41%
70%	13.89%	14.67%	5.98%	3.90%
80%	**26.97%**	**27.62%**	**12.32%**	**7.79%**
90%	48.95%	49.26%	28.60%	22.92%

The weights w in FRCM are evaluated in our testing for different algorithms under ORL and Yale face database (shown in Table 4 and Table 7 respectively). We take the average accuracy for the algorithms from prior cross validation experiments and normalize them with weights by an exponential mapping function:

$$w_i = \frac{\exp(a_i)}{\sum_{i=1}^{5} \exp(a_i)}, \quad (5)$$

where a_i is the average accuracy of expert i. 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. The voting machine assembles the results by calculating the score s of each class as follows:

$$s_j = \sum_{i=1}^{5} w_i * c_i, \forall j \in r_i. \quad (6)$$

We define the score in such a way that only experts with high performance on average and high confidence on the result would take the most significant score in the final decision.

4 Experimental Results

4.1 FDCM

We applied the CBCL face database from MIT for training and testing each of the expert systems to control the condition. Table 1 shows some data for the CBCL face database. The outputs from each single approach are determined by a threshold. When the threshold increases, the detection rate and the number of false detection will increase at the same time. For the FDCM, output value is calculated based on the experts' confidence values. When we change the value of criteria factor δ_i, the sensitivity of the committee machine will be affected. Input images classified as face patterns will be increased if we increase the value of criteria factor δ_i. This property is shown in Fig. 4.
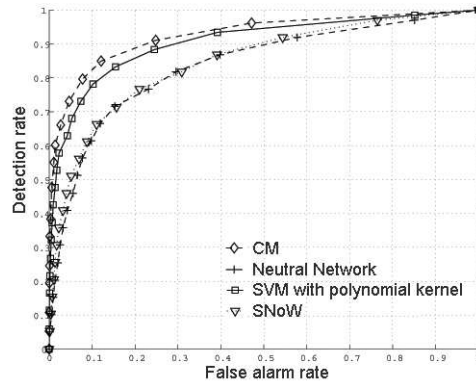
Fig. 4. The ROC curves of three different approaches and committee machine

The Receiver Operating Characteristics (ROC) curves are employed to show the characteristic of each approach. The area under the ROC curve provides a convenient way to compare classifiers. Table 3 shows the false alarm rate of each approach under the same detection rate. The false alarm rate of FDCM is nearly half of the other approach with 80% detection rate. FDCM archives lower false alarm rate than the other methods with same detection rate. Table 2 lists the best operating point of each classifier. Our approach obtains the highest true positive rate while maintaining the lowest false positive rate. By observation, the statistical model of the committee machine can be engaged in the face detection problem and can improve the accuracy in classifying face and non-face pattern.

4.2 FRCM

Two sets of experiments are presented to evaluate the performance of FRCM and the individual algorithms. We adopt leaving-one-out cross validation method for the experiment to produce a thorough result. Table 6 lists the average running time for the algorithms and FRCM on the face databases.

The ORL Database of Faces The experiment is performed on the ORL face database (400 images) from AT&T Laboratories in Cambridge. From Table 4, FRCM (98.8%) achieves 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 inherit better classification ability in general cases. Table 5 shows the details of the underlying data in Image Set 1 and 7 to demonstrate how the committee machine works. We can see the effect of the committee machine in Set 7 that none of the experts has 100% accuracy but FRCM achieves it.

Table 4. ORL recognition results

Set	Eigenface	Fisherface	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%**

Table 5. Result explanation FRCM in Image Set 1 and 7

		Recognized Class(Confidence)					
Set	Image	Eigenface	Fisherface	EGM	SVM	NN	FRCM
1	0	15(0.40)	**0**(0.60)	20(0.20)	15(1.00)	23(0.44)	**15**(0.29)
	34	14(0.60)	**34**(0.80)	28(0.40)	14(1.00)	14(0.63)	**14**(0.46)
7	25	27(0.40)	**27**(1.00)	10(0.20)	25(1.00)	25(0.51)	**25**(0.32)
	34	26(0.40)	**18**(0.60)	34(0.40)	34(1.00)	34(0.37)	**34**(0.36)

Yale Face Database The experiment is performed on Yale face database (165 images) from Yale University. From Table 7, FRCM (86.1%) also outperforms all the individuals on average. The main reason for some non-satisfactory results (i.e., leftlight and rightlight) is due to the fact that Yale database contains variations in strong left and right lighting. 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. Without the lighting variations, FRCM achieves 97.8% accuracy, which is comparable to the ORL result (98.8%).

5 Conclusion

In this paper, we present a heterogeneous committee machine based face processing system which can automatically recognize people from camera. We employ the confidence information on experts' results and weight function on the FDCM and FRCM which can reduce the chance for poor result of certain expert from affecting the ensemble result. The success of the FDCM and FRCM has been demonstrated on the result of CBCL database, and ORL and Yale database, respectively. In our experiment, FDCM achieves 84.1% for true positive rate and 11.4% for false positive rate, which perform better than other three individual approaches. FRCM achieves 98.8% accuracy in ORL test and 97.8% accuracy in Yale test (without lighting variation), which also outperforms other state-of-the-art algorithms. These results show that the use of committee machine works in improving the accuracy of face detection and face recognition.

Table 6. Average running time(s)

	Eigenface	Fisherface	EGM	SVM	NN	FRCM
ORL	2.1	1.5	16.3	6	1.4	27.3
Yale	0.9	0.2	6.5	0.6	0.3	8.5

Table 7. Yale recognition results

Set	Eigenface	Fisherface	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%**

References

1. Jacobs, R.A., Jordan, M.I., Steven, J.N., Geoffrey, E.H.: Adaptive mixtures of local experts. In: Neural Computation. Volume 3. (1991) 79–87
2. Gutta, S., Huang, J.R.J., Jonathon, P., Wechsler, H.: Mixture of experts for classification of gender, ethnic origin, and pose of human faces. IEEE Trans. on Neural Networks **11** (2000) 948–960
3. Huang, F., Zhang, H., Chen, T., Zhou, Z.: Pose invariant face recognition. In: Proc. IEEE Int'l Conf. on Automatic Face and Gesture Recognition. (2000) 245–250
4. Turk, M., Pentland, A.: Face recognition using eigenfaces. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (1991) 586–591
5. Rowley, H., Baluja, S., Kanade, T.: Neural network-based face detection. IEEE Trans. on Pattern Analysis and Machine Intelligence **20** (1998) 23–38
6. Osuna, E., Freund, R., Girosit, F.: Training support vector machines: an application to face detection. In: Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. (1997) 130–136
7. Yang, M.H., Roth, D., Ahuja, N.: A snow-based face detector. In: Advances in Neural Information Processing Systems. Volume 12. (1999) 862–868
8. Sirovich, L., Kirby, M.: A low-dimensional procedure for the characterization of human faces. In: J. Opt. Soc. Amer. A. Volume 4. (1987) 519–524
9. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans. on Pattern Analysis and Machine Intelligence **19** (1997) 711–720
10. Lades, M., Vorbruggen, J., von der Malsburg, C., Wurtz, R.P., Konen, W.: Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. on Computers **42** (1993) 300–311
11. Vapnik, V.: Statistical Learning Theory. John Wiley and Sons, Inc. (1998)
12. Hsu, R.L., Abdel-Mottaleb, M., Jain, A.: Face detection in color images. IEEE Trans. on Pattern Analysis and Machine Intelligence **24** (2002) 696–706