

SIGNIFICANT JET POINT FOR FACIAL IMAGE REPRESENTATION AND RECOGNITION

Sanqiang Zhao and Yongsheng Gao

School of Engineering, Griffith University, Nathan Campus, QLD 4111, Australia
{s.zhao, yongsheng.gao}@griffith.edu.au

ABSTRACT

Gabor wavelet related feature extraction and classification is an important topic in image analysis and pattern recognition. Gabor features can be used either holistically or analytically. While holistic approaches involve significant computational complexity, existing analytic approaches require explicit correspondence of predefined feature points for classification. Different from these approaches, this paper presents a new analytic Gabor method for face recognition. The proposed method attaches Gabor features on a set of shape-driven sparse points to describe both geometric and textural information. Neither the number nor the correspondence of these points is needed. A variant of Hausdorff distance is employed to recognize faces. The experiments performed on AR database demonstrated that the proposed algorithm is effective to identify individuals in various circumstances, such as under expression and illumination changes.

Index Terms—Image representation, face recognition, Significant Jet Point, Hausdorff distance

1. INTRODUCTION

One popular class of techniques used for face recognition is based on Gabor features. These techniques can be roughly classified into holistic approaches and analytic approaches. Holistic approaches employ the whole image after Gabor transformation for feature representation. Liu and Wechsler [1] presented an Enhanced Fisher Linear Discriminant Model (EFM) to the Gabor feature vector for face recognition, and the so-called Gabor-Fisher Classifier (GFC) outperformed both PCA and LDA. Since the 40 Gabor filtered images are concatenated together to form a feature vector, the dimension is huge. This work is later extended to use Independent Component Analysis (ICA) [2] and Kernel PCA [3] to reduce the dimension of the extracted Gabor feature vector and enhance the discriminative power at the same time.

Analytic approaches originated with Dynamic Link Architecture (DLA) [4], which was later extended to Elastic Bunch Graph Matching (EBGM) [5]. In this deformable graph matching method, a topology graph is constructed for

each face, with each node attached one or several Gabor jets. Each component of a jet is a filter response of a specific Gabor wavelet extracted at a predefined feature point. These locally estimated Gabor features are known robust against illumination change, distortion and scaling [6]. In the graph matching process, each node of the template graph is allowed to vary its scale and position according to the appearance variations on the face. Though EBGM showed remarkably competitive performance in the FERET evaluation test [7], the computation involved is rather intensive.

Both DLA and EBGM methods are based on a basic principle that Gabor features should be extracted from predefined fiducial points in a face image. Each fiducial point has clear and exact meaning and refers to the same position across different face images. Face recognition is performed through comparing Gabor features of the same fiducial points over different individuals. Most computational work of EBGM lies in the optimization of locating accurate fiducial points. This paper presents a new approach to circumvent building correspondence between fiducial points. Different from EBGM and DLA, the proposed approach extracts Gabor features from a set of significant points, which are shape-driven and detected from edge map with rich information content on the face image. Neither the number nor the location of significant points is confined, and so diverse facial characteristics of different persons can be represented. A significant point attached with its Gabor jet is called a Significant Jet Point (SJP), which contains useful information and discriminative power for face recognition. Different from most existing analytic Gabor methods that only compare Gabor features from equal numbers of corresponding points, the proposed Jet Point Distance (JPD) is proved to be an effective measurement on different numbers of SJPs. Moreover, unlike neural network methods which inherently suffer from the generalizability problem, the proposed method is suitable for single model face database recognition. The investigation and evaluation of the proposed method is conducted in public available AR face database [8], covering all conditions of human face recognition. The proposed method obtains a significantly higher recognition accuracy than that of Directional Corner Point (DCP) [9] in our experiments, demonstrating the validity of this method

on face representation and recognition. However, it is important to stress that this method is not only restricted to face recognition, but capable of dealing with any other object recognition problems as well.

The remaining part of this paper is organized as follows. Section 2 presents the proposed significant point detection and SJP representation. Section 3 describes the JPD in detail. In Section 4, comparative experiments on AR database are conducted to evaluate the performance of the proposed method. The last section concludes the paper.

2. SIGNIFICANT JET POINT

2.1. Significant Point Detection

Containing geometric information, significant points are detected at those positions which have rich edge information in a face image. In most feature-based methods, feature points are predefined as the locations of eyes, nose, mouth, etc. However, neither the number nor the location of significant points is fixed here. The number of significant points and their locations can vary in order to better represent diverse facial characteristics of different individuals, such as dimples, moles, etc. These diverse features are also important cues that people might use for recognizing faces.

In order to ensure the less demand on storage space and less sensitivity to illumination changes, the significant points should be placed on the significant edge curves with high curvatures. While any general edge detection method can be used to detect the significant points, we choose an edge detector from [10], followed by the Dynamic two-strip algorithm (Dyn2S) [11] to obtain these points. Fig. 1 illustrates two examples of significant points superimposed on the original face images.

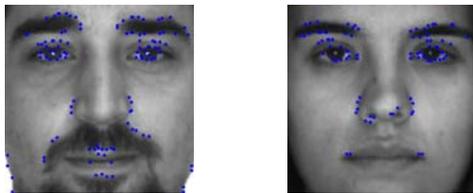


Fig. 1. Significant points.

2.2. Significant Jet Point

Gabor features are one of the widely used image feature descriptors in image processing and object recognition. They can capture the salient visual properties in an image. The 2D Gabor kernel functions used for feature extraction are

$$\psi_j(\mathbf{x}) = \frac{k_j^2}{\sigma^2} \exp\left(-\frac{k_j^2 x^2}{2\sigma^2}\right) \left[\exp(i\mathbf{k}_j \mathbf{x}) - \exp\left(-\frac{\sigma^2}{2}\right) \right], \quad (1)$$

where

$$\mathbf{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_\nu \cos \varphi_\mu \\ k_\nu \sin \varphi_\mu \end{pmatrix}, \quad k_\nu = 2^{\frac{\nu+2}{2}} \pi, \quad \varphi_\mu = \mu \frac{\pi}{8}. \quad (2)$$

The index $j = \mu + 8\nu$ covers a discrete set of five different frequencies $\nu = 0, \dots, 4$ and eight orientations $\mu = 0, \dots, 7$. The width of the Gaussian is controlled by the parameter $\sigma = 2\pi$ [5, 12]. Fig. 2 illustrates the visualized real and imaginary parts of the 40 Gabor kernel functions. A jet J is defined as the set $\{J_j\}$ of 40 convolution coefficients for kernels of different frequencies and orientations obtained at one image pixel $\mathbf{x} = (x, y)$ in an image $I(\mathbf{x})$:

$$J_j(\mathbf{x}) = \int I(\mathbf{x}') \psi_j(\mathbf{x} - \mathbf{x}') d^2 \mathbf{x}'. \quad (3)$$

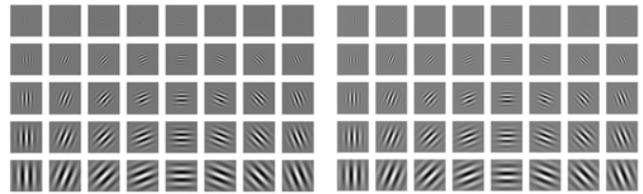


Fig. 2. Gabor kernels. (a) Real part. (b) Imaginary part.

After significant points are detected, Gabor jet is extracted from each point. A Significant Jet Point (SJP) is defined as the Cartesian coordinate $\mathbf{x} = (x, y)$ of a significant point accompanied with a Gabor jet extracted from this point:

$$SJP = (\mathbf{x}; J). \quad (4)$$

Based on the SJP description, a face is represented by a set of significant points with a Gabor jet attached on each point as textural cues. The SJP descriptor, using sparse points, reduces the storage demand of an image and improves the computational efficiency to meet the high-speed requirement in face recognition. It is also expected to be less sensitive to illumination changes due to the fact that it is a feature derived from low-level illumination-insensitive edge map representation. Meanwhile, the Gabor features on the points significantly enhance the discriminative power of the descriptor to improve the recognition accuracy.

3. JET POINT DISTANCE

Generally, face images of a same individual suffer from various external interferences such as illumination, expression, aging, etc. Since general edge detection methods are used to find significant points, it is difficult to find correspondence of SJP pairs between two faces. Hausdorff distance and its variants [13, 9] have been widely utilized as shape comparison metrics on binary images. As a purely geometrics-based method, Hausdorff distance is originally defined between two point data sets. Unlike most shape comparison methods that build a point-to-point correspondence between a model and a test image, Hausdorff distance can be calculated without explicit pairing of points in their respective data sets. In this section, we apply the Hausdorff distance measure to SJP matching

and propose a new Jet Point Distance (JPD) for image recognition. Taking both structural and textural information into consideration, the JPD presents a more reliable dissimilarity measure between two face images represented by SJPs.

3.1. Structure and Texture Measurements

In a SJP, structural information is represented by the coordinates of the significant point. For two points $\mathbf{x}=(x,y)$ and $\mathbf{x}'=(x',y')$, the Euclidean distance $d_E(\mathbf{x},\mathbf{x}')=\sqrt{(x-x')^2+(y-y')^2}$ is used as the structure measurement in this study. The textural information of a SJP is represented by the Gabor jet extracted from the significant point. As complex numbers, a Gabor jet can be written as the polar form $J_j=a_j\exp(i\phi_j)$, where jet magnitudes $a_j(\mathbf{x})$ vary slowly with spatial position and jet phases $\phi_j(\mathbf{x})$ rotate with a rate set by the spatial frequency or wave vector \mathbf{k}_j of the kernels [5, 12]. We thus use a magnitude similarity function to measure two jets J and J' :

$$S(J,J')=\frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'_j{}^2}}, \quad (5)$$

because this jet similarity measurement provides certain robustness to small position variation, which is most likely segmentation error or intra-class variation.

3.2. Jet Point Distance

Given two finite SJP sets $M=\{SJP_1^M, SJP_2^M, \dots, SJP_p^M\}$ representing a model face in the database and $T=\{SJP_1^T, SJP_2^T, \dots, SJP_q^T\}$ representing a test face from input, where P and Q are the numbers of SJPs in M and T respectively. The JPD is defined as

$$JPD(M,T)=\max(d_JPD(M,T),d_JPD(T,M)), \quad (6)$$

where the function $d_JPD(M,T)$ is called the *directed* JPD from set M to T and can be defined as

$$d_JPD(M,T)=\max_{1 \leq p \leq P} \min_{1 \leq q \leq Q} (d_E^2(\mathbf{x}_p^M, \mathbf{x}_q^T) - \lambda S(J_p^M, J_q^T)). \quad (7)$$

This directed JPD function identifies the SJP $SJP_{p^*}^M \in M$ that has the largest distance from any SJP of T and measures the distance from $SJP_{p^*}^M$ to its nearest neighbor in T . The distance between two SJPs is measured through combining Euclidean distance of significant point and magnitude similarity of Gabor jets. This is a compound measurement composed of both structural and textural information. The weight λ is used to balance the contributions of Euclidean distance and jet dissimilarity (inversed from magnitude similarity using a minus). The JPD is the maximum of $d_JPD(M,T)$ and $d_JPD(T,M)$. Therefore, it measures the

degree of mismatch between two SJP sets by measuring the distance of the SJP of M that has the largest distance from any SJP of T , and vice versa. However, this distance measure is very sensitive to outlier SJPs. A few outlier SJPs, even only a single one, can perturb the distance greatly, though the two faces might be very similar.

Realizing that there could be different ways to define the directed Hausdorff distance, Dubuisson and Jain [13] investigated 24 different distance measures and indicated that a modified Hausdorff distance has the best performance. Based on this observation, we define a modified directed JPD as

$$d_JPD(M,T)=\frac{1}{P} \sum_{1 \leq p \leq P} \min_{1 \leq q \leq Q} (d_E^2(\mathbf{x}_p^M, \mathbf{x}_q^T) - \lambda S(J_p^M, J_q^T)). \quad (8)$$

The definition of the undirected JPD is the same as Equation (6). This distance measure can alleviate the sensitivity to outlier SJPs. For a test face image, the face recognition system calculates the JPD between the test face and each model face in the database. The model with minimum distance is considered as the correct return.

4. EXPERIMENTAL RESULTS

The proposed method is assessed on the publicly available AR face database [8]. Because images in some sessions are either missing or corrupted, we eventually obtain 120 complete set of images (65 men and 55 women) in our experiments. All the images are normalized (in scale and rotation) and cropped to 160×160 pixels based on the manually labeled positions of two eyes. Two examples of the cropped faces are illustrated in Fig. 1.

4.1. Determining the Weight λ

In Equation (8), the weight λ balances the contributions of structural and textural measurements. When $\lambda=0$, the JPD reduces to a pure point Hausdorff distance. To determine λ , a series of preliminary experiments is performed using all the neutral expression faces in AR database. The model set is the neutral faces of the first session, and the test set is those of the second session. The top-one recognition rate against the weight λ is displayed in Fig. 3. From the figure, it is found that the recognition rate reaches maximum when $\lambda=160$. The weight $\lambda=160$ is selected and used in the rest of the experiments. This parameter determination method can be employed to achieve optimal settings in other applications with different face database.

In the following, the proposed method is compared with the Directional Corner Point (DCP) method [9] under various situations, using the neutral faces in normal condition taken in the first session as the model set.

4.2. Face Recognition under Various Conditions

The face images under controlled condition in the second

session are first used to evaluate the proposed method. The comparative recognition accuracy is illustrated in Table 1. Although the number of images used in this paper is more than that in DCP, the proposed method still outperforms DCP method by 4.53%.

To compare the recognition accuracy with expression variations, the experiment is also performed on three different sets of images with smiling, angry and screaming expressions in the first session. The results are listed in Table 2. It can be seen from the table that the performance of the proposed method is significantly better than DCP method under all three expression variations, especially under the smiling condition, where the improvement is close to 30%. The recognition rate under the screaming expression is the lowest due to the fact that screaming produces the most considerable physical deformation on a human face, and further heavily distorts the locations of significant points.

We finally perform the experiment under the condition of illumination changes. The AR database contains three different lighting conditions: left light, right light and both lights on. Table 3 displays these experimental results. The proposed method obtains a noticeable recognition accuracy of 100% when either left or right light on. This demonstrates that the proposed method is very tolerant to lighting changes. However, it is still sensitive to extreme lighting, which causes strong specular reflectance on the face skin and thus could erase some significant points.

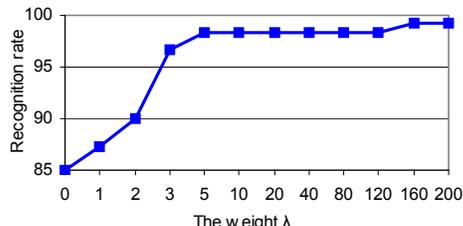


Fig. 3. Recognition rate against the weight λ .

Table 1. Recognition rates under controlled condition

Methods	DCP [9]	The proposed method
Database size	112	120
Recognition rate	94.64%	99.17%

Table 2. Recognition rates under different expressions

Methods	DCP [9]	The proposed method
Database size	112	120
Smiling	63.39%	92.50%
Angry	93.75%	96.67%
Screaming	27.68%	53.33%

Table 3. Recognition rates under different illuminations

Methods	DCP [9]	The proposed method
Database size	112	120
Left light on	87.50%	100.00%
Right light on	89.29%	100.00%
Both lights on	61.61%	75.00%

5. CONCLUSIONS

This paper presents a novel analytic face recognition method using Gabor features on sparse points for face recognition. Taking both structural and textural information into consideration, the proposed face representation and dissimilarity measurement are proved to be not only efficient and reliable, but also robust to illumination, expression and possibly occlusion. We anticipate the proposed method could be extended to other object recognition tasks in the future.

6. ACKNOWLEDGEMENT

This research was supported by the Australian Research Council (ARC) Discovery Grants DP0451091 and DP0877929.

7. REFERENCES

- [1] C. Liu and H. Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition," *IEEE TIP*, vol. 11, pp. 467-476, 2002.
- [2] C. Liu and H. Wechsler, "Independent Component Analysis of Gabor Features for Face Recognition," *IEEE Transactions on Neural Networks*, vol. 14, pp. 919-928, 2003.
- [3] C. Liu, "Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition," *IEEE TPAMI*, vol. 26, pp. 572-581, 2004.
- [4] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz, and W. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Transactions on Computers*, vol. 42, pp. 300-311, 1993.
- [5] L. Wiskott, J.M. Fellous, N. Krüger, and C. von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE TPAMI*, vol. 19, pp. 775-779, 1997.
- [6] T.S. Lee, "Image Representation Using 2D Gabor Wavelets," *IEEE TPAMI*, vol. 18, pp. 959-971, 1996.
- [7] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE TPAMI*, vol. 22, pp. 1090-1104, 2000.
- [8] A.M. Martinez and R. Benavente, The AR Face Database, *CVC Technical Report #24*, June 1998.
- [9] Y. Gao and Y. Qi, "Robust Visual Similarity Retrieval in Single Model Face Databases," *Pattern Recognition*, vol. 38, pp. 1009-1020, 2005.
- [10] R. Nevatia and K.R. Babu, "Linear Feature Extraction and Description," *Computer Graphics and Image Processing*, vol. 13, pp. 257-269, 1980.
- [11] M.K. Leung and Y.-H. Yang, "Dynamic Two-Strip Algorithm in Curve Fitting," *Pattern Recognition*, vol. 23, pp. 69-79, 1990.
- [12] S. Zhao, W. Gao, S. Shan, and B. Yin, "Enhance the Alignment Accuracy of Active Shape Models Using Elastic Graph Matching," *Proceedings of ICBA*, pp. 52-58, 2004.
- [13] M.P. Dubuisson and A.K. Jain, "A Modified Hausdorff Distance for Object Matching," *Proceedings of ICPR*, pp. 566-568, 1994.