Face Recognition Based on Weighted Multi-order Feature Fusion of 2D-FrFT

Xu Wang[(✉)], Lin Qi, Yun Tie, and Enqing Chen

Information Engineering School, Zhengzhou University,
Zhengzhou 450001, China
`iewxalex@l63.com`

Abstract. The fractional Fourier transform (FrFT) features have been known to be effective for face recognition. However, only a few approaches utilize phase feature and they usually perform worse than those using magnitude feature. To investigate the potential of FrFT phase and its fusion between different orders for face recognition, in this paper, we first propose weighted multi-order band fusion of generalized phase spectrum (WMFP) of 2D-FrFT. Compared with the conventional appearance-based face recognition method, the proposed method does not need to perform image-to-vector conversion and can well preserve the discriminative information of the original image. Different from the existing Fourier-based recognition approaches such as Fourier-LDA and local region histogram of 2D-FrFT magnitude and phase (LFMP), the proposed approach merges multiple orders' generalized phase spectrum of 2D-FrFT and gives different weights to different orders simultaneously. Experimental results on two benchmark face databases demonstrate the effectiveness of the proposed method and indicate that our method is better than Fourier-PCA and LFMP, as well as other popular face recognition methods such as Gabor-based linear discriminant analysis (GLDA) and local Gabor binary patterns (LGBP).

Keywords: Face recognition · Two dimensional fractional Fourier transform (2D-FrFT) · Generalized phase spectrum (GPS) · Fusion · Weighted multi-order band fusion of generalized phase spectrum (WMFP)

1 Introduction

Due to the advantages of non-intrusive natural and high uniqueness, as well as the increasing need for security-related applications, face recognition has been an active topic for researchers in the field of computer vision and pattern recognition [1]. A wide range of applications, from public security to personal consumer electronics, have made face recognition one of the most popular research topics. With the past decade, numerous face recognition approaches have been reported in the literature to handle variations in illumination, expression and occlusion etc. Despite a significant level of maturity and a few practical successes, finding efficient and robust algorithms for face recognition is still an active and challenging topic.

As a generalization of Fourier transform (FT), the fractional Fourier transform (FRFT) is a powerful signal analysis and processing tool. The Fourier analysis [2, 3] is

an effective analysis tool for facial images. However, it abnegates the time-domain information of images absolutely so that it will lose some significant information of faces inevitably. The conventional FT can be regard as a rotation in the frequency plane. However, FrFT performs a rotation of signal to any angle. Therefore, the two dimensional fractional Fourier transform (2D-FrFT) contains the time-frequency information of the signal simultaneously. Facial image can be regarded as a gradually changed signal. It has been shown that the 2D-FrFT is an effective analysis tool for facial images [6]. Gao et al. made use of the phase information of the 2D-FrFT for emotion recognition [4]. Wang et al. utilized the phase and Magnitude information of 2D-FrFT for face registration [15]. Kong et al. utilized 2D-FrFT with fuzzy fusion classification algorithm in human emotional state recognition [14].

Motivated by the above-mentioned reason, in this paper, we adopt 2D-FrFT phase information to improve the effect of face recognition. First, we extract the phase information of facial image in 2D-FrFT, which is called the generalized phase spectra (GPS). It has been shown that the lower frequency bands contain more smooth information, while high bands contain the edge information of a facial image [2]. Then, we present an improved two-dimensional separability judgment to select appropriate order parameters for 2D-FrFT. Through selecting the appropriate transform orders and choosing the optimal generalized phase spectrum band, a new spectrum feature can be extracted, which not only contains the smooth information but also includes the edge information of a facial image. In order to make full use of the discriminative information from different orders for face recognition, the proposed approach merges multiple orders' GPS band of 2D-FrFT and gives different weights to different orders simultaneously, as illustrated in Fig. 1. Finally, we perform the classification using the nearest neighbor classifier. In the experiments, two famous public face databases are employed as the test data.
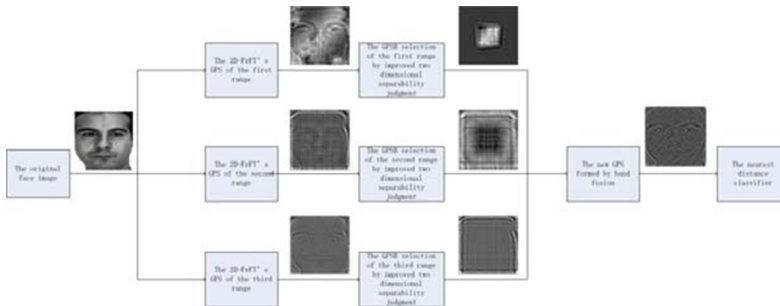

Fig. 1. Illustration of the WMFP.

In the following sections, we will first perform the analysis of FRFT and the definition of the GPS in Sect. 2. Then our face recognition approach is presented in Sect. 3. The experiment results based on the nearest neighbor classifier are summarized in Sect. 4. And finally, conclusions are drawn in Sect. 5.

2   Fractional Fourier Transform

The discrete FrFT [6] is a new time-frequency analysis tool which is developed and widely used in recent years. In essence, the signal representation in fractional Fourier domain integrates the information of time domain and frequency domain. The 2D-FrFT on each image is defined as follows:

$$\mathbf{X}_{(\alpha,\beta)}(m, n) = \sum\nolimits_{p=0}^{M-1} \sum\nolimits_{q=0}^{N-1} x(p, q) K_{(\alpha,\beta)}(p, q, m, n) \qquad (1)$$

$$K_{(\alpha,\beta)} = K_\alpha \otimes K_\beta \qquad (2)$$

$$K_p = \frac{A_\alpha}{2\Delta x} \exp\left(\frac{j\pi(\cot \cot \alpha)m^2}{(2\Delta x)^2} - \frac{j2\pi(\csc \csc \alpha)mn}{(2\Delta x)^2} + \frac{j2\pi(\cot \cot \alpha)n^2}{(2\Delta x)^2} \right) \qquad (3)$$

The transform kernel is $K_{(\alpha,\beta)}$, as defined as (2). 2D-FrFT is equivalent to apply discrete FrFT on the two parameters of a signal $x(p, q)$ successively. The separable kernel K_p with respect to different orders for x and y axes indicates additional degrees of freedom that can be used for specific applications in the feature extraction.

As we all know, two dimensional Fourier transform (2DFT) is not reality preserving. With the 2DFT, the original data is transformed into complex-value which contains both magnitude information and phase information. So does the 2D-FrFT, which is the generalized form of 2DFT. It is inflected in the phenomenon that the transform order can be selected according to different needs. In this paper, since the phase contains a large amount of edge information [4], it is more conducive to identify the face. It is extended from the phase spectrum by the FT. GPS means that a facial image with the 2D-FrFT phase information is distributed on the fractional domain.

3   Proposed Approach

It has been demonstrated that not all phase spectrum bands of FT are beneficial for face recognition [2]. This motivates us to exploit the discriminative nature of different GPS bands. Obviously, how to choose the appropriate GPS bands from different transform orders is the key question. According to the properties of FrFT [6], it is periodic with periodicity equal to 4. From the definition, we give detailed description about the transform order that changes from 0.1 to 1. Pei et al. has given some explanation about the phase information of different transform orders [6]. The transform order p changing from 0.1 to 1 is divided into three ranges according to the inter-class and intra-class difference of phase information. The orders can be divided into three ranges, small order range (0.1 to 0.4), middle order range (0.5 to 0.7) and large order range (0.8 to 1) respectively. The Fig. 2 indicates that the difference of inter-class figures from different ranges is obvious while the intra-class figures from the same range are similar.

As discussed in [3], the small transform orders will lose edge information of an image with the 2D-FrFT, and the large transform orders will contain much noise, such as illumination, expression and so on. As for the middle order, it is the low-abiding, but
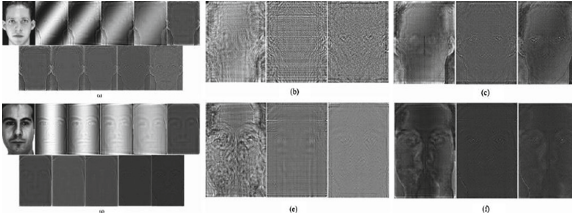
Fig. 2. (a) is the transformed face image from ORL database according to the orders changing from 0–1, (b) and (c) are the intra-class and inter-class average difference of three ranges based on ORL database, as well as (d), (e) and (f) are based on AR database.
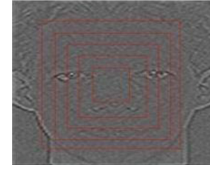
Fig. 3. Illustration of the GPS bands.

we can also think that it cannot make full use of the low spectral phase band and the high spectral phase band yet. To make full use of the GPS bands of different ranges, we select the transform order from each range randomly and then choose the bands by the improved two-dimensional separability judgment. When the division of orders is completed, how to choose the GPS bands from the transform ranges is critical. The two-dimensional separability judgment [2] is adopted to select the frequency band, but it neglects the transformed data structure. In this paper, the improved two-dimensional separability judgment is proposed which takes the transformed data information into account instead of the face images. Suppose that the original image sample set is X, each gray image matrix is sized $M * N$ and expressed by $X(p, q)$, where $1 \leq p \leq M$, $1 \leq q \leq N$. We perform 2D-FrFT as depicted as the Eq. (4):

$$F_{(\alpha,\beta)}(u, v) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} X(p, q) K_{(\alpha,\beta)}(p, q, u, v) \tag{4}$$

For simplicity, we still use X to represent the transformed data. We will choose the GPS bands according to the properties of phase information from different ranges. Let $F_{(\alpha,\beta)}(\mu_0, v_0)$ indicates the zero PSB. Shift $F_{(\alpha,\beta)}(\mu_0, v_0)$ to the center of the image at $(M/2, N/2)$. Since the GPS domain is represented by the matrix form, we use a square ring Ring(k) to represent the kth GPS band, where $0 \leq k \leq M/2$. The four vertexes of Ring(k) are $(u - k, v - k)$, $(u + k, v - k)$, $(u - k, v + k)$ and $(u + k, v + k)$ respectively. So, the kth GPS band denotes:

$$F_{(\alpha,\beta)}(\mu, v) \in Ring(k) \tag{5}$$

Different GPS bands with the above expression way are illustrated in Fig. 3. If we select the GPS band, keep the original values of $F_{(\alpha,\beta)}(\mu, v)$, otherwise set the values of $F_{(\alpha,\beta)}(\mu, v)$ to be zero. The selection step can be expressed by the Eq. (6):

$$F_{(\alpha,\beta)}(\mu, v) = \begin{cases} \text{Original values} & \text{if } F_{(\alpha,\beta)}(\mu, v) \in Ring(k) \\ 0 & \text{if } F_{(\alpha,\beta)}(\mu, v) \notin Ring(k) \end{cases} \tag{6}$$

Then we can evaluate the separability J_k by the improved two-dimensional separability judgment. Let $A_i(i = 1, 2, \ldots, C)$ denote a mean value of w_i class and A denote the total mean value of X. A_i and A are in the form of the matrix. $F_{k(\alpha,\beta)}$ represents the kth GPS band. With regard to the kth GPS band, we can compute the within-class S_w matrix and between-class S_b matrix according to the Eqs. (7) and (8), and the trace ratio is calculated by the Eq. (9).

$$S_b = \sum\nolimits_{i=1}^{C} \left[(A_i - A)(A_i - A)^T \right] \qquad (7)$$

$$S_w = \sum\nolimits_{i=1}^{C} \left[\left(A_i - F_{k(\alpha,\beta)} \right) \left(A_i - F_{k(\alpha,\beta)} \right)^T \right] \qquad (8)$$

$$J_k = \frac{tr(S_b)}{tr(S_w)} \qquad (9)$$

For all the GPS bands, we select the bands with high trace ratio by maximizing the between-class scatter while minimizing the within-class scatter.

4   Experimental Results

In this section, we present the experimental results on public available databases for face recognition, which serve to illustrate the efficacy of the proposed approach. The following describes the details of the experiments and results. The nearest neighbor classifier with the Euclidean distance is applied for recognition.

4.1   Experiments on the FERET Face Database

The FERET database consists of 13539 facial images corresponding to 1565 subjects, who are diverse across ethnicity, gender and age. From the FERET face database, we select a subset includes 1400 images of 200 individuals (each individual has seven images) in the experiment. There are 71 females and 129 males. This database involves the variations in facial expression, illumination and poses ($\pm 15°$ and $\pm 25°$). All the chosen images are resized to 64 * 64 with 256 gray levels per pixel. Figure 4 shows some sample images of one subject from the FERET database. In the experiment, the first three images of each person are used as training samples and the remainder as test samples. So, there are 600 (=200 * 3) training samples and 800(=200 * 4) test samples.

Fig. 4. The samples of cropped face images from the FERET database.

Tables 1 and 2 show the accuracy of different transform orders and the comparison between WMFP and Eigenface, Fisherface, SRC [10], LGBP [14], Fourier + PCA [2] and LFMP [15] separately on the FERET database. We implemented these methods ourselves and tuned the parameters for each method for fair comparison. The results verify the proposed method powerfully. Figure 5 indicates the identity of GPS bands from different transform orders, from which we select 0.4, 0.5 randomly and 0.9 as the band fusion orders. It is a positive correlation between color and the trace ratio of the GPSB. The closer to the white color, the higher trace ratio of GPSB.

Table 1. The recognition rate with the discrimination of different orders on the FERET database

Transform order	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	WMFP
Recognition rate	64%	64%	68%	67%	63%	58%	65%	65%	66%	63%	95%

Table 2. The comparison results on the FERET database

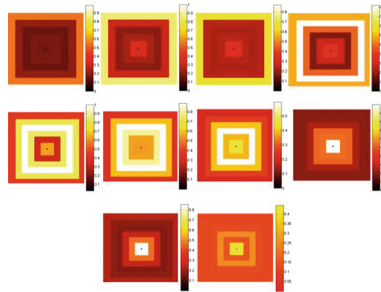| Different methods	Eigenface	Fisherface	SRC	LGBP	Fourier + PCA	LFMP	WMFP
Recognition rate	56.13%	61.2%	76.5%	80.4%	82.94%	90.5%	95%

Fig. 5. Illustration of the GPS bands according to different transform orders on the FERET database.

For the experiments on the FERET database, we test the effectiveness of our method to the overall problems. As can be seen from this table, the proposed WMFP outperforms the 6 compared methods with the lowest gains in accuracy of 5% on the FERET database. We have made three observations from the results listed in Tables 1 and 2: (1) Table 1 shows that different orders of 2D-FrFT may result in different recognition rates. However, the highest rate 68%, which is achieved in order 0.3, is still not satisfying, due to the fact that the single order 2D-FrFT feature is always not enough for recognition. Not only the amount of information, but also the redundancy is increasing when weighted multi-order bands are fused. (2) The Gabor method is very time consuming and more suitable for high-resolution image which imposes a heavy

computational burden on the target device, in particular on mobile devices, which have low computational power. The proposed method is no need to construct the subspace through the feature extraction methods and has less computation cost. (3) WMFP obtains the best recognition performance on all the experiments, which implies that both edge and smooth information of facial image are important for recognition. Compared with other methods, the WMFP is more intuitive and effective with considerable results.

4.2   The Robustness to the Illumination Based on the Extended Yale-B Face Database

The Extended Yale-B database consists of 2414 frontal face images of 38 individuals under various laboratory-controlled lighting conditions. All the test images used in the experiments are manually aligned, cropped, and then resized to 64 * 64 images. The Extended Yale-B set only has little variability of expression, aging, etc. However, its extreme lighting conditions still make it a challenging task for most face recognition methods. The database is divided into five subsets, which can be seen from the Table 3. In this experiment, we select 0.2, 0.5 and 0.8 as the band fusion orders, which are selected randomly from the three transform ranges.

Table 3. Data partition on Extended Yale-B database for various experiments

| Category	Training	Testing
Subset 1	Consisting of 266 images under normal lighting conditions	Consisting of 266 images under normal lighting conditions
Subset 2	Consisting of 266 images under normal lighting conditions	12 images per subject, characterize slight-to-moderate illumination variations
Subset 3	Consisting of 266 images under normal lighting conditions	12 images per subject, characterize moderate-to-drastic illumination variations
Subset 4	Consisting of 266 images under normal lighting conditions	14 images per person
Subset 5	Consisting of 266 images under normal lighting conditions	19 images per person
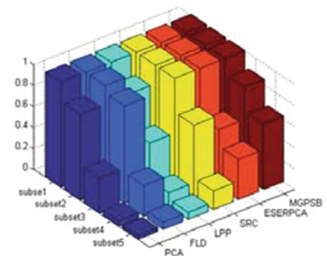
Fig. 6. The comparison on the Extended Yale-B database is revealed in this figure.

Figure 6 shows the performance for varying Subsets. We compare the proposed method with PCA, FLD, LPP, SRC and ESERPCA [11], in light of its capability of handling illumination changes. The three classical methods are sensitive to the illumination varying. Note that SRC, ESERPCA and the proposed method show excellent performance for moderate light variations, yielding 100%. In particular, WMFP obtains the best recognition rate of 62.36% on Subset 5. By contrast, recognition accuracy of the other approaches drops heavily in this Subset. It is not only more visualized to display the performance between different Subsets and methods, but also demonstrate the robustness of the proposed algorithm to the varying illumination.

5 Conclusion

This paper proposes a new method called WMFP which improve the phase information accuracy of 2D-FrFT significantly. The GPS bands are extracted from different orders of 2D-FrFT, which contains both the smooth information of small orders and the edge information of large orders. Following our method, the accuracy of the phase information achieves a preferred result. It also demonstrates the efficacy to the illumination varying which makes full use of the phase information of different transform orders. For future work, we are interested in designing effective classifiers to improve the recognition performance. Moreover, how to combine 2D-FrFT magnitude and phase of face images to further improve the face recognition accuracy appears to be another interesting direction of future work.

REFERENCES

1. Jain, A., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circ. Syst. Video Technol. **14**(1), 4–20 (2004)
2. Jing, X.Y., Tang, Y.Y., Zhang, D.: A Fourier-LDA approach for image recognition. Pattern Recogn. **38**(3), 453–457 (2005)
3. Sao, A.K., Yegnanarayana, B.: On the use of phase of the Fourier transform for face recognition under variations in illumination. Sig. Image Video Process. **4**(3), 353–358 (2010)
4. Gao, L., Qi, L., Chen, E.Q., Mu, X.M., Guan, L.: Recognizing human emotional state based on the phase information of the two dimensional fractional Fourier transform. In: 11th Pacific Rim Conference on Multimedia, pp. 694–794 (2010)
5. Tao, R., Qi, L., Wang, Y.: The theory and Applications of Fractional Fourier Transform. Tsinghua University Press, Beijing (2004)
6. Jing, X.Y., Wong, H.S., Zhang, D.: Face recognition based on discriminant fractional Fourier feature extraction. Pattern Recogn. Lett. **27**(13), 1465–1471 (2006)
7. Pei, S.C., Ding, J.J.: Relations between fractional operations and time-frequency distributions and their applications. IEEE Trans. Sig. Process. **49**(8), 1638–1655 (2001)
8. Choi, S.: Combined features for face recognition. Electron. Lett. **52**(1), 0013–5194 (2016)

9. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)

10. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 210–227 (2009)

11. Luan, X., Fang, B., Liu, L.H., Yang, W.B., Qian, J.: Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion. Pattern Recogn. **47**(2), 495–508 (2014)

12. Liu, C., Wechsler, H.: Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition. IEEE Trans. Image Process. **11**, 467–476 (2002)

13. Zhang, W., Shan, S., Gao, W., et al.: Local Gabor Binary Pattern Histogram Sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In: International Conference on Computer Vision, pp. 786–791 (2005)

14. Kong, M., Qi, L., Zheng, N., et al.: Application of the 2D-FrFT combined with fuzzy fusion classification algorithm in human emotional state recognition. In: TMEE, pp. 1094–1097 (2011)

15. Wang, Y.X., Qi, L., Guo, X., Gao, L.: Face recognition based on histogram of the 2DFrFT magnitude and phase. In: Information Science, Electronics and Electrical Engineering (ISEEE), vol. 3, pp. 1421–1425 (2014). doi:10.1109/6946154