Face Recognition Using Ada-Boosted Gabor Features

Peng Yang[1], Shiguang Shan[1], Wen Gao[1], Stan Z. Li[2], Dong Zhang[2]
[1]*Institute of Computing Technology of Chinese Academy Science*
[2]*Microsoft Research Asia*
[1]*{pyang, sgshan,wgao}@jdl.ac.cn, szli@microsoft.com*

Abstract

Face representation based on Gabor features has attracted much attention and achieved great success in face recognition area for the advantages of the Gabor features. However, Gabor features currently adopted by most systems are redundant and too high dimensional. In this paper, we propose a face recognition method using AdaBoosted Gabor features, which are not only low dimensional but also discriminant. The main contribution of the paper lies in two points: (1) AdaBoost is successfully applied to face recognition by introducing the intra-face and extra-face difference space in the Gabor feature space; (2) An appropriate re-sampling scheme is adopted to deal with the imbalance between the amount of the positive samples and that of the negative samples. By using the proposed method, only hundreds of Gabor features are selected. Experiments on FERET database have shown that these hundreds of Gabor features are enough to achieve good performance comparable to that of methods using the complete set of Gabor features.

1. Introduction

Face recognition has a variety of potential applications in public security, law enforcement and commerce such as mug-shot database matching, identity authentication for credit card or driver license, access control, information security and video surveillance. In addition, there are many emerging fields that can benefit from face recognition, such as human-computer interfaces and e-services, including e-home, tele-shopping and tele-banking. Related research activities have significantly increased over the past few years [1].

The most popular exiting technologies for face recognition include Eigenface (PCA) [2], FisherFace [3], Independent Component Analysis (ICA) [4], Bayesian face recognition [5] and Elastic Bunch Graph Matching (EBGM) [7]. In the FERET test [6], Fisherface, Bayesian matching and EBGM were among the best performers. Especially, the EBGM has attracted much attention because it firstly exploited the Gabor transform to model the local features of faces. However, EBGM takes the complete set of Gabor features, most of which are redundant for classification. For examples, Fasel has pointed out in [8] that the Gabor features used in [7] are not the best ones for the detection of facial landmarks. However, no method has been proposed on how to select the most discriminant Gabor features for recognition purpose. This paper is an attempt to answer this question by introducing the AdaBoost method into the Gabor feature-based face recognition method.

Face recognition is a multi-class problem, therefore, in order to use AdaBoost for classification, as in [5] and [9], we propose to train AdaBoost based on the intra-personal and extra-personal variation in the Gabor feature space. Based on a large database of images, AdaBoost selects a small set of available Gabor features from the extremely large set. The final strong classifier, which combines a few hundreds of weak classifiers (Gabor features), can evaluate the similarity of two face images. The flowchart of recognition process in our system is as following:

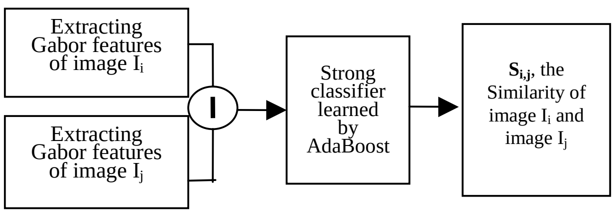


Fig.1. The flowchart of the proposed face recognition method.

A face recognition system comprises two stages: training and testing. In practical applications, the small number of available training face images and the complicated facial variations during the testing stage are the most difficult problems for current face recognition systems. Therefore, a lot of work has been done on training set, including re-sampling, such as [9].

The remaining part of this paper is organized as follows: In section 2, the Gabor representation of face is introduced. Section 3 presents the intra-personal and extra-personal space. Section 4 describes the boosting learning for feature selection and classifier construction. The re-sampling scheme we proposed is conducted in section 5. Experiments and analysis are conducted in section 6, followed by a small discussion, conclusion and future work in section 7.

2. Gaborface

Gabor filter can capture salient visual properties such as spatial localization, orientation selectivity, and spatial frequency characteristics. Considering these excellent capacities and its great success in face recognition [6], we choose Gabor features to represent the face image. Gabor filters are defined as follows:

$$\psi_{u,v}(z) = \frac{\| k_{u,v} \|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2)} \left[e^{i\vec{k}_{u,v} z} - e^{-\sigma^2/2} \right], \quad (1)$$

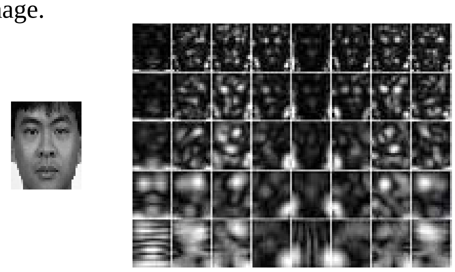
where $k_{u,v} = k_v e^{i\phi_u}$; $k_v = \frac{k_{\max}}{f^v}$ gives the frequency, $\phi_u = \frac{u\pi}{8}, \phi_u \in [0, \pi)$ gives the orientation, and $z = (x, y)$.

$$k_{u,v} = k_v e^{i\phi_u}, \quad (2)$$

where $e^{i \vec{k}_{u,v} z}$ is the oscillatory wave function whose real part and imaginary part are cosine function and sinusoid function respectively. In equation 1, v controls the scale of Gabor filters which mainly determines the center of the Gabor filter in the frequency domain; u controls the orientation of the Gabor filter.

In our experiment we use the Gabor filters with the following parameters: five scales $v \in \{0,1,2,3,4\}$ and eight orientations $u \in \{0,1,2,3,4,5,6,7\}$ with $\sigma = 2\pi$, $k_{\max} = \pi/2$, and $f = \sqrt{2}$. The same parameters are also taken in [7].

The Gaborface, representing one face image, is computed by convoluting it with corresponding Gabor filters. Figure 2 shows the Gaborface representation of a face image.



(a) (b)

Fig.2. Gaborface representation for one face.

The face image is represented by Gaborface, which is used to construct the intra-personal space and the extra-personal space. The construction process will be introduced in the following section.

3. Intra-personal and Extra-personal Space

In FERET96 test, the Bayesian method proposed by Moghaddam and Pentland [5] was the top one performer. Although in FERET97 test it was exceeded by the algorithm of UMD (University of Maryland) [6], it has shown the strong potential in face recognition and other applications of pattern recognition, and has become one of the most widely influential face recognition algorithms.

In nature, the thought of the face recognition method of Moghaddam and Pentland [5] is to convert the multi-class problem into the two-class problem. Basically, face recognition is a multi-class problem. Moghaddam and Pentland [5] used a statistical approach that learned the variations in the different images of an individual to form the intra-personal space, and the variations in the different images of different individuals to form the extra-personal space. Therefore, the multi-class problem is converted into a two-class problem. The estimation of the intra-personal and the extra-personal distributions is based on the assumption that the intra-personal distribution is Gaussian.

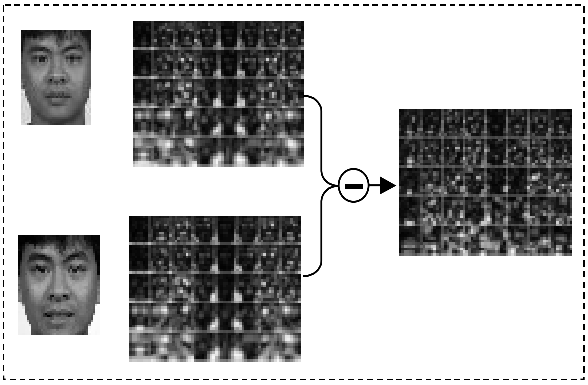
In our system, the definitions of the intra-personal class and the extra-personal class are as follows: $I_{i,k}$ is a face image, where the subscript i means this image belongs to the individual whose ID is i; I_j is a face image of another subject; GI_i means the transformed images got by convoluting I_i with the Gabor filters; GI_j means the transformed images got by convoluting I_j with the same Gabor filters; $H(I_i - I_j) = \|GI_i - GI_j\|$ means the difference of the two images. If $i = j$, $H(I_i - I_j)$ is in the intra-personal space. On the contrary, if $i \neq j$, $H(I_i - I_j)$ is in the extra-personal space. In our system, in the training process, if $i = j$, $H(I_i - I_j)$ is a positive example; otherwise, $H(I_i - I_j)$ is a negative example. Figure 3 shows some different images in intra-personal space and extra-personal space.

In [5], *Maximum a Posterior* (MAP) rule is taken to obtain the two probabilistic similarity measures. Obviously, the intra-personal and extra-personal problem is a two-class problem. As we know, boosting learning is a strong tool to solve two-class classification problems. Noticing the great success of AdaBoost in face detection area, we exploited it in our method to distinguish the intra-personal space from the extra-personal space.

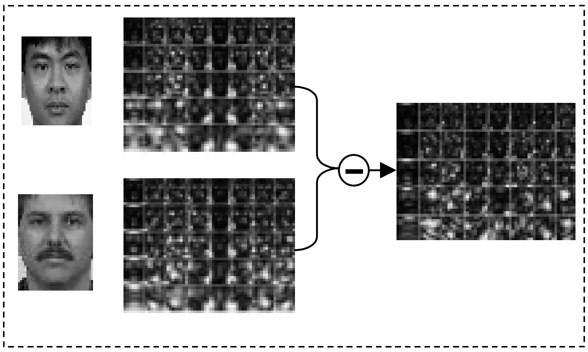
We use AdaBoost to select a small set of Gabor features (or weak classifiers) from the original extremely high dimensional Gabor feature space to form a strong classifier, which is used to calculate the similarity of a pair of Gaborfaces. Equation 3, a strong classifier learned by AdaBoost, is taken to measure their similarity:

$$S(I_i, I_j) = \sum_{m=1}^{M} \alpha_m h_m(I_i, I_j), \tag{3}$$

where a_m is the combining coefficient and $h_m(I_i, I_j)$ is a threshold function. How to derive a_m and $h_m(I_i, I_j)$ will be discussed in the following section.



(a) Intra-personal image



(b) Extra-personal image

Fig.3. Intra-personal image and Extra-personal image represented by Gaborfaces.

4. Learning the most Discriminant Gabor features by AdaBoost

A large number of experimental studies have shown that classifier combination can exploit the discriminating power of individual feature sets and classifiers. With the success of boosting in the application of face detection, boosting, as one of the most commonly used methods of combining classifiers based on statistical re-sampling techniques, has shown strong ability to resolve the two-class problem. For Intra-personal and Extra-personal is used to describe whether two different face images are from the same subject, naturally, AdaBoost, a version of the boosting algorithm, is taken to solve this two-class problem. Therefore, we use AdaBoost to train a strong classifier. The framework of the training process of the proposed method is illustrated in figure 4.

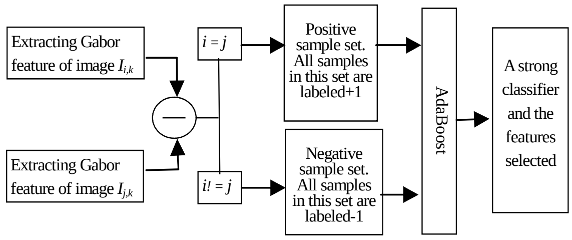


Fig.4. Framework of the proposed training process.

A strong classifier is formed by AdaBoost, which combines a number of weak classifiers. The AdaBoost process is described in Table 1.

Table 1. The AdaBoost algorithm for classifier learning

Given labeled examples Set S and their initial weights ω_1

Do for t=1, … , T:

1. Normalize the weight ω_t
2. For each feature, k, train a classifier h_k with respect to the weighted samples
3. Calculate error, choose the classifier h_t with the lowest error, get α_t , the weight of h_t.
4. Update weights ω_{t+1},

Get the strong classifier $S(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$

$S(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$ of table 1 is re-written as equation (3), $S(x) = S(I_i, I_j) = \sum_{m=1}^{M} \alpha_m h_m(I_i, I_j)$, where $\alpha_m \geq 0$ is the combining coefficient which is used to describe the similarity of I_i and I_j on feature m. Therefore, $S(I_i, I_j)$ is used to evaluate the similarity of image I_i and image I_j on the selected features.

5. Re-sampling from the large pool of extra-person difference

Given a training set that includes N images for each of the K individuals, the total number of image pairs is $\binom{KN}{2}$. A small minority, $K\binom{N}{2}$, of these pairs are from the same individual. Any approach for learning the similarity function should explicitly handle the problem of how to choose limited samples from the overwhelmingly large number of negative samples to deal with the tremendous imbalance of the positive and the negative samples.

A simple proposal to solve this problem is to take a random subset of these pairs for training, but it can not ensure that the random subset could represent all the samples actually, so the re-sampling scheme we proposed is taken to guarantee that all possible samples can be referred during training. Figure 5 is the flowchart of the training procedure, in which S_i is a strong classifier boosted by weak classifiers which are learned from the current training set in the *ith* stage; T_i is the threshold till? the *ith* stage, which ensures to get the false positive and the detection rates that we need; and R_i is the re-sampling operation after the *ith* stage.

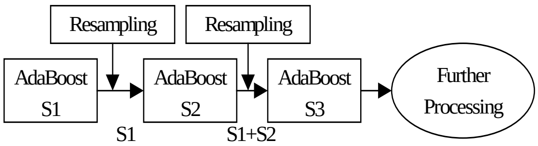


Fig.5. The flowchart of re-sampling procedure.

The ratio of positive samples to negative samples is imbalanced, since the number of negative samples is grossly larger than that of the positive samples. In the training set, the ratio of positive samples to negative samples is kept 1:7. How to re-sampling is a key of our system, it will be introduced in the following.

Because of the imbalanced rate of positive samples to negative samples, all positive samples are reserved in each stage and the negative samples are selected by re-sampling after each stage. Different from the face detection [11], each stage in our system has a false positive rate of about 0.01, which ensures that the weak classifiers learned in this stage are wholly capable of separating the positive samples from the negative samples. Although we can use the completely same steps as [11] to train a cascade of classifiers, the result of it is not as good as the strategy we take in following steps. And this will be further proved by the comparison experiments in section 6.

In [11], after training a stage, re-sampling is also used to select samples. If a negative sample x could pass all of the stages which have been trained, x is selected. In our strategy, x, a negative sample, does not need to pass all of the stages one by one; it just needs to pass the strong classifier *S*, if $S(x) \geq T_i$. So some negative samples trained in previous stages maybe reoccur in the latter stages.

Table 2. Training process with re-sampling scheme we proposed

- Given labeled examples Set, include all positive samples and select negative samples randomly at the rate of 1:7 from whole negative set.
- Do for *t*=1,…,T:
  1. AdaBoost
  2. $S = \sum_{i=1}^{t} S_i$,
  3. Select x randomly from negative set, if $S(x) \geq T_i$, add it to the new negative set for the next round, and S(x) is kept in next stage to get proper threshold T_{i+1}.
- Get a strong classifier $S = \sum_{i=1}^{T} S_i$

6. Experiment and Analysis

We tested the proposed method on the FERET face database, and the training set is also from the training set of FERET database, which includes 1002 images of 429 subjects. All images are cropped and rectified according to the manually located eye positions supplied with the FERET data. The normalized images are 45 pixels high by 36 pixels wide. The training set yields 795 intra-face image pairs and 500,706 extra-face image pairs. At any time, all 795 intra-face pairs and 5000 extra-face pairs are used for training. A new set of 5000 extra-face pairs is selected from the full training set by re-sampling scheme we proposed after one stage of AdaBoost has finished.

The number of Gabor features of each sample is $45 \times 36 \times 5 \times 8 = 64800$, from which the training algorithm would select hundreds of the most discriminant ones. We run AdaBoost in 7 stages, a total of 1108 rounds, and got 1108 features. The first four features learned by our algorithm are shown in figure 6, from which one can find that they are all as intuitively reasonable as the most discriminant Gabor features.

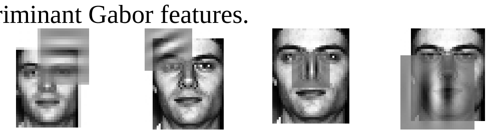


Fig.6. The first four Gabor features selected by the proposed method.

The experimental relationship between the rank-1 recognition rates and the number of weak classifiers is drawn in Figure 7, which is the result when testing the proposed method on the probe set FB and the gallery set FA of the FERET database. There are 1196 images in FA, 1195 images in FB, and all of the subjects have exactly one image in both FA and FB. As it can be seen from Figure.7, with the increase of the selected Gabor features, the rank-1 recognition rate improves from 37.5% with 6 features selected to 95.2% with 700 features selected. With more features exploited, the performance does not improve any longer. The result is comparable with the reported best result on this set in [6]. We also draw in Figure 8 the cumulative match score curve of the proposed method on FB probe set against FA gallery.

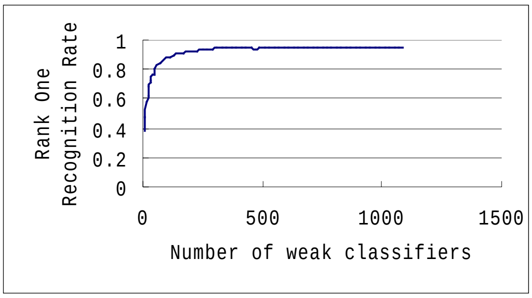


Fig.7. Face recognition performance of the proposed method with respect to the number of weak classifiers.

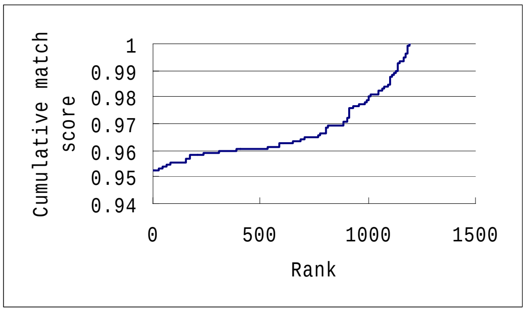


Fig.8. The Cumulative match score of the proposed method when testing on FERET FB probe set.

To prove the advantage of the re-sampling method, all 795 intra-face pairs and 5000 extra-face pairs randomly selected from 500,706 extra-face image pairs are used for training without re-sampling strategy. It means that we run just one stage of AdaBoost, a total of 2000 rounds, and got 2000 features. The same test experiment is done on FB and FA of the FERET database.

Figure 9 shows that the rank-1 recognition rate raises to 92.8% with 1741 features. The rank-1 recognition rate is just 90.6% with 700 features. Comparing the performance of re-sampling method and none re-sampling method, we can draw this conclusion that the re-sampling strategy we proposed is effective.

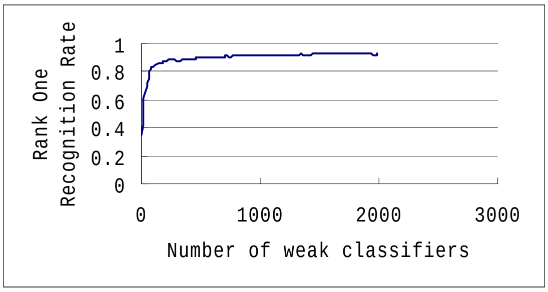


Fig.9. Face recognition performance of the method without re-sampling with respect to the number of weak classifiers

7. Conclusion

In the past few years, face representation based on Gabor features has attracted much attention and achieved great success in face recognition area for several advantages of the Gabor filters including their localizability, orientation selectivity, and spatial frequency characteristics. However, Gabor features currently adopted by most systems are too high dimensional to be used smoothly in a practical system. This paper proposes to tackle this problem by applying the AdaBoost learning approach. And a face recognition method using AdaBoosted Gabor features is proposed. AdaBoosted Gabor features are not only low dimensional but also discriminant. To apply the AdaBoost successfully to face recognition problem, we introduce the intra-face and extra-face difference space in the Gabor feature space to convert the multi-class face recognition problem into a two-class problem. In addition, to deal with the imbalance between the amount of the positive samples and that of the negative samples, a re-sampling scheme is adopted to choose the negative samples. By using the proposed method, only hundreds of Gabor features are selected for classification purpose. The experiments on FERET database have shown that these hundreds of Gabor features are enough to achieve good performance comparable to those methods using the complete set of Gabor features, which has impressively shown the effectiveness of the proposed method.

8. Acknowledge

This research is partially sponsored by NSFC under contract No. 60332010, National Hi-Tech Program of China (No.2001AA114190 and No. 2002AA118010) and the open project of multimedia and intelligent software Lab (No.KP0706200377). This work is also partially sponsored by ISVISION Technologies Co., Ltd and partially supported by NSFC under contract No. 60332010.

Reference

[1] W.Zhao, R.Chellappa and A. Rosenfeld, "Face Recognition: A Literature Survey," UMD CfAR Technical Report CAR-TR948, 2000.

[2] Turk M., and Pentland A., “Eigenfaces for Recognition”, J. Cognitive Neuroscience, vol.3, no. 1, pp. 71-86, 1991.

[3] P.N.Belhumeur, J.P.Hespanha etc. “Eigenfaces vs Fisherfaces: recognition using class specific linear projection”. IEEE trans. On PAMI, vol.20, no.7, pp711-720, July, 1997

[4] Bartlett, M. S., H. M. Lades, et al. “Independent component representations for face recognition”. In SPIE Symposium on Electronic Imaging: Science and Technology; Conference on Human Vision and Electronic Imaging III, San Jose, CA, 1998.

[5] Moghaddam, Pentland. "Beyond Eigenfaces: Probabilistic Matching for Face Recognition", IEEE International Conference on Automatic Face and Gesture Recognition (FG), pps 30-35, April 1998.

[6] PJ Phillips, H. Moon, P. Rauss, and SA Rizvi. “The FERET Evaluation Methodology for Face-Recognition Algorithms.” Proceedings of Computer Vision and Pattern Recognition, Puerto Rico, 137-143, 1997.

[7] L. Wiskott, J. Fellous, N. Kruger, C. von der Malsburg, "Face recognition by elastic bunch graph matching", IEEE Trans. PAMI, vol. 19, no. 7, pp. 775-669, 1997.

[8] IR Fasel, MS Bartlett, and JR Movellan. “A comparison of Gabor filter methods for automatic detection of facial landmarks.” In Proceedings of the 5th International Conference on Face and Gesture Recognition, 2002.

[9] Michael J. Jones and Paul Viola. “Face Recognition Using Boosted Local Features”. MERL Technical Reports. TR 2003-25, April 2003.

[10] Leo Breiman. Bagging predictors. Machine Learning, 24(2):123-140, 1996.

[11] P. Viola, M. Jones. “Robust Real-Time Object Detection.” In Proc. of IEEE Workshop on Statistical and Computational Theories in Computer Vision, 2001.

[12] Xiaoguang Lu and Anil K. Jain, "Resampling for Face Recognition", International Conference On Audio- And Video-Based Biometric Person Authentication (AVBPA'03), pp. 869-877, Guildford, UK, June, 2003

[13] C. Liu, and H. Wechsler, "Gabor Feature Classifier for Face Recognition", Proceedings of ICCV, Vol. 2, pp. 270-275, 2001.