

FACE RECOGNITION BASED ON GRADIENT GABOR FEATURE

Baochang Zhang^{1,2}, Yongsheng Gao², Yu Qiao³

¹ School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, China

² Institute for Integrated and Intelligent System, Griffith University, Australia

³ Graduate School of Frontier Sciences, the University of Tokyo, Japan

Email: b.zhang@griffith.edu.au; yongsheng.gao@griffith.edu.au; qiao@gavo.t.u-tokyo.ac.jp

ABSTRACT

In this paper, a novel Gradient Gabor (GGabor) filter is proposed to extract multi-scale and multi-orientation features to represent and classify faces. Gradient Gabor combines the derivative of Gaussian functions and the harmonic functions to capture the features in both spatial and frequency domains to deliver orientation and scale information. The spatial positions are combined into Gaussian derivatives which allows it to provide more stable information. An Efficient Kernel Fisher analysis method is proposed to find multiple subspaces based on both GGabor magnitude and phase features, which is a local kernel mapping method to capture the structure information in faces. Experiments on two face databases, FRGC Version 1 and FRGC Version 2, are conducted to compare the performances of the Gabor and GGabor features, which show that GGabor can also be a powerful tool to model faces, and the Efficient Kernel Fisher classifier can improve the efficiency of the original kernel fisher method.

1. INTRODUCTION

A good object representation or pattern representation is one of the key issues for a well-designed pattern recognition system [1, 2]. Representation issues include: what representation is desirable for the recognition of a pattern and how to effectively extract the representation from the original input signal.

To achieve the goal of extracting features in a certain scale, Laplacian of Gaussian (LoG) were introduced [3] to simulate the lateral inhibition for edge detection. In order to extract features in certain orientations, oriented filters or steerable filters are introduced by Freeman. However, to be a powerful descriptor, the feature extractor should be anisotropic, which means that it should enhance the feature in a certain scale and orientation simultaneously. Gabor [4] transformation has been widely used as an effective tool in the image processing and pattern recognition tasks. Different from steerable filters [5, 6], Gabor can also include information both in spatial and frequency domains. It can characterize the spatial frequency structure in the image while preserving information of spatial relations, and thus is suitable for extracting the orientation-dependent frequency contents of patterns.

Gabor wavelet is a sinusoidal plane wave with a

particular frequency and orientation, modulated by a Gaussian envelope. Gabor inherits the property of Gaussian, which is a powerful smoothing tool for the image. Moreover, the derivative of Gaussian is regarded as an important tool to extract the edge information for images, which is widely used in image processing and computer vision. In face community, the edge distribution information has been successfully investigated [7]. In order to inherit the property of the derivative of Gaussian, Gradient Gabor (GGabor) is proposed based on the combination of Fourier Transform and the derivative of Gaussian. Gradient Gabor can be considered as a weighted Gabor wavelet, and spatial positions are combined into Gaussian window. After Gradient Gabor feature is extracted, Efficient Kernel Fisher (EKF) discriminant analysis is then proposed to find discriminant subspaces. In this paper, two face databases, FRGC version 1 and FRGC version 2 [8], are used to compare the performances of Gabor and GGabor features, which show that GGabor feature can also be powerful tool to model face objects.

The rest of this paper is organized as follows. Section 2 introduces the Gradient Gabor filter. The recognition method based on efficient kernel Discriminant analysis is described in Section 3. The experiments are given in Section 4. We conclude the paper in Section 5.

2. MOTIVATION OF GRADIENT GABOR

In this section, we will first describe Gabor wavelet, and then define the Gradient Gabor filters.

2.1. Gabor wavelet

The Gabor wavelets (kernels, filters) can be defined as follows [9, 10]:

$$\Psi_{u,v}(z) = (\|k_{u,v}\|^2 / \sigma^2) e^{-(\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2)} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2} \right] \quad (1)$$

where $\vec{k}_{u,v} = \begin{pmatrix} k_v \cos \Phi_u \\ k_v \sin \Phi_u \end{pmatrix}$, $k_v = f_{\max} / 2^{v+2}$, $\Phi_u = u \frac{\pi}{8}$, v is the scales and u is the orientation with $f_{\max} = \pi/2$. In this paper, 4 scales and 4 orientations are used. Gabor wavelet can enhance the features in certain scales and orientations, which is widely used in image processing and object recognition. The Gabor transformation of a given image is defined as its convolution with the Gabor functions:

$$G_{u,v}(z) = I(z) * \Psi_{u,v}(z), \quad (2)$$

where $z = (x, y)$ denotes the image position, the symbol ‘*’ denotes the convolution operator, and $G_{u,v}(z)$ is the convolution result corresponding to the Gabor kernel at scale v and orientation u . The Gabor wavelet coefficient $G_{u,v}(z)$ is a complex, which can be rewritten as:

$$G_{u,v}(z) = A_{u,v}(z) \cdot \exp(i\theta_{u,v}(z)), \quad (3)$$

with one magnitude item $A_{u,v}(z)$, and one phase item $\theta_{u,v}(z)$. It is well known that the magnitude varies slowly with the spatial position, while the phases rotate in some rate with positions, even it can preserve more detailed information. This can cause severe problems for object (face) matching, and it is just the reason that most previous works make use of only the magnitude for face classification. In the following part, we will introduce a new Gradient Gabor filter, which can provide a relatively stable representation of faces.

2.2. Gradient Gabor filters

The Gradient Gabor filters are defined based on the derivative of Gaussian function:

$$\text{G}\Psi_{u,v}(z) = -\frac{\|k_{u,v}\|^4}{\sigma^4} ((x+y)e^{ik_{u,v}z} + C)e^{(-\|k_{u,v}\|^2\|z\|^2/2\sigma^2)}, \quad (4)$$

$$C = -\left(\frac{1}{i\sigma^2} k_{u,v} \exp\left(-\frac{\pi}{\sigma^2} (k_{u,v})^2\right)\right), \quad (5)$$

where C makes Gradient Gabor DC-free. Gabor wavelet is modulated by a Gaussian function, which can be regarded as the weighted Fourier transform. The weights for Gabor wavelets are actually exponentially declining with increasing distance. However, Gradient Gabor is defined based on a weighted Gaussian function, which is not declining in an exponential speed as in Gabor wavelets, because it is slowed down by a linear function as shown in Eq.4 and Fig.1. Different from original Gabor filters, it can be more stable and provide a robust presentation of the face object by using the multi-scale and multi-orientation local features. Samples of 2D Gabor and GGabor filters are shown in Fig.2. In our face recognition system, both magnitude and phase information are combined to enhance its performance.

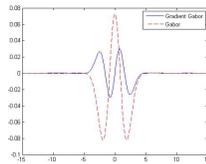


Fig.1 Real parts of 1-D Gradient Gabor and Gabor Filters

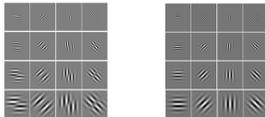


Fig.2 Real parts of 2-D Gradient Gabor(left) and Gabor Filters(right) with 4 scales and 4 orientations

3. ENSEMBLE-BASED EFFICIENT KERNEL FISHER CLASSIFIER

In this part, we will use the ensemble-based kernel fisher discriminant analysis method to find discriminant subspace. As shown in Fig.3, we will make kernel discriminant analysis for each sub-region based Gradient Gabor or Gabor feature.



Fig.3 Face object divided into 48 sub-regions

3.1. Kernel Fisher Analysis

Kernel Fisher discriminant analysis method is used to find a discriminant transformation space, the input data is first projected into an implicit feature space F by the nonlinear mapping $\Phi: x \in R^N \rightarrow f \in F$. In its implementation, Φ is implicit and we will just compute the inner product of two vectors in F by using a kernel function [11, 13]:

$$k(x, y) = (\Phi(x) \cdot \Phi(y)). \quad (6)$$

The between-class scatter matrix S_b and within-class scatter matrix S_w are defined as follows:

$$S_b = \sum_{i=1}^C p(\varpi_i) (u_i - u)(u_i - u)^T, \quad (7)$$

$$S_w = \sum_{i=1}^C p(\varpi_i) E\{((\Phi(x_i) - u_i)(\Phi(x_i) - u_i)^T) | \varpi_i\}, \quad (8)$$

where $u_i = (1/n) \sum_{j=1}^{n_i} \phi(x_{ij})$ denotes the sample mean of class i , u is the mean of all training images, and $p(\varpi_i)$ is the prior probability.

In the original Kernel Fisher method, $w \in F$ should lie in the span of all the samples in F , we define the basis Support Vectors (SVs):

$$SVs = (\phi(x_1), \phi(x_2), \dots, \phi(x_N)), \quad (9)$$

$$w = \alpha SVs^T, \quad (10)$$

where N is the total number of the training samples, and the kernel matrices are defined as follows:

$$K_w = \sum_{i=1}^C p(\varpi_i) E(\eta_j - m_i)(\eta_j - m_i)^T, \quad (11)$$

$$K_b = \sum_{i=1}^C p(\varpi_i) (m_i - \bar{m})(m_i - \bar{m})^T, \quad (12)$$

where $\eta_j = (k(x_1, x_j), k(x_2, x_j), \dots, k(x_n, x_j))^T$,

$m_i = \left(\frac{1}{n_i} \sum_{j=1}^{n_i} k(x_1, x_j), \frac{1}{n_i} \sum_{j=1}^{n_i} k(x_2, x_j), \dots, \frac{1}{n_i} \sum_{j=1}^{n_i} k(x_n, x_j)\right)^T$, and \bar{m} is the mean of all η_j .

In [11, 13], we can find that the definition of w based on all training samples is the curse of the Kernel Fisher

method, which has to save the whole training database. In the following part, we will propose a new Efficient Kernel Fisher (EKF) scheme to solve this problem.

3.2. Efficient Kernel Fisher Analysis For Ensemble-Based Face Recognition

In this part, we try to redefine \mathbf{w} by using the local region features and the clustering centers of the training database, and the new basis Support Vectors (SVs') are:

$$SVs' = (\phi(X_1^1), \phi(X_2^1), X_i^j, \phi(X_L^{C_m})), \quad (13)$$

$$\mathbf{w}' = \alpha' SVs'^T, \quad (14)$$

where c_m is the numbers of clustering centers, and X is the clustering center calculated from the training database by K-means method with $C_m \ll N$. X_i^j is the Gabor or Gradient Gabor feature extracted from the local region $R_{i,i=0,1,\dots,L-1}$ of the j_{th} face image. For Eq.13, we can know that \mathbf{w}' is defined based on both training samples and local region features. In this paper, it should be reasonable for the ensemble-based method, by using local kernel method based on Eq.13, each sub-classifier can preserve the information about the relationship among local features across the training database.

In the classification procedure, v^1, v^2 are the discriminant features vectors corresponding to two face images P_1, P_2 , the cosine similarity rule is used:

$$d(P_1, P_2) = \sum_{i=1}^L \frac{v_i^1 \cdot v_i^2}{\|v_i^1\| \cdot \|v_i^2\|}. \quad (15)$$

From Eq.15, we can easily know that the proposed method is based sum rule, which can actually use the spatial structure information of the face image, therefore, it should be appropriate to face recognition.

4. EXPERIMENTS

To validate the proposed methods, experiment #4 on FRGC Version 1 and Version 2 databases are conducted. For experiment #4 of FRGC Version 1, the training set contains 366 images, the target set (Gallery) contains 943 controlled images, and the query set (Probe) has 943 uncontrolled images. Accordingly, for experiment #4 of FRGC Version 2, the training set contains 12,776 images, the target set contains 16,028 controlled images, and the query set has 8,014 uncontrolled images. In our paper, the polynomial kernel function, $k(x, y) = (x \cdot y)^2$, are used to test the performance of the proposed method.

As shown in [8], the Experiment #4 is designed for indoor controlled still images versus uncontrolled still images, which is the most challenging FRGC experiment. In both experiments, face images are cropped to 64*72 normalization images, which are divided into 8*12-sized sub-images.

4.1. Comparisons based on FRGC Version 1

In this experiment, the mean sample for each class is calculated in the target set. The first experiment is conducted and compared between the original Kernel Fisher and Efficient Kernel Fisher method based on Gabor and Gradient magnitude features. P_GGabor and P_Gabor are based on the phase information, M_GGabor and M_Gabor use the magnitude information, and A_GGabor and A_Gabor use both magnitude and phase information. A_Gabor(K) or A_GGabor(K) means that the methods are based on the original Kernel fisher method, and A_Gabor(E) and A_GGabor(E) are based on the Efficient Kernel Fisher method. Ensemble-GFC is based on the Fisher classifier, details about the method can refer to [12]. The numbers of the clustering centers is 40. From Table.1, we can know that A_Gabor(E) and A_GGabor(E) have achieved a little better performance than the original Kernel fisher method, however, it just needs to save 40 samples, and the old kernel fisher method should use all 366 samples, which makes efficient kernel method more suitable to the real applications. From Fig.5, we can find that the phase part of GGabor achieved a better performance than that of Gabor wavelet in terms of Rank-N recognition rates, which confirms that Gradient Gabor can provide more stable information.

Table.1 the comparative experiments between original Kernel Fisher method and Efficient Kernel Fisher method (Rank-1)

methods	Recognition Rates	Model Size*
A_Gabor(K)	92.9%	366
A_Gabor(E)	93.2%	40
A_GGabor(K)	94.1%	366
A_GGabor(E)	94.3%	40
Ensemble-GFC	91.5%	

*? means the size of the Basis Support Vectors

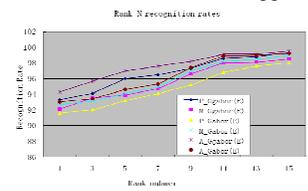


Fig.4. Rank-N (CMC) recognition rates curves based on the EKF method using 40 clustering centers.

4.2. Comparisons based on FRGC Version 2

In this section, another bigger database is used to evaluate the performance of the face recognition system. As for the high complexity of the original kernel fisher method for a large training database, we choose a subset of 5000 as basis support vectors to train kernel fisher discriminant subspaces using the bagging based method as in [11], which are A_Gabor(K) and A_GGabor(K). For the EKF method, 5 sub-classifiers were trained for each 100 clustering centers resulting a total of 500 model size, which

greatly decreased the model size with an acceptable performance as shown in Table 3. Therefore, the EKF method is a more promising way for real applications.

Table 2. The comparative experiments between bagging-based Kernel Fisher and the EKF method.

methods	Recognition Rates	Model Size*
A_Gabor(K)	91.4%	5,000
A_Gabor(E)	90.1%	500
A_GGabor(K)	93.1%	5,000
A_GGabor(E)	92.3%	500

5. SUMMARY AND FUTURE WORK

This paper proposes a new face recognition method based Gradient Gabor feature, in which both magnitude and phase parts are used to represent the face object. The main contributions of the proposed method are: (1) The Gradient Gabor feature is completely new for face recognition; (2) Different from the Gabor phase, the GGabor phase can offer relatively stable information for face recognition task; (3) We have validated the proposed method by conducting the experiments on two face databases, FRGC Version 1 and FRGC Version 2. On both databases, we have achieved better results than Gabor-based results.

Although the performance is successfully verified based on two open databases, the effectiveness of the Gradient Gabor feature can be combined with other methods, such as, Local Binary Pattern method. Due to its excellent performance, we expect that the proposed method is applicable to other object recognition tasks as well.

6. REFERENCES

- [1] A. Leonardis, H. Bischof, Robust recognition using eigenimages, *Computer Vision and Image Understanding*, vol. 78, no.1, pp. 99-118, 2000.
- [2] T.B. Moeslund, E. Granum, A survey of computer vision-based human motion capture, *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231-268, 2001.
- [3] D. Marr, and E. Hildreth, Theory of Edge Detection, in *Proc. of the Royal Society of London, Series B*, vol. 207, pp. 187-217, 1980.
- [4] D. Gabor, Theory of communication, *Journal of the Institution of Electrical Engineers*, vol. 93, part III, no. 26, pp. 429--457, 1946.
- [5] X. Zhang and Y. Jia, Local Steerable Phase(LSP) feature for face representation and recognition, in *proc. of ICCV*, pp.1363-1368, 2006.
- [6] W.T. Freeman, E. H. Adelson, The Design and use of Steerable Filters, *IEEE TPAMI*, pp.891-906, 1991.
- [7] B. Zhang, S. Shan, X. Chen, W. Gao. Histogram of Gabor Phase Patterns: A novel object representation for face recognition, *IEEE TIP*, 2007,1.
- [8] P.J. Philips, P.J. Flynn, T. Scruggs, K.W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaescik Min, Willimam Worek, Overview of the face recognition grand challenges, in *proc. of ICCV*, pp.947-954, 2005.
- [9] C. Liu and H. Wechsler, Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, *IEEE TIP*, vol. 11, no. 4, pp. 467-476, 2002.
- [10] L. Wiskott, J. M. Fellous, N. Kruger, C. von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, *IEEE TPAMI*,19, pp. 775-780, 1997.
- [11] Y. Li, B. Zhang, S. Shan, X. Chen, W. Gao, Bagging based efficient kernel fisher discriminant analysis for face recognition, *In Proc. of ICPR*, pp.523-526, 2006.
- [12] Y. Su, S. Shan, X. Chen, W. Gao, Hierarchical Ensemble of Gabor Fisher Classifier for Face Recognition. in *proc. of International Conference on Automatic Face and Gesture Recognition*, pp. 91-96, 2006.
- [13] C. Liu, Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance, *IEEE TPAMI*, pp.725-737, 2006.