"Eigenphases vs. Eigenfaces"

Marios Savvides, B.V.K. Vijaya Kumar and P.K. Khosla
Electrical & Computer Engineering, Carnegie Mellon University, Pittsburgh PA 15213
msavvid@ri.cmu.edu, kumar@ece.cmu.edu, pkk@ece.cmu.edu

Abstract

In this paper we present a novel method for performing robust illumination-tolerant and partial face recognition that is based on modeling the phase spectrum of face images. We perform principal component analysis in the frequency domain on the phase spectrum of the face images and we show that this improves the recognition performance in the presence of illumination variations dramatically compared to normal eigenface method and other competing face recognition methods such as the illumination subspace method and fisherfaces. We show that this method is robust even when presented with partial views of the test faces, without performing any pre-processing and without needing any a-priori knowledge of the type or part of face that is occluded or missing. We show comparative results using the illumination subset of CMU-PIE database consisting of 65 people showing the performance gain of our proposed method using a variety of training scenarios using as little as three training images per person. We also present partial face recognition results that obtained by synthetically blocking parts of the face of the test faces (even though training was performed on the full face images) showing gain in recognition accuracy of our proposed method.

1. Introduction

There are many challenges to overcome in face recognition[1][2][3][4]; illumination variation[5][6][7][8], facial expression[2], pose[2], occlusion[9]. This paper tries to deal with illumination and occlusion that leads to partial faces being captured using a novel method that takes advantage of frequency representation of face images. We show analytically that performing principal component analysis (PCA)[3] in the frequency domain leads to the same eigenvectors that are obtained as if PCA was performed in the space domain; related via an inverse Fourier transform. However just doing PCA in the frequency domain alone does not produce a gain in recognition performance, rather we show that modeling the complex phase spectrum in the frequency domain produces a face representation that is tolerant to illumination variations, and also can automatically handle occlusions (i.e. missing features) without any special pre-processing. We show empirical results to support this on the CMU PIE database [10] consisting of 65 subjects with 21 different illumination variations per subject. We compare our proposed method to eigenfaces[3], fisherfaces[6] and illumination subspace method[8][6]. We call our method Eigenphases as it models the phase information in the frequency domain representation of the face images.

2. Eigenphase approach

Oppenheim et al [11][12] have shown that phase information of an image retains the most of the intelligibility of an image. Their research also shows that given just the phase spectrum of an image, one can reconstruct the original image up to a scale factor, thus phase information is the most important in representing a 2D signal in the Fourier domain. This is also demonstrated by a simple experiment shown in Figure 1. We have taken two face images; one from person A and one from person B as shown. The Fourier transform of both images were computed, and the respective phase spectrum and magnitude spectrum were extracted. We then synthesized new frequency array using the phase spectrum of person A combined with the magnitude spectrum from person B (the inverse Fourier transform of this array is shown on the bottom row, first column of Figure 1). Similarly we took the phase spectrum from person A and combined it with the magnitude spectrum of
person B (the inverse Fourier transform of this complex array is shown in the bottom row, second column). We observe from the bottom row of Figure 1, that the synthesized face images closely resemble the face image from which the corresponding phase spectrum was extracted from, thus supporting the proposition that phase spectrum contains most of the intelligibility of images.

1.1 Frequency Domain Principal Component Analysis
Since we have established that the complex phase spectrum contains most of the image information, it seems logical to seek to model the image variation by modeling the variation in the phase spectrum of a given sequence of training images. We do this by performing Principal Component Analysis in the frequency domain on the phase spectrums of the training images. So why don’t we do PCA on the whole frequency domain representation? Does doing PCA in the Fourier domain buy us anything?

We show below that indeed the resulting eigenvectors obtained by performing PCA on the Fourier transformed training images are related to the eigenvectors obtained by performing PCA on the original space domain training images via an inverse Fourier transform. This is an important statement as we will show that PCA in the Frequency domain with some simple pre-processing techniques avoids a lot of the problems of space domain PCA.

We will denote $C_s$ and $C_f$ to denote the covariance matrices of the data in the space domain and frequency domain respectively. We demonstrate the proof in 1-D for reasons of simplicity and space but the 2D version is a simple extension. Let $x$ denote the signal of length $d$ samples placed in a column vector, $m$ is the mean vector of all the observed samples $x$ and $DFT$ is the $d \times d$ Fourier transform matrix containing the Fourier basis vectors such that $DFT \cdot x$ is a column vector containing the Fourier transform of $x$.

If we have $N$ training images then the covariance matrix of the Fourier transformed samples is given by:

$$C_f = \sum_{i=1}^{N} \{DFT(x - m)\} \{DFT(x - m)\}^*$$

$$= DFT \cdot X \cdot X^* \cdot DFT^{-1}$$

where $X \cdot X^* = \sum_{i=1}^{N} (x - m)(x - m)^* = C_s$ (2)

PCA diagonalizes the covariance matrix $C_f$ using the orthogonal eigenvectors $v_f$ obtained in Eq. (1).

$$C_f \cdot v_f = \lambda \cdot v_f$$

Substituting $C_f$ for Eq. (1) we get

$$DFT \cdot X \cdot X^* \cdot DFT^{-1} \cdot v_f = \lambda \cdot v_f$$

pre-multiplying Eq. (4) by $DFT^{-1}$ we get

$$XX^* DFT^{-1} v_f = \lambda \cdot DFT^{-1} v_f$$

we now formulate the space domain PCA and noting that the space domain covariance matrix is $C_s=XX^*$

$$C_s \cdot v_s = \lambda \cdot v_s$$

$$XX^* v_s = \lambda \cdot v_s$$

where $v_s$ is the space domain eigenvectors.

Comparing Eq. 8 with Eq. 5 we see by inspection that there is a relation between the space and frequency domain eigenvectors ($v_s$ and $v_f$) related by an inverse Fourier transform as follows:

$$v_s = DFT^{-1} v_f$$

As we have shown (Fig 2) doing PCA in the frequency domain alone does not achieve any advantages, however by eliminating the magnitude spectrum and retaining only the phase spectrum we are able to get a face representation that we will show is tolerant to illumination changes and is robust to missing features. Example eigenphases obtained are shown below in Fig 3.

2. Experiments
To evaluate our proposed method we used the CMU PIE illumination dataset where we used 65 people with 21 different illumination variations captured with no background lighting (example 21 images from a subject are shown in Figure 4). Note that all subjects with the same index image numbers were captured under the same type of illumination conditions as shown in the images of
the subject in Fig 4. To demonstrate the effectiveness of eigenphases we performed three types of experiments as shown in Fig 5. In our experiments we selected faces with different types of illumination and computed the recognition rates on the whole dataset (65 people x 21 faces) using our method and compared the results to that obtained by applying fisherfaces, illumination subspace method (or 3D linear subspace method) and the popular eigenface method.

In the first experiment, we trained on full size images and also tested on full size images, where the training and test images varied in the type of illumination variations. These results are shown in Fig. 6 which shows a plot comparing the recognition rates across 15 different sub-experiments. Each sub-experiment used different types of illumination during training. Table 1 summarizes the types of images used for each sub-experiment. Note however that the sub-experiments are separated into two categories, training images that contained different types of extreme illumination variations and training images captured under near frontal lighting (exhibiting little or none of the extreme lighting). The later scenario is a most probable one in a practical face recognition scenario as we typically will have neutral lighting images of a face, however in the testing phase, the face images might be captured under different and unknown lighting conditions. Fig. 6, shows that the proposed eigenphase method performs very well, clearly outperforming other methods, especially in the scenario where the training images contain near neutral lighting conditions (frontal lighting). In fact, sub-experiment 12, the image indices used for training are 7, 10, 19. We see from Fig 5, that these images are frontal lighting and contain the least amount of lighting variation. In this case the eigenface, illumination subspace method and fisherface algorithm perform poorly with a maximum of 72% recognition rate, while the proposed eigenphases method achieves 97% recognition rates.

Table 1. The Table below shows which type of illumination training images were used in training of the different subspace algorithms for each of the 15 sub-experiments. The training image index corresponding to the type of illumination condition is shown in Fig 1, which is identical across all 65 subjects.

<table>
<thead>
<tr>
<th>Experiment No.</th>
<th>Index # of Training Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>{3, 7, 16}</td>
</tr>
<tr>
<td>2</td>
<td>{1, 10, 16}</td>
</tr>
<tr>
<td>3</td>
<td>{2, 7, 16}</td>
</tr>
<tr>
<td>4</td>
<td>{4, 7, 13}</td>
</tr>
<tr>
<td>5</td>
<td>{1, 2, 7, 16}</td>
</tr>
<tr>
<td>6</td>
<td>{3, 10, 16}</td>
</tr>
<tr>
<td>7</td>
<td>{3, 16, 20}</td>
</tr>
<tr>
<td>8</td>
<td>{5, 6, 7, 8, 9, 10, 18, 19, 20}</td>
</tr>
<tr>
<td>9</td>
<td>{5, 6, 7, 8, 9, 10, 11, 12}</td>
</tr>
<tr>
<td>10</td>
<td>{5, 6, 7, 8, 9, 10}</td>
</tr>
<tr>
<td>11</td>
<td>{5, 7, 9, 10}</td>
</tr>
<tr>
<td>12</td>
<td>{7, 10, 19}</td>
</tr>
<tr>
<td>13</td>
<td>{6, 7, 8}</td>
</tr>
<tr>
<td>14</td>
<td>{8, 9, 10}</td>
</tr>
<tr>
<td>15</td>
<td>{18, 19, 20}</td>
</tr>
</tbody>
</table>

Figure 5. Shows the three types of experiments used where the training images were of full size and the test images were modified as above. (left) Test images are also full size, but variable illumination. (middle) Test images only have the right part of the face, also under variable illumination. (right) Test images only include the eye section shown also under variable illumination.

Figure 6. Recognition Rates comparing the performance using Eigenphases, Fishermfaces, illumination subspace method and eigenfaces on the illumination subset of 65 people of the PIE dataset under variable illumination captured with no background lighting.
In the second experiment we trained the subspace algorithms just as in the first experiment, however in the testing phase we cropped out the test faces as shown in Fig. 5 middle column, to retain only the right half of the faces (note that the test images are partial faces under different lighting illuminations). Fig. 7, shows the resulting recognition rates, where our proposed eigenphase method clearly outperforms the other methods with a minimum margin of recognition accuracy of 23% (sub-experiment 8). The maximum margin of separation is ~72% recorded from sub-experiment 12).

In another experiment we cropped the test images such that only the eye section was visible as shown in the last column of Fig. 5. The results are shown in Fig. 8, where the eigenphase method performs very well with only slight relative degradation in recognition performance. In fact we are still able to achieve recognition rates as high as 99% (with the lowest being 77% for sub-experiment 12).

6. Discussion
We have shown that modeling the variation of the complex phase spectrum of the images in the frequency domain we are able to use a representation that is relatively invariant to illumination variations and is tolerant to occlusions that lead to partial face images. This is important because in the real world, face detectors will detect partial faces either due to occlusion or other factors. We have shown results from three scenarios where we have only illumination variation (Fig. 6), illumination variation and left half of the test faces is missing (Fig. 7), illumination variation and only the eye section of the test faces is visible (Fig. 8). We can clearly see the improvement in recognition accuracy and advantages of our proposed method as it does not require any special preprocessing or any a-priori knowledge to handle illumination variations or occlusions that lead to partial faces during the testing phase. Our method clearly outperforms other face recognition methods such as eigenfaces, illumination subspace (3D linear subspace method) and fisherfaces. Future work includes examining the ability to handle pose.

References

0-7695-2128-2/04 $20.00 (C) 2004 IEEE