

Efficient Image Plane Rotation Invariant Frequency Domain Face Recognition Technique Using Eye Localization

Papia Banerjee¹, Pradipta K. Banerjee², and Asit K. Datta³

¹ Department of Computer Science and Engineering
ABACUS Inst. of Engg. & Mgmt., Hoogly
papia.banerjee24@gmail.com

² Department of Electrical Engineering
Future Institute of Engineering and Management
Kolkata-700150
pradiptak.banerjee@gmail.com

³ Department of Applied Optics and Photonics
University of Calcutta, Kolkata-700009
asitdatta@gmail.com

Abstract. An efficient frequency domain face recognition under arbitrary image plane rotations by single correlation filtering approach is proposed where faces are registered with eye detection. Here eye detection problem is carried out using shift invariant property of correlation filter. The proposed eye detection method includes log-polar transformation, correlation and regression neural network estimation. Proposed system shows the recognition improvement comparing to straightforward correlation filtering and multi-correlation approach.

1 Introduction

Frequency domain face recognition techniques are executed by cross-correlating the Fourier transform of test face image with a synthesized correlation filter (CF), generated from Fourier transform of training images and processing the resulting correlation output via inverse fast Fourier transform [4]. Though CFs are successfully used in face recognition, [4], the performance of CFs degrades drastically while the face images suffer from image plane rotation. In this case the standard CFs fail to detect the authentic person. To improve the performance of CFs it is then necessary to train the system with more images at certain intervals of angular positions, which causes degradation of signal-to-noise ratio. Further question may arise on what is the optimum number of rotated face images required for synthesizing one CF or how many synthesized CFs are enough for accurate recognition for arbitrary rotation of face images. Designing a set of CFs for a single person implies multi correlation scheme in testing phase, which increases time complexity. To overcome these problems single correlation approach is carried out to recognize faces under any in-plane rotation for

properly registered faces which is achieved by automatic localization of eyes. The eye detection process is carried out with log-polar transformation, correlation filter and neural network estimation. The effectiveness of the proposed method is verified with JAFEE database.

2 Proposed Face Registration Method

Fig.(1) shows the detail process of eye detection and face registration. The face image with any rotation ($0^\circ - 360^\circ$) is first log-polar transformed. A variant of unconstrained CF [2] termed as Eye-filter (EF) is designed with the eye templates, cropped from log-polar image in the training phase. The log-polar image is Fourier transformed and cross-correlated with EF. Eye is detected due to the shift invariant property of EF. The location of eye is found by searching the maximum value in the correlation plane. This location of one eye in log-polar is mapped into the original image for the both eyes positions by using the pixel coordinate information. The information of the pixel coordinates in two different domains is then fed to the generalized regression neural network (GRNN)[5]. In the testing stage this network automatically provides the actual position of eyes of rotated faces whenever the eye is detected in log-polar domain by the designed EF. With the actual eye coordinates the angular position of test face is determined and hence rotated in the opposite direction to register the face in its 0° position.

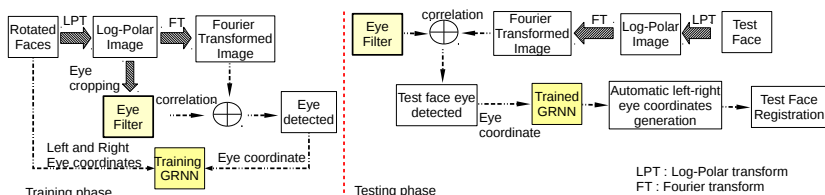


Fig. 1. Block diagram of detail process of face registration

3 Log-Polar Transformation and Eye Filter Synthesis

In addition to mapping in plane rotation into linear shift, log-polar transform offers some other features, such as wide field of view and a highly focused central area [3] as shown in Fig. 2. The measurement of linear shift of eyes in log-polar domain gives the information of the angular position of the face images in original scene. In case of unconstrained minimum average correlation energy (UMACE) [1] filter, conventionally, \mathbf{x}_i (i th image) is used as an exemplar. However, instead of using \mathbf{x}_i , $(\mathbf{x}_i - \beta \mathbf{m})$, (where \mathbf{m} represents mean image in Fourier domain) is introduced to modify the UMACE filter solution.

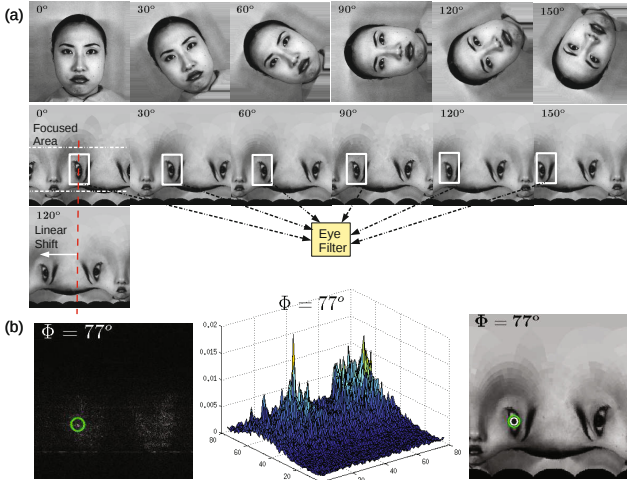


Fig. 2. (a) Log-polar transformation of rotated images and linear shift of eye. (b) Eye detection in log-polar image corresponding to maximum peak in correlation plane.

With this idea the average similarity measure (ASM) [1] can be expressed as

$$ASM_{new} = \frac{1}{N} \sum_{i=1}^N |\bar{\mathbf{X}}_i^* \mathbf{h} - (1 - \beta)\bar{\mathbf{M}}^* \mathbf{h}|^2 = \mathbf{h}^+ \bar{\mathbf{S}}_{new} \mathbf{h} \quad (1)$$

where

$$\bar{\mathbf{S}}_{new} = \frac{1}{N} \sum_{i=1}^N (\bar{\mathbf{X}}_i - (1 - \beta)\bar{\mathbf{M}})(\bar{\mathbf{X}}_i - (1 - \beta)\bar{\mathbf{M}})^* \quad (2)$$

Replacing $\bar{\mathbf{D}}$ of UMACE solution[1], by $\bar{\mathbf{S}}_{new}$ [2], the eye filter solution becomes

$$\mathbf{h}_{EF} = (\bar{\mathbf{S}}_{new})^{-a} (\bar{\mathbf{S}}_{new})^{-b} \mathbf{m} \quad (3)$$

Eq.(3) represents the eye filter (EF) synthesized with the eye templates as shown in Fig.(2). Having designed \mathbf{h}_{EF} filter it is now correlated with the Fourier transformed log-polar image the shifted eye position will be detected and located by the maximum peak value on the resulting correlation surface. Fig.(2b) shows the correlation plane with a peak at desired location.

4 Estimating Eye Coordinates Using GRNN

Corresponding to the eye coordinate $P(u, v)$, shown in Fig.(3a) in log-polar image the two eye coordinates $E(x_1, y_1)$ and $E(x_2, y_2)$ of original image are manually collected in training stage. For different angle positions ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$) of faces, a set of $P(u, v)$, $E(x_1, y_1)$ and $E(x_2, y_2)$ are obtained. These

values are fed to GRNN which can approximate any input-output map. Input layer of GRNN (Fig. 3b) is connected to pattern layer p_i which may contain q nodes, where q is the number of samples within a training data set. Each node represents the input vector, w_j , associated with the j th sample in training data. The signals of the pattern neurons i , are flowing into the third layer termed as summation layer. In this study $E(x_1, y_1)$ and $E(x_2, y_2)$ are used as target vectors where $P(u, v)$ is used for input vector w_i for training of GRNN. Given an unknown sample w' , the approximation for Y can be obtained.

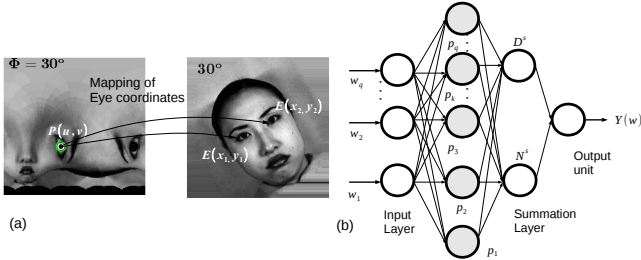


Fig. 3. (a) The mapping of eye coordinate in log-polar domain into original eye coordinates in Cartesian system is shown. (b) The typical GRNN model used for training.

5 Experimental Results

5.1 Eye Detection

Each image of JAFFE database has been manually rotated by angle in the range of $0 - 360^\circ$ with 1° one degree interval, to get $360 \times 213 = 76,680$ number of images for testing purpose. This huge number of images are tested with the proposed method while the training data size contains only (a) 6 eye templates of one individual shown in Fig.(2) to synthesize the \mathbf{h}_{EF} filter and (b) 18 coordinates (3 for each image) set for training of GRNN. With the help of \mathbf{h}_{EF} and GRNN, eyes are found at any arbitrary angular position of test face image. Fig.(4) shows the eye localization of different images at different angular image plane rotations.

Having obtained the eye coordinates in Cartesian system the angular position of the test face is obtained as

$$\theta = \tan^{-1} \frac{y_2 - y_1}{x_2 - x_1} \quad (4)$$

where $\{x_i, y_i\}_{i=1,2}$ are the coordinates of eyes. After calculating θ the image is rotated by an angle $-\theta$ to get the retrieved face image.

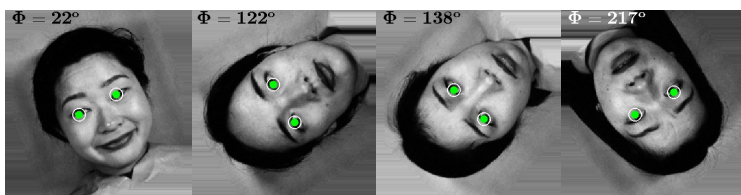


Fig. 4. Accurate eye localization of eyes at different angular positions of different images

5.2 Performance Analysis of Proposed Technique in Single Correlation Approach

At testing stage a standard CF[1] is used. This CF is synthesized with 5 images at 0° position with image index $\{69, 72, 75, 78, 82\}$ of one person (shown in Fig.(4a)). One of the training images is now roated at an angle 70° and then correlated with the designed CF. Fig.(5a) shows the poor PSR [4] value (5.13) obtained in this case. It is due to the fact the CF is designed at 0° position and hence it could not recognize the image with 70° rotation. Next the CF is again designed with the same images but at different angular positions. Each image is rotated from 0° to 360° with equal interval of 10° . Hence one CF is designed

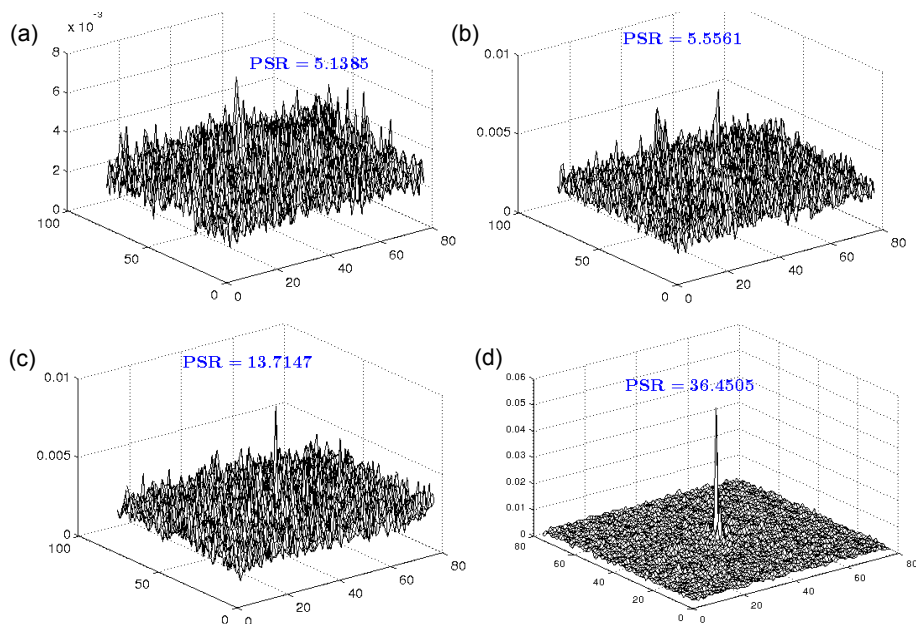


Fig. 5. Correlation planes in response to same test face for (a-b) different CF designs with angular positions (c) corresponds to multi-correlation approach (d) Proposed technique with single correlation approach

with 5×37 images. This CF is now tested with the image with 70° rotation and the correlation output is shown in Fig.(5b).

Again poor PSR value (5.55) is obtained and no prominent peak is found. So increasing the size of the training set size does not resolve the problem here. Next the multi correlation approach is performed. Instead of one CF, for one person 4 CFs are designed. $0^\circ - 360^\circ$ rotated images are divided into 4 sub-angular divisions as $0 - 90^\circ$, $90 - 180^\circ$, $180 - 270^\circ$, $270 - 360^\circ$ and each subdivision contains 10° intervals. With each angular subdivision one CF is synthesized. Such 4 CFs are correlated with the same test image and the corresponding correlation output is shown in Fig.(5c) corresponding to highest PSR value (13.71). Finally the same image with same angular condition is tested with the same CF (designed with only 5 mages at 0° position) after face registration through eye detection. The resulting correlation output is shown in Fig.(5d) where a high PSR value (36.45) is obtained and sharp distinct peak observed at the correlation plane. In case of multi correlation approach 0.28 sec is elapsed and in case of the proposed technique 0.34 sec is consumed. But in term of PSR value and nature of correlation plane the proposed technique of single correlation approach outperforms the multi correlation approach.

6 Conclusions

This paper shows that the single correlation approach can be performed instead of multi correlation process for arbitrarily rotated (in-plane) faces by registering faces with eye detection. Proposed method considers only image plane rotations. How to extend the study in both illumination and pose variation along with image plane rotation needs further investigation.

References

1. Mahalanobis, A., Kumar, B.V., Song, S., Sims, S., Epperson, J.: Unconstrained correlation filter. *Applied Optics* 33, 3751–3759 (1994)
2. Banerjee, P.K., Datta, A.K.: Generalized regression neural network trained preprocessing of frequency domain correlation filter for improved face recognition and its optical implementation. *Optics and Laser Technology* 45, 217–227 (2013)
3. Bone, P., Young, R., Chatwin, C.: Position-, rotation-, scale-, and orientation-invariant multiple object recognition from cluttered scenes. *Optical Engineering* 45 (2006)
4. Kumar, B.V.K., Savvides, M., Xie, C., Venkataramani, K., Thornton, J., Mahalanobis, A.: Biometric verification with correlation filters. *Applied Optics* 43(2), 391–402 (2004)
5. Rutkowski, L.: Generalized regression neural networks in time-varying environment. *IEEE Transaction on Neural Networks* 15(3), 576–596 (2004)