Eigenphase-based face recognition: a comparison of phase-information extraction methods

Slobodan Ribarić, Marijo Maračić

Faculty of Electrical Engineering and Computing,
University of Zagreb, Unska 3, 10 000 Zagreb
E-mail: slobodan.ribaric@fer.hr

Abstract—*Phase information has recently been used quite frequently in the field of face recognition, especially when the facial images are taken under variable illumination conditions. The approach which combines a PCA with the phase information in order to extract the facial features and reduce the dimensionality of the feature space is called eigenphase method. This paper focuses on four modifications to the eigenphase method. The methods differ in terms of their approach to obtaining the phase information from the facial images. The modifications are evaluated on the XM2VTS, Yale and ORL datasets in order to examine their robustness, primarily under variable and normalized illumination conditions and for slightly variable head poses and facial expressions. The results of the experiments are presented and discussed.*

1 Introduction

Face recognition is a rapidly developing area of research. There are already many methods [1] that can produce relatively high face-recognition rates (above 95%) when applied to images obtained under stable illumination conditions. Illumination, pose and the expression on the human face are unconstrained in real-world environments, which tends to reduce face-recognition rates. There are, however, many methods that attempt to solve the problem of variable illumination in the images [2], [3]. One such method is based on eigenphases [4]. In general, phase information contains most of the information useful for image analysis, intelligibility, reconstruction and recognition [5]. The eigenphase method uses the phase information combined with a Principal Component Analysis (PCA) [6] to extract image features that may result in successful face recognition. Different approaches can be used to extract phase information from an image [5], [7], [8].

Oppenheim and Lim [5] emphasized the importance of phase information over magnitude information for signal intelligibility when a signal is expressed in the frequency domain. The effect was demonstrated by showing that the reconstructed images are still intelligible in the image domain if their magnitude spectrum in the frequency domain is "removed". Two ideas about how to "remove" the magnitude spectrum in the frequency domain were presented. The first idea suggests that the magnitude should be set to unity throughout the whole of the spectrum. The second idea is to set the magnitude throughout the spectrum to the average values of the magnitude spectrum over an arbitrary image set.

The eigenphase method was described by Savvides et al. [4]. The method supplements the idea presented in [5], that the image intelligibility is mainly related to the phase information by establishing that the illumination of the image mostly influences the magnitude spectrum, but not the phase spectrum of the image. The eigenphase method was applied to the face-recognition problem using the CMU PIE face dataset, which contains image sets that vary only in terms of the illumination conditions. The facial images were transformed to the frequency domain and the PCA was then applied to the phase spectra of the training images. It was demonstrated that the PCA, by itself, has the same effect on the recognition rates, regardless of the domain in which it is applied (the frequency or image domain). The recognition rates of the eigenphase method were higher than the rates of the LDA, PCA and 3D Linear Subspace methods in all the experiments. The authors concluded that the higher recognition rates of the eigenphase method were due to the phase spectrum's invariance to the illumination conditions and partial occlusions.

Zaeri et al. showed that the eigenphase method can be efficiently implemented in devices with limited memory [7], [8]. Their method combined MPEG-7 Fourier Feature Descriptor (FFD) vectors with the eigenphase method and phase-information binarization producing an efficient face-recognition method with a high recognition rate. The method was tested on the XM2VTS and ORL datasets and the reported recognition rates were better than those achieved using PCA in the image domain, eigenphase and MPEG-7 methods.

This paper presents four different phase-information extraction approaches and gives the results of preliminary face-recognition experiments based on modifications to the eigenphase method.

2 Eigenphase-based face recognition

The eigenphase method transforms the facial images into the frequency domain, where the PCA is applied to the phase information to extract the features for recognition.

The eigenphase-based face-recognition process can be divided into the following steps: image preprocessing, image transformation to the frequency domain, phase-feature extraction and classification.

2.1 Image preprocessing and transformation to the frequency domain

The facial images are converted to grayscale, resized to 64 x 64 pixels and converted into vectors by scanning them row by row. Each sample vector consisting of 1 x Z components (Z = 4096) is denoted as $\mathbf{x}_m$, m = 1, …, N, where N is the total number of samples (i.e., facial images) in the dataset.

The samples are then transferred to the frequency domain by:

$$\hat{\mathbf{f}}_m = \hat{\mathbf{F}}_{DFT} \cdot \mathbf{x}_m \tag{1}$$

where $\hat{\mathbf{f}}_m$ is a sample vector in the frequency domain, m = 1, 2, …, N and $\hat{\mathbf{F}}_{DFT}$ is a Fourier-transform matrix.

2.2 Phase-feature extraction and classification

The phase features are extracted from the phase information by using the PCA. The phase information can be obtained from sample vectors transformed into the frequency domain by using three phase-extraction approaches:

(i) (MagUn approach) Phase information is obtained by setting all the magnitudes of the components $\hat{f}_{mn}$ of the vector $\hat{\mathbf{f}}_m$ to unity [5]:

$$\hat{f}'_{mn} = \frac{\hat{f}_{mn}}{\left|\hat{f}_{mn}\right|} = 1 \cdot e^{j\varphi_{mn}} \tag{2}$$

where m = 1, 2, …, N; n = 1, 2, …, Z and | | denotes the absolute value of $\hat{f}_{mn}$.

(ii) (MagAv approach) Phase information is obtained by setting the magnitudes of the components $\hat{f}_{mn}$ of the complex vector $\hat{\mathbf{f}}_m$ to the average magnitudes $\bar{A}_n$ (n = 1, 2, …, Z) of the corresponding components throughout the arbitrary image set [5]. Usually, this image set is formed from the training subset of the dataset (its cardinality being $N_1 < N$):

$$\hat{f}'_{mn} = \frac{\hat{f}_{mn}}{\left|\hat{f}_{mn}\right|} \cdot \frac{1}{N_1} \sum_{k=1}^{N_1} \left|\hat{f}_{kn}\right| = \bar{A}_n \cdot e^{j\varphi_{mn}} \tag{3}$$

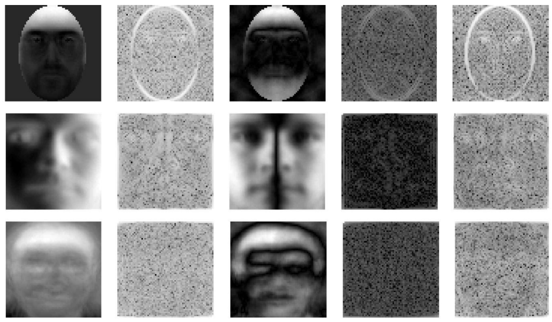
where m = 1, 2, …, N; n = 1, 2, …, Z.

(iii) (PhAngle approach) The phase spectrum is obtained from the components of the complex vector $\hat{\mathbf{f}}_m$ in the form of an angle determined by the well-known atan2 variant of the arctangent function.

$$f'_{mn} = \operatorname{atan2}\left(\operatorname{Im}\left(\hat{f}_{mn}\right), \operatorname{Re}\left(\hat{f}_{mn}\right)\right) \tag{4}$$

where m = 1, 2, …, N and n = 1, 2, …, Z. By $\mathbf{f}'_m$ we denote an angle sample vector with the components f'_{mn}.

Although the form of the phase information is different depending on the chosen approach (for the PhAngle approach the sample vectors are real) the following steps of the method remain the same. The PCA is used to extract the sample features and reduce the dimensionality of the sample space.



a) b) c) d) e)

Fig. 1. The eigenvectors corresponding to the highest eigenvalues. Column a) represents eigenfaces calculated on the normalized XM2VTS, the cropped Yale and the ORL datasets, respectively. Columns b), c), d) and e) represent eigenphases after conversion back to the image domain calculated by using the MagUn, MagAv, PhAngle and CovUn approaches, respectively.

In face recognition the number of training samples N_1 is usually much smaller than the dimensionality of the samples Z, which means that the feature space can have N_1 - 1 dimensions at most. It is simpler to calculate M eigenvectors where $1 \le M \le N_1 - 1$ using the N_1- by-N_1 matrix $\hat{\mathbf{C}}_T$ than using the covariance matrix $\hat{\mathbf{C}}$:

$$\bar{\hat{\mathbf{f}}} = \frac{1}{N_1} \cdot \sum_{p=1}^{N_1} \hat{\mathbf{f}}_p \tag{5}$$

$$\hat{\mathbf{A}} = \left[\hat{\mathbf{f}}_1 - \bar{\hat{\mathbf{f}}}, \hat{\mathbf{f}}_2 - \bar{\hat{\mathbf{f}}}, ..., \hat{\mathbf{f}}_{N1} - \bar{\hat{\mathbf{f}}}\right] \tag{6}$$

$$\hat{\mathbf{C}} = \sum_{p=1}^{N_1} \left(\hat{\mathbf{f}}_p - \bar{\hat{\mathbf{f}}}\right) \cdot \left(\hat{\mathbf{f}}_p - \bar{\hat{\mathbf{f}}}\right)^* = \hat{\mathbf{A}} \cdot \hat{\mathbf{A}}^* \tag{7}$$

$$\hat{\mathbf{C}}_T = \sum_{p=1}^{N_1} \left(\hat{\mathbf{f}}_p - \bar{\hat{\mathbf{f}}}\right)^* \cdot \left(\hat{\mathbf{f}}_p - \bar{\hat{\mathbf{f}}}\right) = \hat{\mathbf{A}}^* \cdot \hat{\mathbf{A}} \tag{8}$$

The eigenvectors of the matrix $\hat{\mathbf{C}}_T$ are denoted as $\hat{\mathbf{u}}_r$, where $1 \le r \le N_1 - 1$. Every eigenvector $\hat{\mathbf{u}}_r$ can be transformed to the corresponding eigenvector $\hat{\mathbf{v}}_q$ of the covariance matrix $\hat{\mathbf{C}}$ according to:

$$\hat{\mathbf{v}}_q = \frac{\hat{\mathbf{A}} \cdot \hat{\mathbf{u}}_r}{\sqrt{\lambda_r}} \tag{9}$$

where q = 1, 2, …, $N_1 - 1$ and $\lambda_r = \lambda_q$ is the corresponding eigenvalue. The eigenvectors $\hat{\mathbf{v}}_q$ (they correspond to the highest eigenvalues; Fig. 1) represent the axes of a new, transformed sample coordinate system in the frequency domain that forms the feature space. Like with the eigenface, each eigenvector $\hat{\mathbf{v}}_q$ can be rearranged into an image whose size is equal to the size of the original image. Such a form of representation for an eigenvector is named the eigenphase.

The projections of the samples into the feature space are calculated as:

$$\hat{\mathbf{v}}_m = \hat{\mathbf{V}}^* \cdot \left(\hat{\mathbf{f}}_m - \bar{\hat{\mathbf{f}}}\right) \quad (10)$$

where m = 1, 2, …, N. $\hat{\mathbf{v}}_m$ is a projected sample, $\hat{\mathbf{V}}^*$ is a conjugate transpose of $\hat{\mathbf{V}}$, where $\hat{\mathbf{V}} = [\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, ..., \hat{\mathbf{v}}_M]$ is a reduced eigenvector matrix.

According to the phase-extraction approach described in [7] it is possible to obtain phase information from the covariance matrix $\hat{\mathbf{C}}$ during the PCA process (CovUn approach). The phase information is extracted by setting the magnitude of the complex components of the covariance matrix in the frequency domain $\hat{\mathbf{C}}$ to unity:

$$\hat{C}'_{kl} = \frac{\hat{C}_{kl}}{\left|\hat{C}_{kl}\right|} = 1 \cdot e^{j\varphi_{kl}} \quad (11)$$

where k = 1, 2, …, Z; l = 1, 2, …, Z. The first $M \leq N_1 - 1$ eigenvectors with the largest associated eigenvalues define the feature subspace.

The 1-NN (nearest-neighbor) rule based on the Euclidean distance is used as the classifier in all the face-recognition experiments.

3 Experimental setups and results

Three different datasets were used to evaluate the performance of the modifications of the eigenphase method based on the phase-information extraction approaches: the XM2VTS database [9], [10] the Extended Yale Face Database B [11] and The ORL Database of Faces [12].

Fig. 2. Dataset examples. First row shows image examples from the normalized XM2VTS dataset; second row, from the cropped Yale dataset; and third row, from the ORL dataset.

A few processed images from each of the datasets mentioned above are depicted in Fig. 2. The performance of the four different phase-information extraction approaches described in Section 2 was tested on these datasets.

To allow a recognition-rate comparison in the experiments, the results of the PCA method in the image domain are shown along with the results of the eigenphase approaches.

3.1 Experiments

The goal of the experiments was to compare the recognition rates of the eigenphase approaches as a function of the feature-space dimensionality. The dimensionality of the feature space was iteratively increased up to 199 dimensions for all four approaches. This upper bound is determined by the training set for the ORL database. Half of each dataset was used as a training set. The results achieved on all the datasets are shown in Figs. 3 - 5.

It is clear from Figs. 3 - 5 that the ratios of the recognition rates of the approaches change depending on the conditions that cause variations in the dataset.

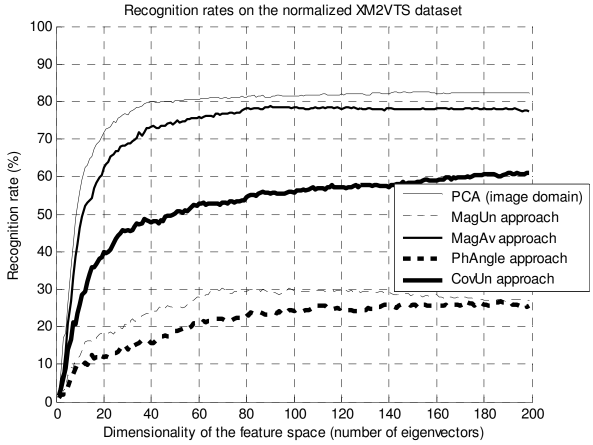


Fig. 3. Recognition rates for the normalized XM2VTS dataset as a function of the feature-space dimensionality.

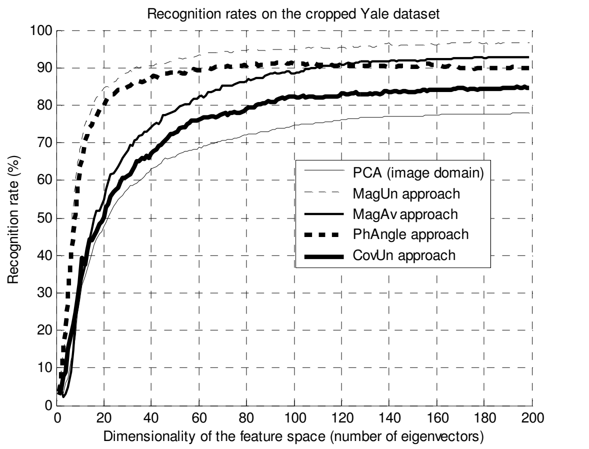


Fig. 4. Recognition rates for the cropped Yale dataset as a function of the feature-space dimensionality.

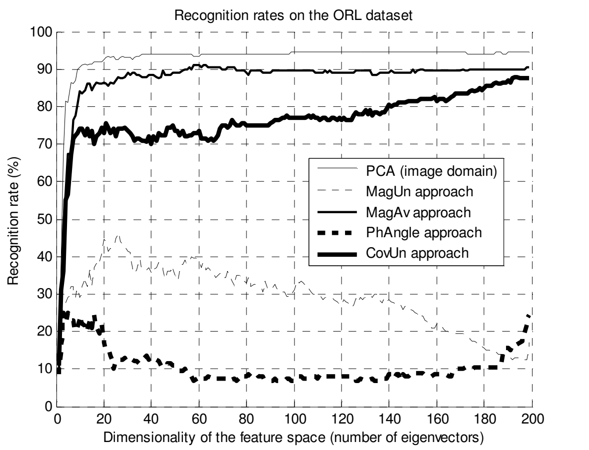


Fig. 5. Recognition rates for the ORL dataset as a function of the feature-space dimensionality.

When variations in the dataset are caused by head positioning or facial expressions, while illumination variations are small (the normalized XM2VTS dataset; Fig. 3), the MagAv approach significantly outperforms all the other eigenphase approaches, with a recognition rate of around 78%. Note that the PCA in the image domain outperforms the MagAv approach. When variations in the dataset are mostly caused by illumination conditions (the cropped Yale dataset; Fig. 4) the MagUn approach achieves the best recognition rate, around 96%, and the MagAv approach achieves around 92%. In the case when the variations in the images originate from all the aforementioned sources (the ORL dataset; Fig. 5) the MagAv approach outperforms the other eigenphase approaches, but its recognition rate is lower than the recognition rate of the PCA in the image domain. The performance of the MagUn and PhAngle approaches is unacceptably low for the datasets whose images exhibit head positioning and facial expression variations (the normalized XM2VTS and ORL datasets). The MagAv approach performs well on all the examined sources of variation. The reason for this is the fact that the average magnitude in the frequency domain falls off at higher frequencies, which happens naturally in typical images [5]. We assume that the average magnitude tends to absorb and average all the variations from the image set, becoming robust to variation effects. The MagUn approach assumes the uniformity of the magnitude spectrum, while the PhAngle approach completely ignores the magnitude spectrum, which suppresses only the effect of illumination and their success is therefore limited to the cropped Yale dataset. It is important to note that in these experiments the variations within the datasets are adequately represented in each of the training sets because each of them contains half of the dataset. Under these conditions the PCA in the image domain performs well. The eigenphase approaches were able to significantly outperform the recognition rate of the PCA in the image domain when the illumination conditions vary significantly (Fig. 4). The recognition rate of the PCA can be improved to around 94% by omitting the first 10 principal components when the illumination conditions vary significantly.

4 Conclusion

Four different approaches to phase-information extraction were presented. Three of them extract the phase information directly from the Fourier transformation of the facial images: the MagUn approach - the magnitude spectrum is set to unity, the MagAv approach - the magnitude spectrum is set to the average magnitude spectrum, and the PhAngle approach - calculates the phase spectrum based on the atan2 function. The fourth approach, called CovUn, uses a covariance matrix of the Fourier transformation of the facial images. The application of the PCA to the phase information obtained with the above approaches is called the eigenphase method. The modifications of the eigenphase method were evaluated on the normalized XM2VTS, the cropped Yale and the ORL datasets. The results were different depending on the source of the variation in the dataset images. The MagAv approach, which combines the average magnitude spectrum with image phase information, resulted in relatively high and stable face-recognition rates under all the presented image variations (variations caused by illumination conditions, head positioning and facial expressions). The recognition rates of the MagUn approach achieved the best recognition rate and demonstrated robustness when the dataset exhibited large illumination variations. The approaches MagUn and PhAngle did not perform well on the datasets where variations of the head pose and facial expressions were present.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld, “Face Recognition: A Literature Survey”, ACM Computer Surveys, Vol. 35, No. 4, December 2003, pp. 399-458.

[2] X. Zou, J. Kittler and K. Messer, “Illumination Invariant Face Recognition: A Survey”, First IEEE International Conference on Biometrics: Theory, Applications, and Systems, September 2007, pp. 1-8.

[3] R. Gopalan, D. Jacobs, “Comparing and combining lighting insensitive approaches for face recognition”, Computer Vision and Image Understanding, Elsevier, Vol. 114, No. 1, January 2010, pp. 135-145.

[4] M. Savvides, B. V. K. Vijaya Kumar and P. K. Khosla, “Eigenphases vs. Eigenfaces”, Proceedings of the 17th International Conference on Pattern Recognition, 2004. Vol. 3, pp. 810-813.

[5] A. V. Oppenheim and J. S. Lim, “The Importance of Phase in Signals”, Proceedings of the IEEE, Vol. 69, No. 5, May 1981, pp. 529-541.

[6] M. A. Turk, A. P. Pentland, “Face recognition using eigenfaces”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91, 3-6 June 1991, pp. 586 -591.

[7] N. Zaeri, F. Mokhtarian, A. Cherri, “Binarized Eigenphases for Limited Memory Face Recognition Applications”, Proceedings of the 5th International Conference on Computer Vision Systems, 2007, http://biecoll.ub.uni-bielefeld.de/volltexte/2007/14/pdf/ ICVS2007-63.pdf.

[8] N. Zaeri, “Eigenphases for Corrupted Face Images”, ACTEA '09. International Conference on Advances in Computational Tools for Engineering Applications, 2009, pp. 537-540.

[9] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre; “XM2VTSbd: The Extended M2VTS Database”, Proceedings 2nd Conference on Audio and Video-base Biometric Personal Verification (AVBPA99)” Springer Verlag, New York, 1999, pp. 72-77.

[10] S. Ribarić, I. Fratrić, K. Kiš; "A Novel Biometric Personal Verification System Based on the Combination of Palmprints and Faces", Informatica, Vol. 19, No. 1, 2008, pp. 81-100.

[11] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, “From few to many: illumination cone models for face recognition under variable lighting and pose”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.23, no.6, pp. 643-660, Jun 2001.

[12] F.S. Samaria, A.C. Harter, “Parameterization of a stochastic model for human face identification”, Proceedings of the Second IEEE Workshop on Applications of Computer Vision, 1994., vol., no., pp. 138-142, 5-7 Dec 1994.