A SVM Face Recognition Method Based on Optimized Gabor Features

Linlin Shen[1], Li Bai[2] and Zhen Ji[1]

[1] Faculty of Information & Engineering, Shenzhen University, China, 518060
{llshen, jizhen}@szu.edu.cn
[2] School of Computer Science & Information Technology, University of Nottingham, UK, NG8 1BB,
bai@cs.nott.ac.uk

Abstract. A novel Support Vector Machine (SVM) face recognition method using optimized Gabor features is presented in this paper. 200 Gabor features are first selected by a boosting algorithm, which are then combined with SVM to build a two-class based face recognition system. While computation and memory cost of the Gabor feature extraction process has been significantly reduced, our method has achieved the same accuracy as a Gabor feature and Linear Discriminant Analysis (LDA) based multi-class system.

Keywords: Gabor features, Support Vector Machine, Linear Discriminant Analysis

1 Introduction

Automatic recognition of human faces has been an active research area in recent years because it is user-friendly and unintrusive and it does not require elabrated collaboration of the users, unlike fingerprint or iris recognition. In addition to the importance of advancing research, it has a number of commercial and law-enforcement applications such as surveillance, security, telecommunications and human-computer intelligent interaction.

Gabor wavelet based recognition algorithms have been shown to be advantageous over many other methods in the literature. For example, the Elastic Bunch Graph Matching (EBGM) algorithm has shown very competitive performance and was ranked the top performer in the FERET evaluation [1]. In a recent face verification competition (FVC2004), both of the top two methods used Gabor wavelets for feature extraction. The application of Gabor wavelets for face recognition was pioneered by Lades et al.'s work since Dynamic Link Architecture (DLA) was proposed in 1993 [2]. In this system, faces are represented by a rectangular graph with local features extracted at the nodes using Gabor wavelets, referred to as Gabor jets. Wiskott et al [3] extended DLA to EBGM, where graph nodes are located at a number of facial landmarks. Since then, a large number of elastic graph based methods have been proposed [4-7]. Chung et al. [8] use the Gabor wavelet responses over a set of 12 fiducial points as input to a Principal Componet Analysis (PCA) algorithm, yielding a feature vector of 480 components. They claim to have improved the recognition rate up to 19% with this method compared to that by a raw PCA. All of these methods can be classified as analytic approaches since the local features extracted from selected points in faces are used for recognition. Recently, Gabor wavelets have also been applied in global form for face recognition [9, 10]. Liu et al. [9] vectorize the Gabor responses and then apply a downsampling by a factor of 64 to reduce the computation cost of the following subspace training. Their Gabor-based enhanced Fisher linear discriminant model outperforms Gabor PCA and Gabor fisherfaces. These holistic methods normally use the whole image after Gabor wavelets processing for feature representation. A more detailed survey on Gabor wavelet based face recognition methods can be found in [11].

Despite the success of Gabor wavelets based face recognition systems, the huge dimension of Gabor features extracted using a set of Gabor wavelets demands large computation and memory costs, which makes them impractical for real applications [11]. For the same reason, Support Vector Machine (SVM) has also seldom been applied to face recognition using Gabor features. Some works in the literature have tried to tackle this problem by (1) downsampling the images [12], (2) considering the Gabor responses over a reduced number of points [8], or (3) downsampling the convolution results [9, 10]. Strategies (2) and (3) have also been applied together [13]. However, these methods suffer from a loss of information because of the downsampling, or dimension reduction. Furthermore, the feature dimension after downsampling might still be too large for the fast training of SVM. Our works [14] have also shown that facial landmarks like eyes, nose and mouth might not be the optimal locations to extract Gabor features for face recognition.

In this paper, we propose a general SVM face recognition framework using optimized Gabor features. The most signifiant positions for extracting features for face recognition are first learned using a boosing algorithm, where the optimized Gabor responses are computed and used to train a two-class based SVM for identification. Since only the most important features are used, the two-class SVM based identification algorithm is both efficient and robust.

2 Gabor Wavelets and Feature Extraction

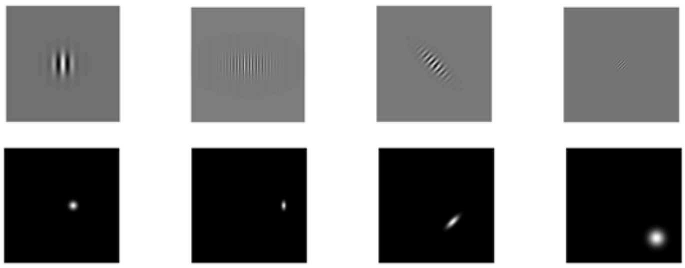
In the spatial domain, the 2D Gabor wavelet is a Gaussian kernel modulated by a sinusoidal plane wave [11]:

$$\begin{aligned} g(x,y) &= w(x,y)s(x,y) = e^{-(\alpha^2 x'^2 + \beta^2 y'^2)} e^{j2\pi f x'} \\ x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned} \tag{1}$$

where f is the central frequency of the sinusoidal plane wave, θ is the anti-clockwise rotation of the Gaussian and the plane wave, α is the sharpness of the Gaussian along the major axis parallel to the wave, and β is the sharpness of the Gaussian minor axis perpendicular to the wave. To keep the ratio between frequency and sharpness constant, $\gamma = \frac{f}{\alpha}$ and $\eta = \frac{f}{\beta}$ are defined and the Gabor wavelets can now be rewritten as:

$$\varphi(x, y) = \frac{f^2}{\pi\gamma\eta} g(x, y) = \frac{f^2}{\pi\gamma\eta} e^{-(\alpha^2 x'^2 + \beta^2 y'^2)} e^{j2\pi f x'} \tag{2}$$

Fig.1 shows four Gabor wavelets with different parameters in both spatial and frequency domain.



(a) (b) (c) (d)

Fig.1. Gabor filters with different parameters $\Pi(f, \theta, \gamma, \eta)$ in spatial domain (the 1st row) and frequency domain (the 2nd row), (a) $\Pi_a(0.1,0,1,1)$; (b) $\Pi_b(0.3,0,6,3)$; (c) $\Pi_c(0.2,3\pi/4,3,1)$; (d) $\Pi_d(0.4,3\pi/4,2,2)$

Once a set of Gabor wavelets have been designed, image features at different locations, frequencies and orientations can be extracted by convolving the image $I(x, y)$ with the filters:

$$O_{\Pi(f,\theta,\gamma,\eta)}(x, y) = I * \varphi_{\Pi(f,\theta,\gamma,\eta)}(x, y) \tag{3}$$

The number of scales and orientations may vary in different systems. We use in this paper a wavelet bank with 5 scales and 8 orientations to extract image features:

$$f_u = f_{\max} / \sqrt{2}^u, u = 0,...,4 \qquad \theta_v = \frac{v}{8}\pi, v = 0,...,7 \tag{4}$$

The results S are thus the convolutions of an input image $I(x, y)$ with all of the 40 wavelets:

$$S = \{O_{u,v}(x, y) | u \in \{0,...,4\}, v \in \{0,...,7\}\} \tag{5}$$

where $O_{u,v}(x, y) = \| I * \varphi_{\Pi(f_u, \theta_v)}(x, y) \|$.

When the convolution results $O_{u,v}(x, y)$ over each pixel of the image are concatenated to form an augmented feature vector, the size of the vector could be very large. Take an image of 24×24 for example, the convolution result will give 24×24×5×8=23,040 features. To make SVM applicable to such a large feature vector, Qin and He [13] reduced the size of feature vector by including only the convolution reults over 87 manually marked landmarks. However, locating the 87 landmarks itself is a difficult problem, and the manually selected positions might not be the optimal ones for face recognition. Furthermore, wavelets with the same parameters are used at different landmarks, which is not the optimal way to feature extraction. In this paper, a boosting based feature selection process is used to choose the most useful features, which are then given as input to SVM to learn an efficient and robust face identification system.

3 The OG-SVM Classifier

Ever since its invention, SVM has been widely applied in classification and pattern recognition. One of the main reasons for the widespread applications of SVM is that its decision function is only based on the dot product of the input feature vector with the Support Vectors (SVs) [15], i.e. it has no requirements on the dimension of the feature vector. Theoretically features with any dimension can be fed into SVM for training. However in practical implementation, features with large dimension, e.g. Gabor features, could bring substantial computation and memory cost to the SVM training and classification process. In our experiments, the SVM training process did not even complete after 74 hours when a set of Gabor features of dimension 23,040 was used, due to the large computation and memory costs.

To make the SVM classifier both efficient and accurate, we propose to use optimized Gabor features for classification. As shown in Fig.2, the system starts with the Gabor feature extraction, as described in section 2. The extracted Gabor features and associated class labels for all of the training samples are then fed into the boosting algorithm to eliminate those non-discriminative features, which are not significant for classification. Once the most important positions with tuned Gabor wavelets are identified, the optimized Gabor features can be extracted and used to train the classifier, namely, the OG-SVM classifier. Using the optimized features, the boosting algorithm also learned a reasonably good classifier - Boosted Classifier (BC). However, the nonlinear OG-SVM classifier achieved further improvement on classification accuracy, with similar efficiency.

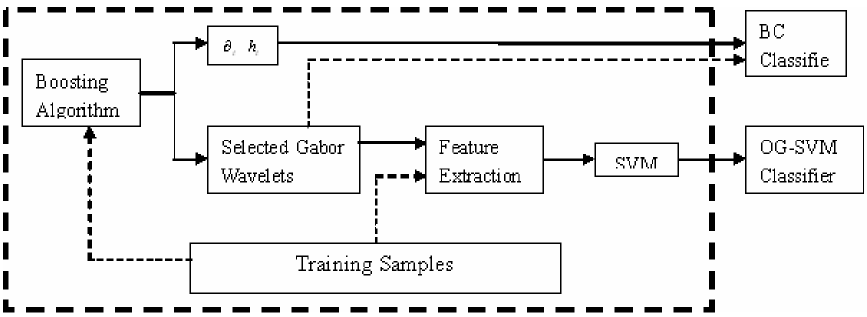


Fig.2. Learning process of the proposed OG-SVM classifier

3.1 Boosting Based Gabor Feature Selection

Introduced by Freud and Schapire [16], boosting algorithms have been used successfully for selecting Haar-like features for general object detection [17, 18]. The essence of boosting algorithms is to select a number of 'weak' classifiers, which are then linearly combined into a single strong classifier. The algorithm operates as follows: for a two-class problem, m labelled training samples are given as $(x_i, y_i), i = 1,2,..,m$, where $y_i \in \{-1,1\}$ is the class label associated with sample $x_i \in R^N$. A large number of weak classifiers $h: R^N \rightarrow \{-1,1\}$ can be generated to form a weak classifier pool for training. In each of the iterations, the space of all possible weak classifiers is searched exhaustively to find the one that contributed the least to the overall classification error. The error is then used to update the weights associated with each sample such that the wrongly classified samples have their weights increased. The algorithm thus focuses on difficult training samples, increasing their representation in successive training sets. When a weak classifier is designed to use only a single feature to make decisions, boosting is equivalent to feature selection.

To apply the boosting algorithm to Gabor feature selection, we simplify the task of a multi-class face recognition problem to a two-class problem: selecting Gabor features that are effective for intra- and extra-person space discrimination. Such selected Gabor features should be robust for face recognition, as intra- and extra-person space discrimination is one of the major difficulties in face recognition. Two spaces, intra- and extra-person spaces are defined, with intra-person space measuring respectively dissimilarities between faces of the same person and extra-person space dissimilarities between different people. For a training set with L facial images captured for each of the K persons to be identified, $K\binom{L}{2}$ samples could be generated for class $Intra$ while $\binom{KL}{2} - K\binom{L}{2}$ samples are available for class $Extra$. More details about the Gabor feature selection process can be found at [14].

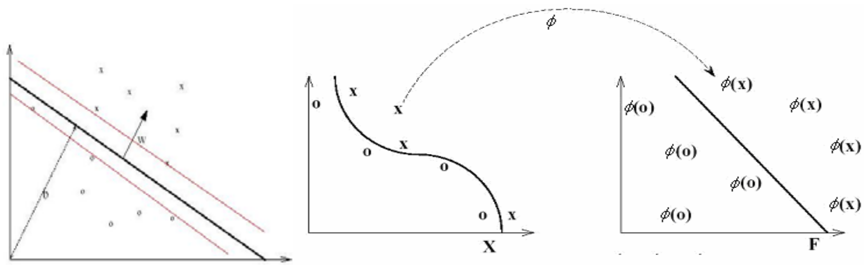
Upon completion of T boosting iterations, T weak classifiers are selected to form the final strong classifier. The resulting strong classifier $H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$, called BC in this paper, is a weighted linear combination of all the selected weak classifiers, with each weak classifier using certain Gabor feature for decision. At the same time, T most significant Gabor features for face recognition have also been identified.

3.2 Support Vector Machine

Once the optimized features are selected, they can be given to SVM for classifier training. Based on an observed feature $x \in R^N$, SVM is basically a linear hyperplane classifier $f(x) = \langle w, x \rangle + b$ aimed at solving the two class problem [19]. As shown in Fig.3a, the classifier can separate the data from two classes very well when the data is linearly separable. Since there might be a number of such linear classifiers available, SVM chooses the one with the maximal margin, which is defined as the width that the boundary could be increased by before hitting a data point. The distance between the two thin lines (boundary) in the figure thus defines the margin of the linear SVM with data points on the boundary known as Support Vectors (SV). The linear classifier $f(x)$ with maximized margin can be found using quadratic problem (QP) optimization techniques as below:

$$f(x) = sign\left(\sum \alpha_k y_k \langle x_k, x \rangle + b\right) \quad (6)$$

where $x_k \in R^N$ are the support vectors learned by SVM.



(a) (b)

Fig.3. A hyperplane classifier in 2-dimension feature space (a and mapping of the data (b

For non-linearly separable data, a nonlinear mapping function $\phi : R^N \to F, x \to \phi(x)$ is used to map it into a higher dimension feature space where a linear classifier can be applied. Fig.3b shows an example using the kernel method to train a non-linear SVM. Using the kernel trick [15], the non-linear SVM is now:

$$f(x) = sign\left(\sum \alpha_k y_k k(x_k, x) + b\right) \quad (7)$$

where $k(\mathbf{x}_k, \mathbf{x})$ is a kernel function, e.g., a polynomial kernel and a RBF kernel etc.

3.3 Identification

As shown in Fig.2., once the boosting iterations and the SVM learning process are completed, two classifiers, i.e. BC and OG-SVM, are created using the *T* selected Gabor features. Though trained to discriminate intra-person and extra-person spaces, they could also be used for recognition (identification) as follows: given a gallery $\{q_j\}$ of m known individuals and a probe p to be identified, both classifiers will first compute the Gabor feature differences $\{x_j = [d_1 \cdots d_t \cdots d_T]\}$ between the probe and each of the gallery images, and then calculate an intra-person confidence score using respective decision functions:

$$\delta_j = \begin{cases} \sum_{t=1}^{T} \alpha_t h_t(x_j), & \text{BC} \\ \sum_k a_k y_k k(x_k, x_j) + b, & \text{SVM} \end{cases} \quad (8)$$

the probe is then identified as person j that gives the maximum confidence score δ_j.

4 Experimental Results

4.1 The Database

The FERET database is used to evaluate the performance of the proposed method for face recognition. The database consists of 14051 eight-bit grayscale images of human heads with views ranging from frontal to left and right profiles. 600 frontal face images corresponding to 200 subjects are extracted from the database for the experiments - each subject has three images of 256×384 with 256 gray levels. The images were captured at different times under different illumination conditions and contain various facial expressions. Two images of each subject are randomly chosen for training, and the remaining one is used for testing. The following procedures were applied to normalize the face images prior to the experiments:

- each image is rotated and scaled to align the centers of the eyes,
- each face image is cropped to the size of 64×64 to extract facial region,
- each cropped face image is normalized to zero mean and unit variance.

4.2 The Results

In this experiment, classification and recognition performance of the proposed two-class classifier, OG-SVM, will be tested and evaluated against that of BC and other methods, e.g. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Gabor features are first selected by boosting algorithm using the training set, and then used to train BC and OG-SVM (see Fig.2. for the process). The training set thus consists of 200 intra-person difference samples and 1,600 extra-person difference samples.

Since both BC and OG-SVM are trained to discriminate intra-person and extra-person differences, we first evaluate their classification performances on the training set. Fig.4 shows the classification error of BC and OG-SVM with different kernel functions, which are computed as the ratio between the number of wrongly classified difference samples and the number of training samples. One can observe from the figure that the performances of both classifiers improve when the number of features increases. However, the performance of OG-SVM is much more stable than BC.

While OG-SVM with RBF kernel achieves the lowest classification error rate (0.44%) when 140 features are used, OG-SVM with linear kernel shows similar performance.

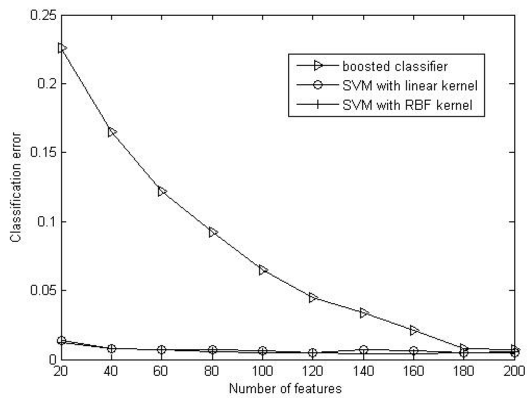


Fig.4. Classification performances of OG-SVM and BC

The classifiers are then applied to the test set (200 images, 1 image per person) for face identification and their performances are shown in Fig.5. Similarly, OG-SVM achieves higher recognition rate than BC when different number of features are used. The highest recognition accuracy of 92% is achieved by OG-SVM with linear kernel when 120 Gabor features are used. The results also suggest that the difference of OG-SVM using RBF kernel and linear kernel is quite small, when the features selected by boosting algorithm are considered.

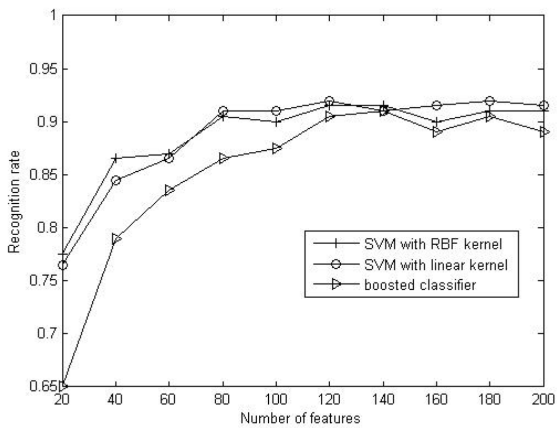


Fig.5. Recognition performances of OG-SVM and BC

To show the efficiency and accuracy of the proposed method, we also compare its performance with other Gabor feature based approaches in Table 1. While PCA and LDA are also well known as Eigenface and Fisherface methods, details of Downsmaple Gabor + PCA and Downsample Gabor + LDA can be found in [10]. In the implementation, downsampling with rate 16 was used to reduce the dimension of extracted Gabor features before they are input to PCA, or LDA for further processing. The table shows that the proposed OG-SVM achieved similar accuracy with Downsmaple Gabor + LDA, but with much fewer feature dimension and much less feature extraction costs. In our experiments (a normal PC with P4 3.0 GHz CPU), while it takes 100ms to train the OG-SVM classifier, the system can averagely identify 50 faces per second.

Table 1. Accuracy and efficiency of OG-SVM

Methods	Recognition Rate	No. of Convolutions for Gabor Feature Extraction	Dimension of Features
PCA	60%	N/A	$64\times64=4096$
LDA	76%	N/A	$64\times64=4096$
Downsample Gabor + PCA	80%	$64\times64\times40=163,840$	10,240
Downsample Gabor + LDA	**92%**	$64\times64\times40=163,840$	10,240
BC	90%	120	120
OG-SVM	**92%**	120	120

5 Conclusions

We have proposed in this paper a novel SVM face recognition method based on optimized Gabor features. While some methods in the literature consider the responses at landmark points only, our method uses a boosting algorithm to find the most significant positions and wavelet to extract features for face recognition. The features thus extracted are efficient. While downsampling could be used to reduce the dimension of features before they are fed into PCA, or LDA for further processing, it could introduce loss of important information. Furthermore, complex feature extraction process has to be used to extract high dimensional features before downsampling. By combining boosting selected Gabor features with SVM, our method not only substantially reduces computation and memory cost of the feature extraction process, but also achieves the same performance as that of Downsample Gabor + LDA, when FERET database is used for testing.

Acknowledgement. Research funded by SZU R/D Fund 200746.

References

1. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J.: The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence **22,** 1090-1104, (2000)
2. Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., Von der Malsburg, C., Wurtz, R.P., Konen, W.: Distortion invariant object recognition in the Dynamic Link Architecture. IEEE Transactions on Computers **42,** 300-311, (1993)

3. Wiskott, L., Fellous, J.M., Kruger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence **19**, 775-779, (1997)
4. Mu, X.Y., Hassoun, M.H.: Combining Gabor features: summing vs.voting in human face recognition. 2003 IEEE International Conference on Systems, Man and Cybernetics (2003) 737
5. Duc, B., Fischer, S., Bigun, J.: Face authentication with Gabor information on deformable graphs. IEEE Transactions on Image Processing **8,** 504-516, (1999)
6. Jiao, F., Gao, W., Chen, X., Cui, G., Shan, S.: A face recognition method based on local feature analysis. Proc. of the 5th Asian Conference on Computer Vision, 188-192, (2002)
7. Liao, R., Li, S.: Face recognition based on multiple facial features. Proc. of the 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 239-244, (2000)
8. Chung, K.C., Kee, S.C., Kim, S.R.: Face recognition using principal component analysis of Gabor filter responses. International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Corfu, Greece, 53-57, (1999)
9. Liu, C.J., Wechsler, H.: Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition. IEEE Transactions on Image Processing **11,** 467-476, (2002)
10. Shen, L., Bai, L., Fairhurst, M.: Gabor wavelets and General Discriminant Analysis for face identification and verification. Image and Vision Computing **25,** 553-563, (2007)
11. Shen, L., Bai, L.: A review on Gabor wavelets for face recognition. Pattern Analysis and Applications **9,** 273-292, (2006)
12. Zhang, W., Shan, S., Gao, W., Chang, Y.Z., Cao, B., Yang, P.: Information fusion in face identification. Proceedings of the 17th International Conference on Pattern Recognition, Vol. 3, 950-953, (2004)
13. Qin, J., He, Z.-S.: A SVM face recognition method based on Gabor-featured key points. Proceedings of the 4th International Conference on Machine Learning and Cybernetics, Vol. 8, 5144-5149, (2005)
14. Shen, L., Bai, L.: MutualBoost learning for selecting Gabor features for face recognition. Pattern Recognition Letters **27,** 1758-1767, (2006)
15. Cristianini, N., Shawe-Taylor, J.: An introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press (2000)
16. Freund, Y., Schapire, R.: A short introduction to boosting. Journal of Japanese Society for Artifical Intelligence **14,** 771-780, (1999)
17. Lienhart, R., Maydt, J.: An extended set of Haar-like features for rapid object detection. Proc. IEEE Conference on Image Processing, 900-903, (2002)
18. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, 511-518, (2001)
19. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery **2,** 121-167, (1998)