SIViP (2010) 4:353–358
DOI 10.1007/s11760-009-0125-4

ORIGINAL PAPER

On the use of phase of the Fourier transform for face recognition under variations in illumination

Anil Kumar Sao · B. Yegnanarayana

Received: 17 November 2008 / Revised: 20 February 2009 / Accepted: 7 July 2009 / Published online: 21 July 2009

Abstract In this paper, we propose a representation of the face image based on the phase of the 2-D Fourier transform of the image to overcome the adverse effect of illumination. The phase of the Fourier transform preserves the locations of the edges of a given face image. The main problem in the use of the phase spectrum is the need for unwrapping of the phase. The problem of unwrapping is avoided by considering two functions of the phase spectrum rather than the phase directly. Each of these functions gives partial evidence of the given face image. The effect of noise is reduced by using the first few eigenvectors of the eigenanalysis on the two phase functions separately. Experimental results on combining the evidences from the two phase functions show that the proposed method provides an alternative representation of the face images for dealing with the issue of illumination in face recognition.

Keywords Fourier transform · Face recognition · Phase wrapping · Eigenanalysis

1 Introduction

The objective of face recognition task is to recognize a person using his/her face image information [1]. It is a biometric system. Biometric systems are automated methods to recognize a person using his/her physiological characteristics [2]. Other physiological characteristics such as iris pattern [3], finger print [4], and behavioral patterns such as voice [5], handwriting [6] and key-stroke pattern [6], are also used in biometric systems. Issues involved in face recognition are variations due to pose, expression and illumination [1,7]. In this paper, we address the issue of variation due to illumination in face recognition.

Changes in illumination in the face recognition are normally dealt with either by modeling the effect of illumination on faces or by extracting features that are less sensitive to illumination. Modeling the effect of illumination requires samples of the face image at various illumination conditions [8,9]. Model building may add some artifacts and smear some unique information of the person's face image. These artifacts may in turn decrease the performance of face recognition. Extracting features that are less sensitive to illumination requires suitable representation of the face image. Some of the representations can be found in [10–12], and one of the them is the edge map [13]. There are two issues that need to be addressed in using the edge map for face recognition. These are: (1) computation of the edge map, and (2) matching the edge maps of two face images. Edge maps are obtained by computing the intensity gradient, and then thresholding the gradient. Selection of the threshold value becomes a major issue in the representation of an edge map. If the threshold value is low, spurious edges may show up in the edge map. On the other hand, a high threshold value may remove some important edge information. One way to address this threshold issue is to use continuous edge gradient (edginess) representation of face images [14]. In matching of the edge maps, even small deviation in the edge contour of the same person's face image may lead to poor matching. Some matching techniques proposed in the literature include distance transform [15] and Hausdorff distance [16]. These techniques

A. K. Sao (✉)
Speech and Vision Laboratory, Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai 600036, Tamil Nadu, India
e-mail: anil.sao@gmail.com

B. Yegnanarayana
International Institute of Information Technology, Hyderabad 500032, Andhra Pradesh, India
e-mail: yegna@iiit.ac.in

use only the spatial information of an edge map without considering the inherent local structural characteristics inside the edge map [13]. The problem of matching in continuous edge gradient representation is addressed to some extent by using a potential field derived from the continuous edge gradient [14].

One can avoid the issues of computation and matching of edge maps by representing the edge map in the Fourier domain. Phase of the Fourier transform corresponds to the relative locations of events such as lines and edges [17,18]. But computation of the phase of the Fourier transform using arctan leads to the problem of phase wrapping [17]. This issue was addressed in the literature using phase unwrapping and group-delay processing [17,19]. Use of the phase of the Fourier transform of a given face image for face recognition was studied in [20], where the phase was computed directly using arctan function. In this paper, we propose a method which does not use the phase of the Fourier transform directly, but uses the information present in the phase effectively.

The organization of the paper is as follows: The proposed representation of the face image is described in Sect. 2. Eigen-analysis on the proposed representation is discussed in Sect. 3 to derive a compressed representation of the phase information. Section 4 gives the results of experimental studies. A summary of the work and conclusion are given in Sect. 5.

2 Phase of Fourier transform

The most common way of representing an image in the spatial domain is by a two-dimensional array of positive numbers, corresponding to the gray levels of the pixels. An image can also be represented in the frequency domain as the discrete Fourier transform (FT) of the two-dimensional array of pixels [21]. The Fourier representation involves complex numbers, i.e., the magnitude and phase parts. The relative importance of the magnitude and phase of the FT of a signal/image under different situations was studied in [17,18]. It is difficult to visualize how the information in these two components are related, because the magnitude and phase are not directly comparable.

Let us represent an image by $x[n_1, n_2]$, $n_1 = 0, 1, \ldots, R-1$, $n_2 = 0, 1, \ldots, C-1$. Here R and C are the number of rows and columns of the given image, respectively. The discrete Fourier transform (DFT) [21] of $x[n_1, n_2]$ is given by

$$\begin{aligned} X[k_1, k_2] &= \mathrm{DFT}\,\{x[n_1, n_2]\} \\ &= X_r[k_1, k_2] + X_i[k_1, k_2] \\ &= |X[k_1, k_2]| \exp[\theta[k_1, k_2]], \end{aligned} \tag{1}$$

where $|X[k_1, k_2]| = \sqrt{(X_r[k_1, k_2])^2 + (X_i[k_1, k_2])^2}$, and $\theta[k_1, k_2] = \arctan\left\{\frac{X_i[k_1,k_2]}{X_r[k_1,k_2]}\right\}$ are the magnitude and the phase of the DFT, respectively. The real and imaginary parts of the DFT are denoted by X_r and X_i, respectively. The original image can be obtained from $X[k_1, k_2]$ by inverse DFT relation [21] which is abbreviated as IDFT. The information contained in the magnitude and phase of the DFT can be visualized using magnitude-only synthesis of the face image $x_m[n_1, n_2] = \mathrm{IDFT}\{|X[k_1, k_2]|\}$ and the phase-only synthesis of the face image $x_p[n_1, n_2] = \mathrm{IDFT}\{\exp[j\theta(k_1, k_2]]\}$, respectively. The images $x_m[n_1, n_2]$ and $x_p[n_1, n_2]$ are shown in Fig. 1. The phase-only face image retains more edge information (crucial features) of the original face image as compared to the magnitude-only face image.

One of the properties of edges is that they are less sensitive to illumination. Hence the phase of the Fourier transform of a given image can be used to address the issue of illumination in face recognition. But computation of the phase spectrum using arctan leads to the problem of phase wrapping [21]. One way to address this issue is to use a function of the phase spectrum, instead of the phase spectrum directly. We can write

$$\begin{aligned} \exp[j\theta[k_1, k_2]] &= \cos[\theta[k_1, k_2]] + j\sin[\theta[k_1, k_2]] \\ &= \frac{X[k_1, k_2]}{|X[k_1, k_2]|}, \end{aligned} \tag{2}$$

where,

$$\cos[\theta[k_1, k_2]] = X^c[k_1, k_2] = \frac{X_r[k_1, k_2]}{|X[k_1, k_2]|},$$

and

$$\sin[\theta[k_1, k_2]] = X^s[k_1, k_2] = \frac{X_i[k_1, k_2]}{|X[k_1, k_2]|}. \tag{3}$$

We can use the cosine and sine functions of the phase to avoid the phase wrapping problem. These components may contain some complementary information of the face image, as shown quantitatively in the experiments section. They are used separately in template matching based method for face recognition.

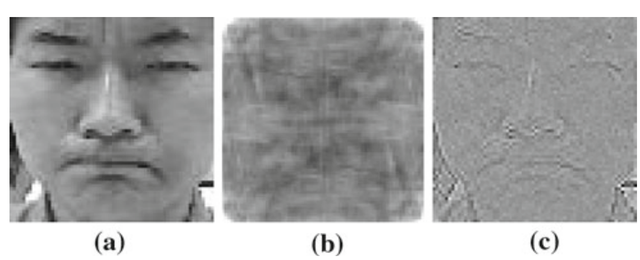

Fig. 1 **a** Gray-level face image. **b** Magnitude-only synthesis of face image. **c** Phase-only synthesis of face image

3 Eigenanalysis using Fourier phase

The cosine and sine functions of the phase spectrum accentuate the high frequency components. Hence they emphasize noise also. The effect of noise can be reduced using eigenanalysis [22]. Let the training face images for the person i be denoted by set D_i. The cosine and sine functions of the phase spectrum are computed using (3). The DFT of a real image exhibits conjugate symmetry [21]. Hence only the non-redundant coefficients (the shaded region in Fig. 2) of the cosine and sine functions of the phase spectrum are used in the eigenanalysis. Let $\mathbf{\Psi}^c \in \mathbb{R}^{N\times m} = [\psi_1^c, \ldots, \psi_m^c]$ and $\mathbf{\Psi}^s \in \mathbb{R}^{N\times m} = [\psi_1^s, \ldots, \psi_m^s]$ be the eigenvectors corresponding to the m largest eigenvalues derived using the covariance matrices of $\mathbf{X}^c$ and $\mathbf{X}^s$ representations of the given training face image, respectively. Here $N = \frac{RC}{2} + 2$. The $\mathbf{X}^c$ and $\mathbf{X}^s$ are the vector representations of the non-redundant coefficients of the cosine and sine functions of the phase spectrum of the given face image $x[n_1, n_2]$, respectively. The eigenvectors are used to represent the face image approximately as follows:

$$\begin{aligned}\mathbf{a}_{\mathbf{x}}^{c,m} &= (\mathbf{\Psi}^c)^t\mathbf{X}^c\\ \mathbf{a}_{\mathbf{x}}^{s,m} &= (\mathbf{\Psi}^s)^t\mathbf{X}^s.\end{aligned} \tag{4}$$

These new representations are used for matching in a face recognition task. The effect of noise is reduced as only the first m ($m \le N$) coefficients are considered for matching. The proposed representations ($\mathbf{X}^s$ and $\mathbf{X}^c$) have another advantage in the context of eigenanalysis, as the size of of resulting covariance matrix ($(\frac{RC}{2}+2) \times (\frac{RC}{2}+2)$) is approximately one fourth as compared to the covariance matrix $(RC \times RC)$ obtained using gray level values of the face image. Thus, the estimation of the covariance matrix (eigenvectors) may be more accurate for same number of training face images.

Let $d_{i,\mathbf{y}}^{c,m}$ denote the minimum Euclidean distance obtained for a given test face image $y[n_1, n_2]$ using the cosine function of the phase spectrum of the available training face images of the i^{th} person. That is

$$d_{i,\mathbf{y}}^{c,m} = \min_{\mathbf{x}\in D_i} ||\mathbf{a}_{\mathbf{y}}^{c,m} - \mathbf{a}_{\mathbf{x}}^{c,m}||_2. \tag{5}$$

Similarly, the minimum Euclidean distance is computed using the sine function of the phase spectrum of the test and training face images, and is denoted by $d_{i,\mathbf{y}}^{s,m}$. The identity (i^*) of a given face image is obtained using the combined Euclidean distance as follows:

$$i^* = \arg\min_i \left[(d_{i,\mathbf{y}}^{s,m})^2 + (d_{i,\mathbf{y}}^{c,m})^2\right]^{1/2}. \tag{6}$$

Performance (η) is computed as

$$\eta = \frac{\text{Number of correctly identified face images}}{\text{Total number of available test face images}} \times 100. \tag{7}$$

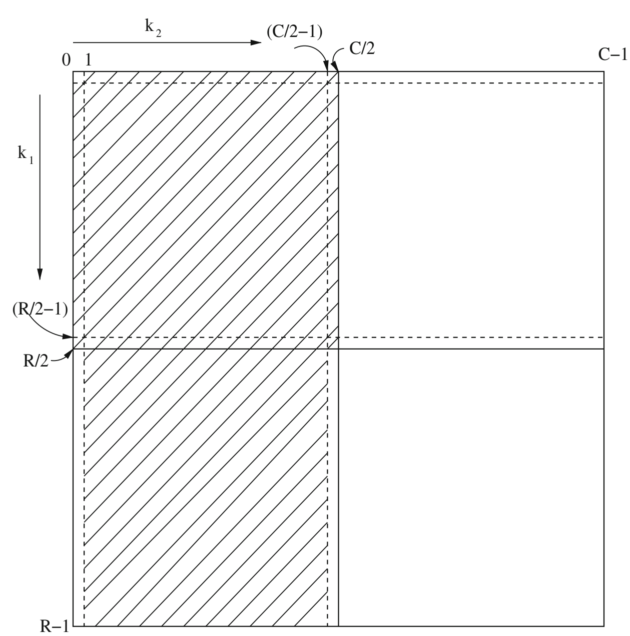

Fig. 2 DFT coefficients $X[k_1, k_2]$ in the shaded area determine the remaining coefficients

3.1 Significance of DFT coefficients

The spacing of the edges will be inversely proportional to the frequency in the phase of the Fourier transform. Thus the low frequency DFT coefficients correspond to events/edges separated by large spacing, and the high frequency DFT coefficients for events/edges separated by small spacing. The effect of the different DFT coefficients can be seen in the phase-only synthesis of the face image, by making the first l DFT coefficients zero along both the axes in the frequency domain, and preserving the remaining DFT coefficients. For different choices of l, the phase-only synthesis of face images in Fig. 3 show the effect of the DFT coefficients on the spacing of edges in the images.

Matching true class face images having some variation with respect to training images can be improved by removing some high frequency DFT coefficients. This is an advantage because noise and events with small spacing are given less importance. Experiments were conducted by considering only the first k DFT coefficients along both the axes of the $X^c[k_1, k_2]$ and $X^s[k_1, k_2]$ representations of the given training face image. Only non-redundant coefficients are used for eigenanalysis. The recognition performance is improved by removing some high frequency DFT coefficients. But removing many DFT coefficients results in loss of information in the face image, and hence the performance also degrades.

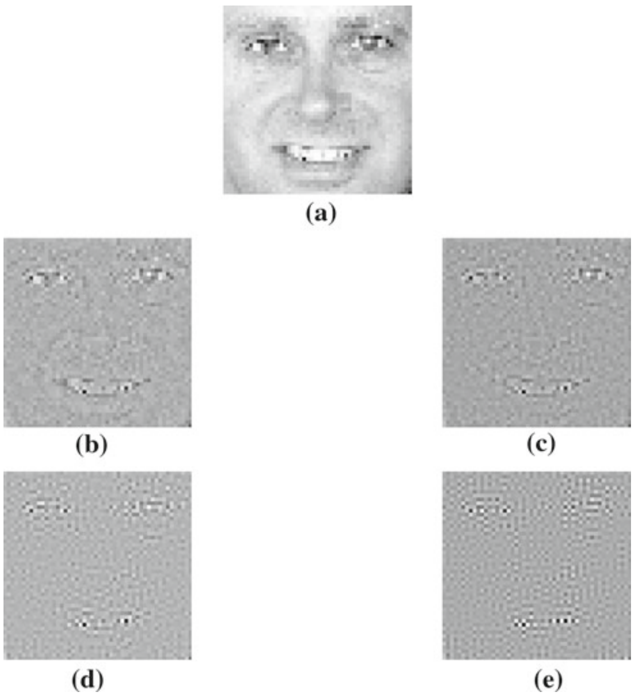

Fig. 3 **a** Gray-level face image. Phase-only synthesis face image for **b** $l = 5$, **c** $l = 10$, **d** $l = 15$, and **e** $l = 20$

3.2 Padding with zeros in computation of the phase of the Fourier transform

One of the properties of DFT is that zero padding in one domain results in increase in the number of samples in the other domain [21]. Hence zero padding in the spatial domain results in an increase in the number of samples in the frequency domain. The most common form of zero padding is to append strings of zeros at the end of the 2-D sequence. After appending, the new 2-D sequence is of size $fR \times fC$, where f is a scaling parameter, and R and C are the number of rows and columns in the original face image, respectively. The resulting 2-D sequences of the face images are used in face recognition. For high values of f the finer resolution of the phase spectrum is captured, and this may help in discrimination of faces.

4 Experiments

Performance of the proposed representation is evaluated on three face databases, namely, illumination variation set of FacePix database [23,24], PIE-NL data set (subset of PIE database) [25], and Yale-B database (http://cvc.yale.edu/projects/yalefacesB.html). There are face images of 30 persons in the FacePix database. The illumination set in this database was captured with the person looking directly into the camera, while the light source was moved around the subject. The light source was moved at 1° interval from −90° to 90°. These images are denoted by $L^1, \ldots, L^{181}$. The size of the face images are rescaled to 50×50 in our experiments. We have taken one face image (L^{91}) of each person for training, and the remaining face images are used for testing. Thus 30 training face images are used to compute the eigenvectors. The performance (η) is obtained using the combined Euclidean distance for different values of m as shown in Fig. 4. Note that the performance is shown by using up to 30 eigenvectors, as there are only 30 training face images. The performance is also shown for the cosine and sine functions of the phase spectrum separately. The two functions of the phase spectrum seem to contain some complementary information, as the performance improves by combining the evidences from the two functions. In general the performance increases with m, but after some value of m the performance reaches a maximum value. The performance can be improved further by padding the image with zeros before computing the DFT, and by removing some high frequency DFT coefficients. In fact the performance (η) improves from 60 to 72.6% by using $f = 2$ for the zero padding parameter, and $k = 20$ for choosing the number of DFT coefficients. We have repeated the experiments with different sets of the training face images. Performance comparison with other methods [23,24] is given in Table 1. The results show that the proposed method performs better than the existing methods. For comparison, the performance using the edginess-based method [26] is also given. In the case of one training face image per person (second column of Table 1), the performance of the edginess-based representation is better than other representations. This is because the edginess-based representation is independent of the number of training images, whereas in the FT phase representation the smearing of the edges is realized using the first few eigenvectors, which are obtained using eigenanalysis on only a small number (30) of the training face images.

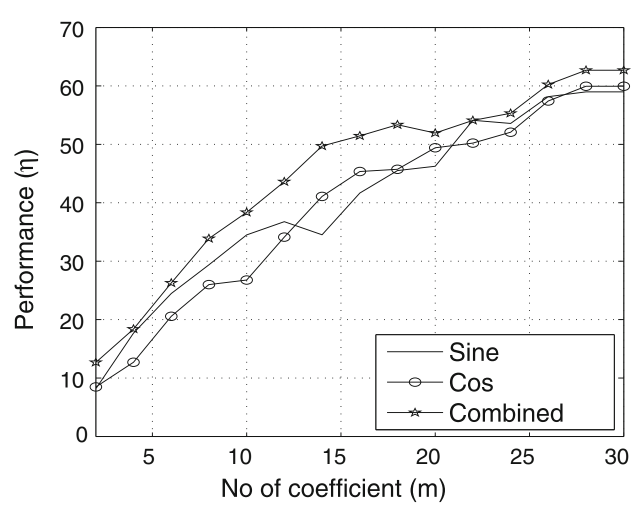

Fig. 4 Performance with one training face image

Table 1 Average recognition rate (in %) for different sets of reference face images under illumination variation of FacePix database

Methods	Set of reference face images		
	L^{91}	L^{1}, L^{91} and L^{181}	L^{1}, L^{46}, L^{91}, L^{136} and L^{181}
Principal component analysis	48.84	71.71	90.33
Linear discriminant analysis	53.04	79.52	94.92
Hidden Markov model	19.26	37.38	59.37
Bayesian information criteria	49.80	79.10	93.54
Edginess-based method	81.43	94.32	99.72
Proposed	72.6	97.1	99.33

When the number of the training images are increased, the performance of the FT phase representation has improved significantly compared to the edginess-based representation, as can be seen from the results in third and fourth columns of Table 1, where the number of training images are 90 and 150, respectively.

The PIE-NL data set contains 65 persons, each having 21 face images. These face images are acquired under different illumination conditions by controlling 21 flashes. These face images are denoted by $I^1, I^2, \ldots, I^{21}$. Table 2 compares the performance of the proposed method with the method discussed in [20] for different sets of training face images. The method given in [20] uses the phase of the FT computed using arctan in the eigenanalysis. The results in the table show that the proposed method uses the phase information effectively, and avoids the phase unwrapping as well.

The Yale-B face database contains face images of 10 persons, each having 576 viewing conditions (nine different poses and 64 different lighting conditions from negative azimuth to positive azimuth). Two sets of training data are formed: Set 1 containing $35 \times 10 = 350$ images of negative azimuth (with 35 images for each person), and set 2 containing $29 \times 10 = 290$ images of positive azimuth (with 29 images for each person). The eigenvectors are derived using the data set 1, and the performance is evaluated using the data set 2. The performance is also obtained by interchanging the training and test data sets. The resulting performance using the proposed method is given in Table 3, along with other existing methods [27]. One can observe from Tables 1, 2, and 3 that the phase of the Fourier transform is robust to illumination variation in face recognition.

Table 2 Recognition rate (in %) using PIE-NL data set. Here I^7, I^{10}, and I^{19} are face images with frontal lighting. I^3 and I^{16} are face images with left shadow and right shadow, respectively

Methods	Set of training face images		
	I^7	I^7, I^{10} and I^{19}	I^3, I^{16} and I^7
Eigenphase [20]	–	97	100
Proposed	97.73	99.05	100

Table 3 Average recognition rate (in %) for different sets of reference face images under illumination variation of Yale-B face database

Methods	Set of reference face images	
	Set 1	Set 2
(PCA) Principal component analysis	79.31	80.29
(2DPCA) 2-D principal component analysis	82.76	86.88
(BDPCA) Bidirectional PCA	82.76	88.57
(W-BDPCA) Whitened BDPCA	87.59	91.14
(ICA) Independent component analysis	82.41	84.57
(EICA) Enhance component analysis	86.90	85.14
(RC-ICA) Row column independent component analysis	91.38	92.57
Proposed	93.6	96

5 Summary

This paper highlights the significance of the FT phase for face recognition using template matching. The FT phase preserves the edge information. An advantage of this representation is that it does not require any thresholding for computation of the edge maps. Moreover, the representation is better suited for matching the edge information of two face images, compared to other representations. The proposed method avoids computation of unwrapped phase. Noise in the function of the phase spectrum is reduced using eigenanalysis. The importance of edges in matching two face images is controlled by selecting a subset of the Fourier coefficients for the eigenanalysis. Experimental results show that the proposed method is less sensitive to illumination. The results are also comparatively better than some of the earlier methods proposed using the phase information of an image.

It is obvious that in template matching method matching edge information is not possible if there is variation in pose and expression in the image. A simple way of matching with these variations is by smearing the edges, in which case the discriminatory information itself may be lost. It is necessary to incorporate the person-specific information in the FT phase, and highlight that information while matching. It is also worthwhile to augment the existing methods of face representation with the FT phase representation in order to improve the performance, as the FT phase may provide some complimentary information of the face image.

References

1. Chellappa, R., Wilson, C., Sirohey, S.: Human and machine recognition of faces. A Survey. **83**, 705–740 (1995)
2. Tolba, A., El-Baz, A.H., El-Harby, A.A.: Face recognition: a literature review. Int. J. Signal Process. **2**, 88–103 (2005)
3. Daugman, J.: How iris recognition works. In: Proc. IEEE Conf. ICIP (2002)
4. Maltoni, D., Jain, A.K., Prabhakar, S., Maio, D.: In: Handbook of Fingerprint Recognition. Springer, New York (2003)
5. Yegnanarayana, B., Prasanna, S.R.M., Zachariah, J.M., Gupta, C.S.: Combining evidence from source, suprasegmental and spectral features for a fixed text speaker verification system. IEEE Trans. Acoust. Speech Signal Process. **13**, 575–582 (2005)
6. Jain, A.K.: Biometric recognition. Nature **449**, 38–40 (2007)
7. Deniz, O., Castrilon, M., Hernandwez, M.: Face recognition using independent component analysis and support vector machine. Pattern Recognit. Lett. **22**, 2153–2157 (2003)
8. Gross, R., Matthews, I., Baker, S.: Appearance-based face recognition and light-fields. IEEE Trans. Pattern Anal. Mach. Intell. **26**, 449–465 (2004)
9. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.J.: Face recognition using Laplacianface. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 328–340 (2005)
10. 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)
11. Jing, X.Y., Tang, Y.Y., Zhang, D.: A Fourier-LDA approach for image recognition. Pattern Recognit. **38**, 453–457 (2005)
12. Gundimada, S., Asari, V.K.: Facial recognition using multisensors images based on localized kernel eigen spaces. IEEE Trans. Image Process. **18**, 1314–1325 (2009)
13. Gao, Y., Leung, M.K.H.: Face recognition using line edge map. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 765–779 (2002)
14. Sao, A.K., Yegnanarayana, B., Kumar, B.V.K.V.: Significance of image representation for face recognition. In: Signal Image and Video Processing, pp. 225–237 (2007)
15. Paglieroni, D.W.: Distance transforms: Properties and machine vision applications. Proc. Comput. Vision Graphics Image Process. **54**, 56–74 (1992)
16. Takacs, B.: Comparing face images using modified Hausdorff distance. Pattern Recognit. **31**, 1873–1881 (1998)
17. Yegnanarayana, B., Saikia, D., Krishnan, T.: Significance of group delay functions in signal reconstruction from spectral magnitude or phase. IEEE Trans. Acoust. Speech Signal Process. **ASSP-32**, 610–623 (1984)
18. Oppenheim, A.V., Lim, I.S.: The importance of phase in signals. Proc. IEEE **69**, 529–562 (1981)
19. Tribolet, J.M.: A new phase unwrapping algorithm. IEEE Trans. Acoust., Speech, Signal Processing, pp. 170–177 (1977)
20. Savvides, M., Kumar, B.V.K.V., Khosla, P.K.: Eigenphases versus Eigenfaces. In: ICPR (3), pp. 810–813 (2004)
21. Oppenheim, A.V., Schafer, R.W.: Discrete-time Signal Processing. Prentice Hall of India, New Delhi (1997)
22. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cognitive Neurosci. **3**, 71–86 (1991)
23. Black, J., Gargesha, M., Kahol, K., Kuchi, P., Panchanathan, S.: A framework for performance evaluation of face recognition algorithm. In: Internet Multimedia System, Boston (2002)
24. Little, G., Krisgna, S., Black, J., Panchanathan, S.: A methodology for evaluating robustness of face recognition algorithms with respect to change in pose and illumination angle. In: Proc. Int. Conf. Acoustics, Speech and Signal Processing, Philadelhia (2005)
25. Sim, T., Baker, S., Bsat, M.: The CMU pose illumination, and expression PIE database of human faces. Tech. Report CMU-RI-TR-01-02, Robotics Institute, Carnegie Mellon University (2001)
26. Sao, A.K., Yegnanarayana, B.: Face verification using template matching. IEEE Trans. Inf. Forensic Security **2**, 636–641 (2007)
27. Gao, Q., Zhang, L., Zhang, D.: Sequential row column independent component analysis for face recognition. Neurocomputing **72**, 1152–1159 (2009)