We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Frequency Domain Face Recognition

Marios Savvides, Ramamurthy Bhagavatula, Yung-hui Li
and Ramzi Abiantun
Department of Electrical and Computer Engineering, Carnegie Mellon University
United States of America

1. Introduction

In the always expanding field of biometrics the choice of which biometric modality or modalities to use, is a difficult one. While a particular biometric modality might offer superior discriminative properties (or be more stable over a longer period of time) when compared to another modality, the ease of its acquisition might be quite difficult in comparison. As such, the use of the human face as a biometric modality presents the attractive qualities of significant discrimination with the least amount of intrusiveness. In this sense, the majority of biometric systems whose primary modality is the face, emphasize analysis of the spatial representation of the face i.e., the intensity image of the face. While there has been varying and significant levels of performance achieved through the use of spatial 2-D data, there is significant theoretical work and empirical results that support the use of a frequency domain representation, to achieve greater face recognition performance. The use of the Fourier transform allows us to quickly and easily obtain raw frequency data which is significantly more discriminative (after appropriate data manipulation) than the raw spatial data from which it was derived. We can further increase discrimination through additional signal transforms and specific feature extraction algorithms intended for use in the frequency domain, so we can achieve significant improved performance and distortion tolerance compared to that of their spatial domain counterparts.

In this chapter we will review, outline, and present theory and results that elaborate on frequency domain processing and representations for enhanced face recognition. The second section is a brief literature review of various face recognition algorithms. The third section will focus on two points: a review of the commonly used algorithms such as *Principal Component Analysis* (PCA) (Turk and Pentland, 1991) and *Fisher Linear Discriminant Analysis* (FLDA) (Belhumeur et al., 1997) and their novel use in conjunction with frequency domain processed data for enhancing face recognition ability of these algorithms. A comparison of performance with respect to the use of spatial versus processed and un-processed frequency domain data will be presented. The fourth section will be a thorough analysis and derivation of a family of advanced frequency domain matching algorithms collectively known as *Advanced Correlation Filters* (ACFs). It is in this section that the most significant discussion will occur as ACFs represent the latest advances in frequency domain facial recognition algorithms with specifically built-in distortion tolerance. In the fifth section we present results of more recent research done involving ACFs and face recognition. The final

Source: Face Recognition, Book edited by: Kresimir Delac and Mislav Grgic, ISBN 978-3-902613-03-5, pp.558, I-Tech, Vienna, Austria, June 2007

section will be detail conclusions about the current state of face recognition including further future work to pursue for solving the remaining challenges that currently exist.

2. Face Recognition

The use of facial images as a biometric stems naturally from human perception where everyday interaction is often initiated by the visual recognition of a familiar face. The innate ability of humans to discriminate between faces to an amazing degree causes researchers to strive towards building computer automated facial recognition systems that hope to one day autonomously achieve equal recognition performance. The interest and innovation in this area of pattern recognition continues to yield much innovation and garner significant publicity. As a result, face recognition (Chellappa et al., 1995; Zhao et al., 2003) has become one of the most widely researched biometric applications for which numerous algorithms and research work exists to bring the work to a stage where it can be deployed.

Much initial and current research in this field focuses on maximizing separability of facial data through dimensionality reduction. One of the most widely known of such algorithms is that of PCA also commonly referred to as *Eigenfaces* (Turk and Pentland, 1991). The basic algorithm was modified in numerous ways (Grudin, 2000; Chen et al., 2002, Savvides et al., 2004a, 2004b; Bhagavatula & Savvides, 2005b) to further develop the field of face recognition using PCA variants for enhanced dimensionality reduction with greater discrimination. PCA serves as one of the universal benchmark baseline algorithms for face recognition. Another family of dimensionality reduction algorithms is based on LDA (Fisher, 1936). When applied to face recognition, due to the high-dimensionality nature of face data, this approach is often referred to as *Fisherfaces* (Belhumeur et al., 1997). In contrast to *Eigenfaces*, *Fisherfaces* seek to maximize the relative between-class scatter of data samples from different classes while minimizing within-class scatter of data samples from the same class. Numerous reports have exploited this optimization to advance the field of face recognition using LDA (Swets, D.L. & Weng, J., 1996; Etemad & Chellappa, 1996; Zhao et al. 1998, 1999).

Another actively researched approach to face recognition is that of ACFs. Initially applied in the general field of *Automatic Target Recognition* (ATR), ACFs have also been effectively applied and modified for face recognition applications. Despite their capabilities, ACFs are still less well known than the above mentioned algorithms in the field of biometrics. Due to this fact most significant work concerning ACFs and face recognition comes from the contributions of a few groups. Nonetheless, these contributions are numerous and varied ranging from general face recognition (Savvides et al., 2003c, 2004d; Vijaya Kumar et al., 2006) large scale face recognition (Heo et al., 2006; Savvides et al., 2006a, 2006b), illumination tolerant face recognition (Savvides et al., 2003a, 2003b, 2004a, 2004e, 2004f), multi-modal face recognition (Heo et al., 2005), to PDA/cell-phone based face recognition (Ng et al., 2005).

However, regardless of the algorithm, face recognition is often undermined by the caveat of limited scope with regards to recognition accuracy. Although performance may be reported over what is considered a challenging set of data, it does not necessarily imply its applicability to real world situations. The aspect of real world situations that is most often singled out is that of scale and scope. To this end, large scale evaluations of face recognition algorithms are becoming more common as large scale databases are being created to fill this need. One of the first and most prominent of such evaluations is the *Face Recognition Technology* (FERET) database (Phillips et al., 2000) which ran from 1993 to 1997 in an effort to develop face recognition algorithms for use in security, intelligence, and law enforcement.

Following FERET, the *Face Recognition Vendor Test* (FRVT) (Phillips et al., 2003) was created to evaluate commercially available face recognition systems. Since its conception in 2000, FRVT has been repeated and expanded to include academic groups in 2002 and 2006 to continue evaluation of modern face recognition systems. Perhaps the most widely known and largest evaluation as of yet is the *Face Recognition Grand Challenge* (FRGC) (Phillips et al., 2005) in which participants from both industry and academia were asked to develop face recognition algorithms to be evaluated against the largest publicly available database. Such evaluations have served to better simulate the practical real-world operational scenarios of face recognition.

3. Subspace Modelling Methods

Image data, and particularly facial image data is typically represented in a very high dimensional space, thus a significant amount of data needs to be processed requiring significant computation and memory. In this case, we try to reduce the overall dimensionality of the data by projecting it onto a lower dimensional space that still captures most of the variability and discrimination. Several techniques have been proposed for the latter option such PCA, and *Fisher Discriminant Analysis* (FLDA) (Belhumeur et al., 1997).

3.1 Principal Component Analysis

PCA is among the most widely used dimensionality reduction technique. It enables us to extract a lower dimensional subspace that represents the principal directions of variations of the data with controlled loss of information. Also known as the *Karhunen Loeve Transform* (KLT) or *Hotelling Transform*, its application in face recognition is most commonly known as *Eigenfaces*.

The aim of PCA is to find the principal directions of variation within a given set of data. Let $\mathbf{X}$ denote a $d \times N$ matrix containing N data samples of dimension d vectorized along each column. PCA looks for $k < d$ principal components projections such that the projected data $\left\{ y_i = \boldsymbol{\omega}_i^{\mathrm{T}} \mathbf{X} \right\} \in \mathbb{R}^D$ has maximum variance. In other words, we look for the d unit norm direction vectors $\boldsymbol{\omega}_i \in \mathbb{R}^D$ that maximize the variance of the projected data or equivalently best describe the data. These projection vectors form an orthogonal basis that best represent the data in a least-squared error sense. The variance is defined as

$$\begin{aligned} \mathrm{Var}(\mathbf{y}) &= \mathrm{Var}\left(\boldsymbol{\omega}^{\mathrm{T}}\mathbf{x}\right) \\ &= \mathbf{E}\left[\left(\boldsymbol{\omega}^{\mathrm{T}}\mathbf{x} - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\mu}\right)^2\right] \\ &= \boldsymbol{\omega}^{\mathrm{T}}\mathbf{E}\left[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\right]\boldsymbol{\omega} \\ &= \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\omega} \end{aligned} \tag{1}$$

such that $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} = 1$, and $\boldsymbol{\Sigma}$ is defined as

$$\boldsymbol{\Sigma} = E\left[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\right] \tag{2}$$

where $\boldsymbol{\mu} = E[\mathbf{x}]$. We can estimate the covariance matrix $\hat{\boldsymbol{\Sigma}}$ and the mean $\hat{\boldsymbol{\mu}}$ from the N available data samples as

$$\begin{aligned}\hat{\boldsymbol{\Sigma}} &= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^{\mathrm{T}} \\ &= \frac{1}{N}\mathbf{X}\mathbf{X}^{\mathrm{T}}\end{aligned} \tag{3}$$

$$\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \tag{4}$$

where $\mathbf{X}$ now denotes the zero-mean data matrix. To maximize this objective function under the constraint $\|\boldsymbol{\omega}\| = 1$, we utilize the following Lagrangian optimization:

$$L(\boldsymbol{\omega}, \boldsymbol{\lambda}) = \boldsymbol{\omega}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}\boldsymbol{\omega} - \boldsymbol{\lambda}\left(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} - 1\right) \tag{5}$$

To find the extrema we take the derivative with respect to $\boldsymbol{\omega}$ and set the result to zero. Doing so we find that:

$$\hat{\boldsymbol{\Sigma}}\boldsymbol{\omega}_i = \boldsymbol{\lambda}_i\boldsymbol{\omega}_i \tag{6}$$

Premultiplying Eq. (6) by $\boldsymbol{\omega}_i^{\mathrm{T}}$ we get more insight

$$\boldsymbol{\omega}_i^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}\boldsymbol{\omega}_i = \boldsymbol{\lambda}_i\boldsymbol{\omega}_i^{\mathrm{T}}\boldsymbol{\omega}_i \longrightarrow \boldsymbol{\omega}_i^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}\boldsymbol{\omega}_i = Var\{\mathbf{y}_i\} = \boldsymbol{\lambda}_i \tag{7}$$

This corresponds to a standard eigenvalue-eigenvector problem, hence the name *Eigenfaces*.

The directions of variation we are looking for are given by the eigenvectors $\boldsymbol{\omega}_i$ of $\hat{\boldsymbol{\Sigma}}$, and the variances along each direction are given by the corresponding eigenvalues $\boldsymbol{\lambda}_i$ as shown from the above equation. Thus we first choose the eigenvectors (or *Eigenfaces*) with the largest eigenvalues. Moreover, because the covariance matrix is symmetric and positive semi-definite, the eigenvectors produced from Eq. (6) will yield an orthogonal basis. In other words, PCA is essentially a transformation from one coordinate system to a new orthogonal coordinate system which allows us to perform dimensionality reduction and represent the data in the least squared error sense. We apply PCA to face images taken from the Carnegie Mellon University Pose-Illumination-Expression (CMU PIE) No-Light database (Sims et al., 2003) to visualize the resulting *Eigenfaces*. Figure 1 shows the mean image followed by the first 6 dominant *Eigenfaces* computed from this dataset.

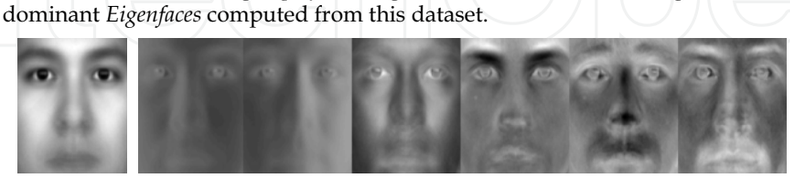


Figure 1. From left to right: PIE No-Light database mean face image followed by the first 6 *Eigenfaces*

3.2 Fisher Linear Discriminant Analysis

Despite its apparent power, PCA has several shortcomings with regards to discriminating between different classes primarily because PCA is optimal for finding projections that are optimal for representation but not necessarily for discrimination.

First developed for taxonomic classifications, LDA (Fisher, 1936) tries to find the optimal set of projection vectors $\boldsymbol{\omega}_i$ that maximize the projected between-class scatter while simultaneously minimizing the projected within-class scatter. This is achieved by maximizing the criterion function equal to the ratio of the determinant of the projected scatter matrices as defined below:

$$J_{\mathrm{FLDA}}(\mathbf{W}) = \frac{\left|\mathbf{W}^{\mathrm{T}}\mathbf{S}_{\mathbf{B}}\mathbf{W}\right|}{\left|\mathbf{W}^{\mathrm{T}}\mathbf{S}_{\mathbf{W}}\mathbf{W}\right|} \tag{8}$$

Where $\mathbf{S}_{\mathbf{B}}$ and $\mathbf{S}_{\mathbf{W}}$ are defined as

$$\mathbf{S}_{\mathbf{B}} = \sum_{i=1}^{c}(\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^{\mathrm{T}} \tag{9}$$

$$\mathbf{S}_{\mathbf{W}} = \sum_{i=1}^{c}\sum_{j=1}^{N_i}\left(\mathbf{x}_j^i - \boldsymbol{\mu}_i\right)\left(\mathbf{x}_j^i - \boldsymbol{\mu}_i\right)^{\mathrm{T}} \tag{10}$$

where N_i, $\boldsymbol{\mu}_i$, and $\boldsymbol{\mu}$ are the number of training images for i^{th} class, the mean of the i^{th} class, and the global mean of all classes respectively. To maximize the Fisher criterion we follow a similar derivation to that of Eq. (5) yielding the following generalized eigenvalue-eigenvector problem:

$$\mathbf{S}_{\mathbf{B}}\boldsymbol{\omega}_i = \boldsymbol{\lambda}\mathbf{S}_{\mathbf{W}}\boldsymbol{\omega}_i \tag{11}$$

whose standard eigenvalue-eigenvector problem equivalent is

$$\mathbf{S}_{\mathbf{W}}^{-1}\mathbf{S}_{\mathbf{B}}\boldsymbol{\omega}_i = \boldsymbol{\lambda}_i\boldsymbol{\omega}_i \tag{12}$$

When applying FLDA to face recognition, the data dimensionality d is typically greater than the total number of data samples N. This situation creates rank deficiency problems in $\mathbf{S}_{\mathbf{w}}$. More specifically, note that $\mathbf{S}_{\mathbf{B}}$, being the sum of c outer product matrices has at most rank $c-1$. Similarly, $\mathbf{S}_{\mathbf{W}}$ is not full rank but of rank $N-c$ at most (when $N \ll d$). To avoid this singularity condition, one can perform PCA on the data to reduce its dimensionality to $N-c$ and then perform FLDA as shown in Eq. (13). The final resulting basis is called *Fisherfaces* (Belhumeur et al., 1997) as given by Eq. (14).

$$\mathbf{W}_{\mathrm{FLDA}} = \arg\max_{\mathbf{W}} \frac{\left|\mathbf{W}^{\mathrm{T}}\mathbf{W}_{\mathrm{PCA}}^{\mathrm{T}}\mathbf{S}_{\mathbf{B}}\mathbf{W}_{\mathrm{PCA}}\mathbf{W}\right|}{\left|\mathbf{W}^{\mathrm{T}}\mathbf{W}_{\mathrm{PCA}}^{\mathrm{T}}\mathbf{S}_{\mathbf{W}}\mathbf{W}_{\mathrm{PCA}}\mathbf{W}\right|} \tag{13}$$

$$\mathbf{W}_{\mathrm{Fisherface}}^{\mathrm{T}} = \mathbf{W}_{\mathrm{FLDA}}^{\mathrm{T}}\mathbf{W}_{\mathrm{PCA}}^{\mathrm{T}} \tag{14}$$

3.3 Frequency Domain Extensions

It has been shown (Oppenheim et al., 1980) that phase information of an image holds the most salient information. In (Hayes et al., 1981), it is shown that one can reconstruct the original signal up to a scale factor given only phase information of the signal. This concept was exploited in face recognition to improve performance over standard algorithms (Savvides et al., 2004b). Figure 2 shows images of two different subjects; each image is split in Fourier domain between magnitude and phase. Figure 2 shows that when the first subject's Fourier magnitude spectrum is coupled with the second subject's Fourier phase spectrum, the resulting image in spatial domain shows significantly more similarity to the second subject compared to the first subject.

(a) (b) (c) (d)

Figure 2. (a) Original image of first subject (b) Original image of second subject (c) Spatial domain image synthesized from combination of Fourier magnitude spectrum of first subject with Fourier phase spectrum of second subject (d) Spatial domain image synthesized from combination of Fourier magnitude spectrum of second subject with Fourier phase spectrum of first subject

However, performing PCA in the frequency domain alone does not constitute any breakthrough, this is because the eigenvectors obtained in the frequency domain are merely the Fourier transform of their spatial domain counterparts. We begin this derivation by defining the standard 2-D *Discrete Fourier Transform* (DFT) pair which is fundamental to the rest of our discussion. Given an 2-D discrete input signal $x[m,n]$ of size $M \times N$ we denote its Fourier transform as $X[k,l]$ whose Fourier transform pair is defined as follows:

$$\begin{aligned} & x[m,n] \underset{F^{-1}}{\overset{F}{\longleftrightarrow}} X[k,l] \\ & X[k,l] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x[m,n] e^{\frac{-i2\pi km}{M}} e^{\frac{-i2\pi ln}{N}} \\ & x[m,n] = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} X[k,l] e^{\frac{i2\pi km}{M}} e^{\frac{i2\pi ln}{N}} \end{aligned} \tag{15}$$

where $i = \sqrt{-1}$, operator F is defined as the forward DFT, and the operator F^{-1} is the inverse DFT.

The estimated covariance matrix of the data in Fourier domain $\hat{\mathbf{\Sigma}}_f$ is given by Eq. (16) where $\mathbf{F}$ is the $d \times d$ Fourier transform matrix containing the DFT basis vectors. The estimated covariance matrix of the data in Fourier domain is given as

$$\begin{aligned}\hat{\mathbf{\Sigma}}_f &= \frac{1}{N}\sum_{i=1}^{N}\left\{\mathbf{F}\left(\mathbf{x}_i-\hat{\boldsymbol{\mu}}\right)\right\}\left\{\mathbf{F}\left(\mathbf{x}_i-\hat{\boldsymbol{\mu}}\right)\right\}^{+} \\ &= \mathbf{F}\hat{\mathbf{\Sigma}}_s\mathbf{F}^{-1}\end{aligned} \tag{16}$$

As was with standard PCA, the eigenvectors $\boldsymbol{\omega}_f$ of $\hat{\mathbf{\Sigma}}_f$ are given by

$$\mathbf{F}\hat{\mathbf{\Sigma}}_s\mathbf{F}^{-1}\boldsymbol{\omega}_f = \boldsymbol{\lambda}\boldsymbol{\omega}_f \tag{17}$$

Premultiplying each side by $\mathbf{F}^{-1}$ we get

$$\hat{\mathbf{\Sigma}}_s\mathbf{F}^{-1}\boldsymbol{\omega}_f = \boldsymbol{\lambda}\mathbf{F}^{-1}\boldsymbol{\omega}_f \tag{18}$$

Comparing Eq. (18) to Eq. (6) we conclude that $\boldsymbol{\omega}_s = \mathbf{F}^{-1}\boldsymbol{\omega}_f$ where $\boldsymbol{\omega}_s$ is an *Eigenface* in spatial domain. We have thus proved that modeling data in the frequency domain does not bring any advantages so far. This fact brings to doubt the usefulness of such a transform with respect to PCA and FLDA without any further processing. However, the ability to distinguish using the magnitude and phase spectrums is the key advantage of the Fourier domain. By modelling the subspace of the phase and magnitude spectrums separately, we can gain further insight and properties of the data otherwise unattainable in the space domain.

3.3.1 Phase Spectrum

It has been shown (Savvides et al., 2004b) that by performing PCA on the phase spectrum alone and disregarding the magnitude spectrum the resulting subspace is more robust with respect to illumination variation. The resulting principal components derived from this new subspace are termed *Eigenphases* in analogy to *Eigenfaces*. It was shown that *Eigenphases* outperform *Eigenfaces* and *Fisherfaces* when trying to recognize not only full faces but also partial or occluded faces as depicted in Figure 4.

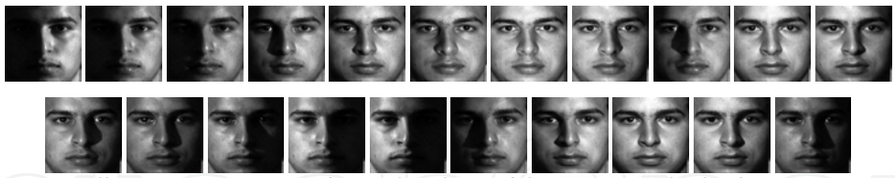
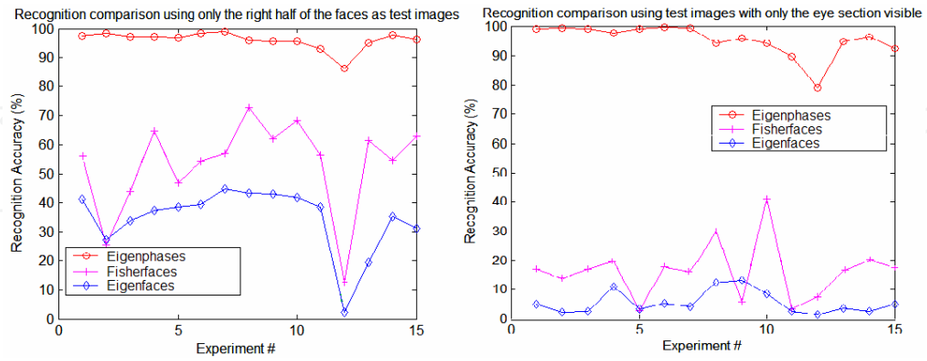


Figure 3. All twenty-one images of a single subject of the PIE No-Light database

(a) (b) (c)

Figure 4. Various occlusions on an example PIE No-Light subject (a) full face (b) right half-face (c) eye section



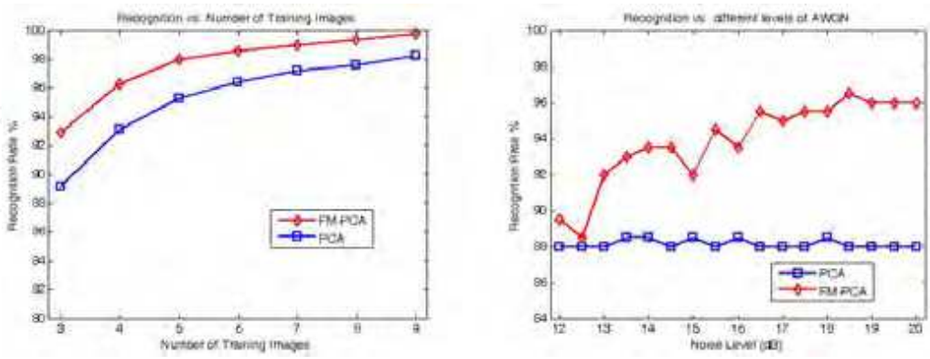
(a) (b)

Figure 5. Rank-1 identification rates obtained by *Eigenphases*, *Eigenfaces*, and *Fisherfaces* for two different experiments each using different types of partial faces. (a) right half face (b) eye-section face

In this work, comparisons between Rank-1 identification rates obtained from *Eigenphases*, *Eigenfaces*, and *Fisherfaces* are made when using whole and partial faces. Training is done on multiple subsets of the PIE database while testing is performed over the whole database. Fifteen different training subsets each representing different types of illumination with the first seven having the most or harshest illumination variation with the remaining eight containing near frontal lighting which are considered the most neutral lighting conditions. Figure 5 depicts the recognition rates obtained with the three different methods using half-faces and eye-sections. These results show that not only do *Eigenphases* outperform *Eigenfaces* and *Fisherfaces* for all experiments by a wide margin, but they also demonstrate minimal performance degradation for half-faces and eye-section faces. This added occlusion robustness is a very attractive property in real-world applications where missing data and poor data quality are common problems.

3.3.2 Magnitude Spectrum

In contrast, if PCA is performed on the magnitude spectrum only, it has been shown (Bhagavatula & Savvides, 2005a) that the resulting subspace holds many advantages over spatial subspaces. Using the Olivetti Research Laboratory (ORL) database, which is noted for significant pose variation, it was shown that the *Fourier Magnitude Principal Component Analysis* (FM-PCA) subspace yielded higher recognition rates across a range of experiments. These experiments included varying the number of training images whose comparison to spatial domain PCA or *Eigenfaces* is illustrated in Figure 6 (a). It was also shown that FM-PCA is more robust to noise as demonstrated in Figure 6 (b). This was verified by corrupting the testing images with varying levels of *Additive White Gaussian Noise* (AWGN). In similar fashion, it was demonstrated that *Fourier Magnititude Fisher Linear Discriminant Analysis* (FM-FLDA) clusters data better than traditional *Fisherfaces* with decreased within-class scatter and increased between-class scatter. FM-FLDA yields higher recognition rates for varying image sizes and resolutions in comparison to spatial FLDA or *Fisherfaces* as tabulated in Table 1.



(a) (b)

Figure 6. Comparisons of identification rates of spatial domain PCA and FM-PCA under varying conditions (a) varying number of training images (b) varying degrees of AWGN noise corrupting the testing images

In addition to increased performance, Fourier Magnitude feature subspaces hold another key advantage. They are shift invariant, as a direct result of the properties of Fourier transform. If the image is shifted in the spatial domain, that shift will translate into a linear-phase change in frequency domain and not in its magnitude. This makes Fourier Magnitude subspaces robust to errors in registration, where the input images are not correctly centred which could cause significant recognition errors. To demonstrate this property, face recognition experiments have been done (Bhagavatula & Savvides, 2005a) by shifting images in both horizontal and vertical directions up to ±5 pixels. These results verify that FM-FLDA and FM-PCA recognition accuracies are not affected, while their spatial domain counterparts are severely affected.

Image Size	32 × 32	64 × 64	112 × 92	128 × 128
FM-Fisher	80.8%	83.2%	84.6%	84.4%
Traditional Fisher	77.7%	78.5%	77.3%	74.0%

Table 1. Recognition accuracies with different image resolutions

4. Advanced Correlation Filters (ACFs)

4.1 Advanced Correlation Filter Basics

The previous sections of this chapter have shown the power of frequency domain representations of data when used in conjunction with techniques and algorithms usually applied to spatial domain representations. However, none of the preceding concepts have been derived from a purely frequency domain approach. By developing algorithms whose focus is on the frequency domain representation of information we can achieve significant gains in performance. One such family of algorithms that have and are still being developed is that of *Correlation Filters* (CFs). CFs have a long and rich history in optics, automatic target recognition, and pattern recognition in general. More recently a new family of CF's termed ACFs (Vijaya Kumar, 1992) have evolved to become the cutting edge of this general family of algorithms. The numerous and varied types of ACFs offer many attractive qualities such as

shift invariance, normalized outputs, and noise tolerance. Their derivations require some knowledge in such fields as linear algebra, signal processing, and detection and estimation theory. We will assume that readers will have sufficient background in these fields and only elucidate on background information when is necessary. We will also now limit our discussion to two-dimensional applications which include facial recognition using grayscale imagery.

To begin the discussion we define a few fundamental terms and conventions that will be used repeatedly for the span of this section. The application of a CF or ACF to an input image will yield a correlation plane. The centre or origin of correlation plane will be considered to be the spatial position (0, 0). Analysis of the correlation plane to some metric of performance or confidence will usually involve calculation and identification of the largest value or peak in the correlation plane.. The simplest CF is the *Matched Filter* (MF), commonly used in applications such as communication channels and radar receivers where the goal is detecting a known signal in additive noise. The concept of noise is a very important aspect of pattern recognition problems. To characterize noise we define the quantitative measure of *Power Spectral Density* (PSD). Using this characterization of noise the MF is developed with the goal of maximizing the *Signal-to-Noise-Ratio* (SNR). Fundamentally this is equivalent to describing a filter whose application to an input signal will minimize the effect of specific type of noise while maximizing the output value when presented with the desired input signal. We will not develop the MF, however multiple other sources provide detailed derivations for varying applications and should be consulted for more information. We will use this fundamental concept of maximizing the response of the desired signal or pattern and minimizing the effects of noise as a guideline in our derivation of ACFs.

One of the fundamental differences between typical CFs and ACFs is the ability to synthesize ACFs from multiple instances of training data or in the case of face recognition, multiple facial images and by doing so, to be able to recognize all instances which are present in the training data. The desire or hope here is that the training data sufficiently represents or captures the potential distortion or variation that might be presented to the recognition system. With respect to face recognition systems this is an extremely desirable quality because the human face is subject to numerous variations both intrinsic and extrinsic. By allowing such variations to be at least partially represented through the use of representative training data we can increase both performance and robustness of face recognition systems.

4.2 Correlation Basics

Before we can derive any ACF we must first lay the framework of correlation with respect to 2D imagery. The standard definition of discrete 2-D correlation between an input 2-D signal $x(m,n)$ and a 2-D filter $h(m,n)$ resulting in 2D correlation output plane $y(m,n)$ is as follows:

$$\begin{aligned} y(m,\,n) &= x(m,\,n) \otimes h(m,\,n) \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x(m+k,\,n+l) h(k,\,l) \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x(k,\,l) h(k-m,\,l-n) \end{aligned} \tag{19}$$

We will only consider the case of discrete correlation as this is the case of interest in face recognition systems although the analog domain provides some desirable qualities and generalizations. However, for our purposes the desired properties of both correlation and the Fourier transform are present in the discrete domain. Using the Fourier transform and its properties as discussed previously we can express Eq. (19) in the frequency domain as

$$\begin{aligned} y(m,n) &= x(m,n) \otimes h(m,n) \\ &= F^{-1}\left\{X(k,l) \cdot H^{*}(k,l)\right\} \end{aligned} \tag{20}$$

where $X(k,l)$ and $H(k,l)$ are the 2-D Fourier transforms of $x(m,n)$ and $h(m,n)$ respectively. The symbols F^{-1}, $\cdot$, and * represent the inverse Fourier transform, the element by element (point to point) multiplication of the two 2-D signals, and the element by element conjugation respectively. Correlation in the frequency domain is vastly preferred to correlation in the spatial domain with regards to the number computational floating point operations required.

4.3 Synthetic Discriminant Functions

One of the first ACFs to incorporate such a composite design is the Synthetic Discriminant Function filter (Hester & Casasent, 1980). The design of the *Synthetic Discriminant Function* (SDF) filter is that the filter is created such that it yields a correlation plane whose output at the origin yields a pre-specified value. By introducing such a constraint on the output we not only allow for normalized comparisons but also a degree of discrimination into our filters. This framework refers to the ability to use a single filter to recognize different patterns or classes with sufficient discrimination as opposed to using a single filter for each class or image sample (as with the case of MFs). For example, in a two class problem we would like to design a filter yields an output value of 1 for class 1 while yielding an output value of 0 for class 2. We can achieve this by constraining the correlation plane outputs (at the origin) to be 1 for all training data from class 1 and 0 for all training data from class 2.

Our derivation of the SDF filter begins with an outline of the basic variables and problem definition. Let us assume that we have N facial training images $x_i(m,n)$ of size $d_1 \times d_2$. Define u_i to be the output value of the correlation plane $y_i(m,n)$; that is the result of applying the filter $h(m,n)$ to the training image $x_i(m,n)$. Please note that the output value of the correlation plane is considered to be the value of the correlation plane at the origin or equivalently $y_i(0,0)$. Thus we can define the following equation,

$$u_i = y_i(0,0) = \sum_{m=1}^{d_1}\sum_{n=1}^{d_2} x_i(m,n)h(m,n), \quad 1 \le i \le N \tag{21}$$

The above equation explicitly demonstrates the correlation operation and the constraint on the correlation plane output value at the origin. However, for convenience we can rewrite the above equation into a more compact vector format. Suppose we take a training image $x_i(m,n)$ (of dimensions d_1 x d_2) and place its entries (vectorize) from left to right and top to bottom into a column vector $\mathbf{x}_i$ of length $d = d_1$ x d_2 and similarly for $h(m,n)$ into column vector $\mathbf{h}$ whose length is also d. We can now express Eq. (21) in the following form,

$$u_i = \mathbf{x}_i^{\mathrm{T}}\mathbf{h}, \quad 1 \le i \le N \tag{22}$$

where $^{\mathrm{T}}$ is the transpose operation. We now have a system of N linear equations which encourages us to express them as the product of a matrix and a vector in order to take advantage of matrix algebra. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ be matrix of size $d \times N$ whose columns are the training image vectors. Likewise, let $\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]^{\mathrm{T}}$ be a column vector of length N whose entries are the desired output values. Now we can express this system of linear equations as the following matrix vector product:

$$\mathbf{u} = \mathbf{X}^{\mathrm{T}}\mathbf{h} \tag{23}$$

A unique solution for **h** can be found by assuming that **h** is a linear combination of the training images, i.e. the columns of **X**. In matrix vector form this can be represented as

$$\mathbf{h} = \mathbf{X}\mathbf{a} \tag{24}$$

where **a** is a column vector of length N whose entries are weightings for the linear combination of the columns of **X**. Substituting Eq. (24) into Eq. (23) we form the following equation:

$$\mathbf{u} = \mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{a} \tag{25}$$

From the above equation we can uniquely determine **a** to equal

$$\mathbf{a} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{u} \tag{26}$$

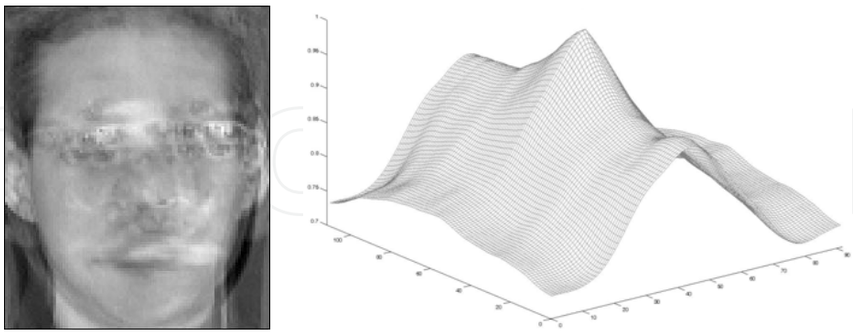
where $^{-1}$ is the standard matrix inverse. Subsequent substitution of the above equation into Eq. (24) yields a solution for the SDF filter **h** which is as follows:

$$\mathbf{h} = \mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{u} \tag{27}$$

Eq. (27) expresses the SDF filter **h** as a column vector of length d in the space domain as opposed to the frequency domain.

We use the SDF filter to demonstrate some key characteristics of correlation in general and also some specific qualities of composite correlation. The images shown in Figure 7 are those of a set of training images taken from the ORL face database. We have used these training images to design an SDF filter whose correlation with any of the training images will yield a correlation plane whose output value, i.e. peak will equal 1. Figure 8 (a) shows the resulting SDF filter point spread function (2D-impulse response) , while Figure 8 (b) demonstrates the result of correlating the filter to one of the training images.

Figure 7. Facial training images taken from single subject in the ORL database



(a) (b)

Figure 8. (a) SDF filter derived from training images in Figure 7 (b) Mesh plot of correlation plane produced from application of SDF filter to one of the training images

As can be seen in these figures, the design of the filter guarantees a correlation plane whose peak equals 1 when applied to one of the training images. We make special note of the fact that we no longer specify the value of 1 to be at the origin but merely be the value of the peak (maximum value in the correlation plane) which corresponds to the location of the detected pattern. This consideration reflects the fact that correlation is a shift-invariant operation assuming the pattern of interest is still completely contained within the input image.

4.4 Minimum Average Correlation Energy Filter

Our discussion and development of the SDF filter has motivated us to address the issue of sidelobes whose presence is significant detriment to performance of any ACF. As such we will now derive the *Minimum Average Correlation Energy* (MACE) filter (Mahalanobis et al., 1987) whose design will not only allow us to achieve constrained peaks as in the SDF filter but also suppress sidelobes in order to yield sharp distinct peaks. This is fundamentally a minimization of the sidelobe heights. One approach is to minimize the correlation plane energy which will subsequently suppress sidelobes. We define the term *Average Correlation Energy* (ACE) for the same N training images in the previous section as

$$\text{ACE} = \frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{d_1}\sum_{n=1}^{d_2}\left|y_i(m,n)\right|^2 \tag{28}$$

where the variables d_1, d_2, and $y_i(m,n)$ retain their definitions from our development of the SDF filter. Eq. (28) can be represented in the frequency domain by applying Parseval's Theorem. Letting $Y_i(k,l)$ be the 2-D Fourier transform of $y_i(m,n)$ we express Eq. (28) as

$$\text{ACE} = \frac{1}{N\cdot d}\sum_{i=1}^{N}\sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\left|Y_i(k,l)\right|^2 \tag{29}$$

where d again is the total dimensionality of a training image. Since $y_i(m,n)$ is the result of the correlation between an input image $x_i(m,n)$ and our MACE filter $h(m,n)$ we can use Eq. (20) to rewrite the above equation into the following form:

$$\text{ACE} = \frac{1}{N \cdot d} \sum_{i=1}^{N} \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} \left| X_i(k,l) \right|^2 \left| H(k,l) \right|^2 \tag{30}$$

It should be noted that it is at this point in the derivation where the role of the frequency domain representations of both the data and the filter are fundamental to the filter design. Later ACF designs will also utilize the quantitative measure of ACE along with other such measures. For now let us to proceed to again represent Eq. (30) in matrix vector form. Let **h** be a column vector of length d whose elements are taken from $H(k,l)$ and $\mathbf{X}_i$ be a diagonal matrix of size $d \times d$ whose non-zero elements are taken from $X_i(k,l)$. Using these frequency domain terms we can express Eq. (30) as

$$\text{ACE} = \frac{1}{N \cdot d} \sum_{i=1}^{N} \left(\mathbf{h}^{+} \mathbf{X}_i \right) \left(\mathbf{X}_i^{*} \mathbf{h} \right) \tag{31}$$

where the symbol $^{+}$ indicates the conjugate transpose. We can compress this expression further by defining a new diagonal matrix **D** of size $d \times d$ as follows:

$$\mathbf{D} = \frac{1}{N \cdot d} \sum_{i=1}^{N} \mathbf{X}_i \mathbf{X}_i^{*} \tag{32}$$

This allows us to express the quantity of ACE in very concise manner as

$$\text{ACE} = \mathbf{h}^{+} \mathbf{D} \mathbf{h} \tag{33}$$

Our goal in the design of the MACE filter is the minimization of the ACE of the training images while still satisfying the peak constraints we have specified. To accomplish this we must express these constraints in the frequency domain as well. Due to the fact that inner products in the frequency domain (at the origin only) are equivalent to inner products in the spatial domain, we can rewrite the peak constraints expressed in Eq. (23) as

$$\mathbf{X}^{+} \mathbf{h} = d \cdot \mathbf{u} \tag{34}$$

where **X** is a matrix of size $d \times N$ whose columns are the vector representations of the FTs of the training images. Thus, the filter **h** which minimizes Eq. (33) while satisfying the constraints expressed in Eq. (34) is our MACE filter. This constrained optimization can be solved using Lagrange multipliers, which can be found in the original paper (Mahalanobis et al., 1987), which yield the final solution to the frequency domain filter **h**:

$$\mathbf{h} = \mathbf{D}^{-1} \mathbf{X} \left(\mathbf{X}^{+} \mathbf{D}^{-1} \mathbf{X} \right)^{-1} \mathbf{u} \tag{35}$$

The notation and form of the solution allows for simple and efficient calculation of the filter in column vector form from which a simple reshaping operation can be done to recover the 2-D frequency domain filter of size $d_1 \times d_2$. Correlation of the filter with an input image now requires one less Fourier transform as the filter is already represented and stored in the

frequency domain. Using the same training images from our derivation of the SDF filter we can create a MACE filter whose output correlation planes will not contain the problematic sidelobes.

Visualizing the point spread function of the MACE filter itself does not reveal much insight without more significant analysis, but the goals of ACE minimization and constrained peaks are achieved as shown in Figure 9. Not only is the peak equal to 1 as specified, but the sidelobes are drastically suppressed when compared to those in the SDF filter's correlation plane in Figure 8 (b). Noise tolerance can be built in as discussed in the next section.

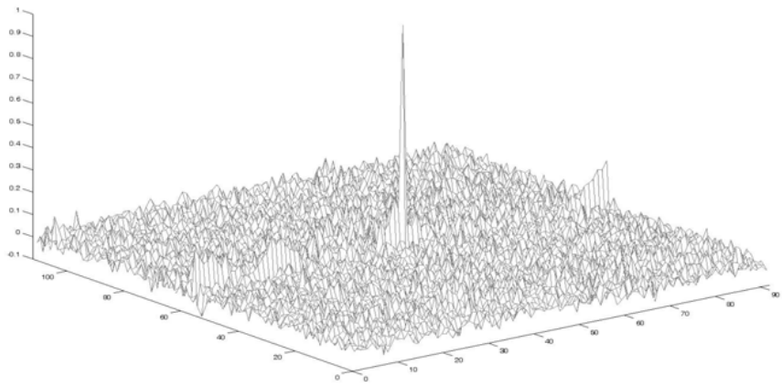


Figure 9. Mesh plot of correlation plane produced from application of MACE filter to one of the training images

4.5 Minimum Variance Synthetic Discriminant Function

Through our derivations of the SDF and MACE filters we have shown that in order to achieve high discriminative ability in our filters we must be able to control the correlation plane through constraints and sidelobe energy minimizations. However, in any practical application we must always take into consideration the factor of noise introduced from varying sources. Whether it is sensor noise or noise caused by background clutter, the presence of noise can have significant impact on any face recognition system. As such we would like to introduce into our ACF designs some degree of noise tolerance. Let us formalize the problem with the following equation:

$$\begin{aligned}(\mathbf{x}+\mathbf{v})^{\mathrm{T}}\mathbf{h} &= \mathbf{x}^{\mathrm{T}}\mathbf{h}+\mathbf{v}^{\mathrm{T}}\mathbf{h} \\ &= u+\delta\end{aligned} \tag{36}$$

where $\mathbf{x}$ is an image vector and $\mathbf{v}$ is the additive noise vector whose responses to the filter vector $\mathbf{h}$ are u and δ respectively. The variations in the outputs of our filter are due to δ and therefore δ is the quantity we wish to suppress. For the rest of the derivation we will assume that our noise processes are stationary. We will also assume that our noise is zero mean without any loss of generality. To suppress the effect of variation in our filter outputs due to noise we aim to minimize the variance of the output noise term δ. Denote this variance as the *Output Noise Variance* (ONV) whose definition is

$$\begin{aligned} \text{ONV} &= E\left\{\delta^2\right\} \\ &= E\left\{\left(\mathbf{v}^{\mathrm{T}}\mathbf{h}\right)^2\right\} \\ &= E\left\{\mathbf{h}^{\mathrm{T}}\mathbf{v}\mathbf{v}^{\mathrm{T}}\mathbf{h}\right\} \\ &= \mathbf{h}^{\mathrm{T}}E\left\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\right\}\mathbf{h} \\ &= \mathbf{h}^{\mathrm{T}}\mathbf{C}\mathbf{h} \end{aligned} \tag{37}$$

where **C** is the covariance matrix of the input noise. We take note of the independence of ONV from the image vector **x** which implies that its definition is identical for all images of interest.

Let us now consider the training images we used in developing the SDF filter whose derivation focused on achieving certain constraints placed on output peak values. We would now like to not only achieve those same constraints expressed in Eq. (23) but also minimize the ONV amongst our training images. This formulation lends itself to the use of Lagrange minimization almost identical to that used in the formulation of the MACE filter to yield the following filter solution:

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{X}\left(\mathbf{X}^{+}\mathbf{C}^{-1}\mathbf{X}\right)^{-1}\mathbf{u} \tag{38}$$

The above filter is referred to as the *Minimum Variance Synthetic Discriminant Function* (MVSDF) filter (Vijaya Kumar, 1986). While the MVSDF filter does achieve minimum ONV amongst its training images, it does not suppress ACE and as such suffers from unsuppressed sidelobes. In later ACF designs we will show how to achieve an optimal tradeoff between ONV and ACE minimization in order to provide varying degrees of simultaneous noise tolerance and sidelobe suppression.

4.6 Maximum Average Correlation Height Filter

All of the ACFs we have described to this point have been designed with some constraint or optimization in mind that is meant to introduce distortion tolerance into our filters. However, this is but one way and perhaps not the best way to create distortion tolerance. There is no formalized relationship between the constraints we have described so far and the degree of distortion tolerance incorporated into the filter. A more intuitive approach is to remove these constraints to allow for more solutions. In essence this is akin to generalizing the solution space which will hopefully contain solutions to non-training images. This would result in a greater degree of distortion tolerance when compared to ACFs derived using hard constraints.

To address the issue of distortion tolerance it is necessary to first quantize the amount of distortion present in a set of filtered images. To this end we define the *Average Similarity Measure* (ASM) over a set of N filtered images $y_i(m,n)$ as

$$\text{ASM} = \frac{1}{N}\sum_{i=1}^{\mathrm{N}}\sum_{m}\sum_{n}\left(y_i(m,n) - \overline{y}(m,n)\right)^2 \tag{39}$$

where we define $\bar{y}(m,n)$ as the average image whose exact definition is

$$\bar{y}(m,n)=\frac{1}{N}\sum_{j=1}^{N}y_j(m,n) \tag{40}$$

ASM is a measure of the average variation amongst a set of correlation surfaces. As was with previous ACFs we recognize the fact that the above spatial domain equation is equivalently expressed in the frequency domain by applying Parseval's theorem. Let $Y_i(k,l)$ be the 2D-Fourier transform of $y_i(m,n)$ and $\bar{Y}(k,l)$ be the 2D-Fourier transform of $\bar{y}(m,n)$. Also, because we are primarily concerned with the frequency domain let us express $Y_i(k,l)$ and $\bar{Y}(k,l)$ as the column vectors $\mathbf{y}_i$ and $\bar{\mathbf{y}}$ respectively. Eq. (39) is equivalently represented in the frequency domain as

$$\begin{aligned}\text{ASM}&=\frac{1}{N\cdot d}\sum_{i=1}^{\mathrm{N}}\sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\left|Y_i(k,l)-\bar{Y}(k,l)\right|^2\\&=\frac{1}{N\cdot d}\sum_{i=1}^{\mathrm{N}}\left|\mathbf{y}_i-\bar{\mathbf{y}}\right|^2\end{aligned} \tag{41}$$

We must now introduce the filter itself into this metric to allow for optimization with respect to the filter coefficients. Let us consider the ASM over a set of correlation surfaces which are the result of filtering a set N training images $x_i(m,n)$ with the filter $h(m,n)$. As such let us express the Fourier transforms of the i^{th} training image and the filter as $X_i(k,l)$ and $H(k,l)$ respectively. Also, define $\bar{X}(k,l)$, the average Fourier transform of the N training images, as

$$\bar{X}(k,l)=\sum_{i=1}^{N}X_i(k,l) \tag{42}$$

We proceed by representing $X_i(k,l)$, $\bar{X}(k,l)$, $H(k,l)$ as column vectors $\mathbf{x}_i$, $\bar{\mathbf{x}}$, and $\mathbf{h}$ respectively. Let us now define the diagonal matrices $\mathbf{X}_i$ and $\bar{\mathbf{X}}$ whose non-zero elements are taken respectively from $\mathbf{x}_i$ and $\bar{\mathbf{x}}$. Using these matrices we can express $\mathbf{y}_i$ and $\bar{\mathbf{y}}$ as

$$\mathbf{y}_i=\mathbf{X}_i^*\mathbf{h} \tag{43}$$

$$\bar{\mathbf{y}}=\bar{\mathbf{X}}^*\mathbf{h} \tag{44}$$

Substituting the above equations in to Eq. (41) we have the following equivalent expression:

$$\begin{aligned}\text{ASM}&=\frac{1}{N\cdot d}\sum_{i=1}^{\mathrm{N}}\left|\mathbf{X}_i^*\mathbf{h}-\bar{\mathbf{X}}^*\mathbf{h}\right|^2\\&=\frac{1}{N\cdot d}\sum_{i=1}^{\mathrm{N}}\mathbf{h}^+\left(\mathbf{X}_i-\bar{\mathbf{X}}\right)\left(\mathbf{X}_i-\bar{\mathbf{X}}\right)^*\mathbf{h}\\&=\mathbf{h}^+\mathbf{S}\mathbf{h}\end{aligned} \tag{45}$$

where the diagonal matrix **S** is defined as

$$\mathbf{S} = \frac{1}{N \cdot d} \sum_{i=1}^{\mathrm{N}} \left(\mathbf{X}_i - \bar{\mathbf{X}}\right)\left(\mathbf{X}_i - \bar{\mathbf{X}}\right)^* \tag{46}$$

We have now expressed the distortion metric of ASM as a function of the filter and the training images. However, while minimizing distortion we also wish to maximize the filter's response to authentic patterns/faces. Unlike the MACE filter we have no constraint on the peak value and thus our desire is to maximize the correlation peak value over the entire set of training images i.e., maximize the average peak value. We denote this quantity by the measure of *Average Correlation Height* (ACH) whose definition is

$$\begin{aligned} \mathrm{ACH} &= \frac{1}{N} \sum_{i=1}^{N} y_i(0,0) \\ &= \frac{1}{N \cdot d} \sum_{i=1}^{N} \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} Y_i(k,l) \\ &= \frac{1}{N \cdot d} \sum_{i=1}^{N} \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} X_i^*(k,l) H(k,l) \end{aligned} \tag{47}$$

whose matrix vector formulation utilizing previously defined vectors $\mathbf{x}_i$, $\bar{\mathbf{x}}$, and $\mathbf{h}$ is

$$\begin{aligned} \mathrm{ACH} &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i^+ \mathbf{h} \\ &= \bar{\mathbf{x}}^+ \mathbf{h} \end{aligned} \tag{48}$$

While our immediate goal is to suppress ASM while maximizing ACH it is of course also desirable to suppress ONV as defined earlier. This simultaneous minimization maximization problem lends itself to a Rayleigh quotient representation as follows:

$$\begin{aligned} J(\mathbf{h}) &= \frac{|\mathrm{ACH}|^2}{\mathrm{ASM+ONV}} \\ &= \frac{\left|\mathbf{m}^+ \mathbf{h}\right|^2}{\mathbf{h}^+ \mathbf{S} \mathbf{h} + \mathbf{h}^+ \mathbf{C} \mathbf{h}} \\ &= \frac{\left|\mathbf{m}^+ \mathbf{h}\right|^2}{\mathbf{h}^+ (\mathbf{S}+\mathbf{C}) \mathbf{h}} \end{aligned} \tag{49}$$

The filter **h** that maximizes this ratio is the dominant eigenvector of $(\mathbf{S}+\mathbf{C})^{-1}\mathbf{m}\mathbf{m}^+$ which is

$$\mathbf{h} = \alpha (\mathbf{S}+\mathbf{C})^{-1} \mathbf{m} \tag{50}$$

where α is a normalizing coefficient. The above filter solution is termed the *Maximum Average Correlation Height* (MACH) filter (Mahalanobis et al., 1994). The MACH filter is often used in ATR applications where its tolerance for noise and distortion addresses the issue of

sensor noise and background clutter while maintaining the ability to resolve sharp and distinct correlation peaks necessary for accurate target detection and recognition. These same issues are paralleled in many face recognition applications and as such the same characteristics of the MACH filter are desired.

4.7 Optimal Tradeoff Filters

We have thus far developed ACFs whose derivation incorporate different desirable qualities such as the MACE filter's ability to resolve sharp correlation peaks, or the MVSDF filter's tolerance for noise. However, while these filter solutions provide these attractive properties they inherently create deficiencies in other aspects. For example, the MACE filter while being able to resolve sharp peaks, has little tolerance for noise while the MVSDF filter's tolerance for noise is offset by its relative inability to generate sharp correlation peaks. The fundamental issue concerning these particular ACFs is their singular focus on optimality with respect to one aspect of distortion. Depending on the application, a more preferred approach might be to design a filter whose optimality in these varying aspects is variable. In other words, we desire a filter which maintains a tradeoff between peak sharpness and noise tolerance. Termed *Optimal Tradeoff* (OT) filters, we will not go through the complete derivation in the interest of conciseness and its similarity to previous derivations.

The OT filter counterpart for the MACE and MVSDF filter is referred to as the *Optimal Tradeoff Synthetic Discriminant Function* (OTSDF) filter (Vijaya Kumar, 1994). It is obtained by minimizing a weighted sum of ACE and ONV which are the metrics for the MACE and MVSDF filters respectively. The resulting filter solution is

$$\mathbf{h} = (\alpha\mathbf{D} + \beta\mathbf{C})^{-1}\mathbf{X}\left[\mathbf{X}^{+}(\alpha\mathbf{D} + \beta\mathbf{C})^{-1}\mathbf{X}\right]^{-1}\mathbf{u} \tag{51}$$

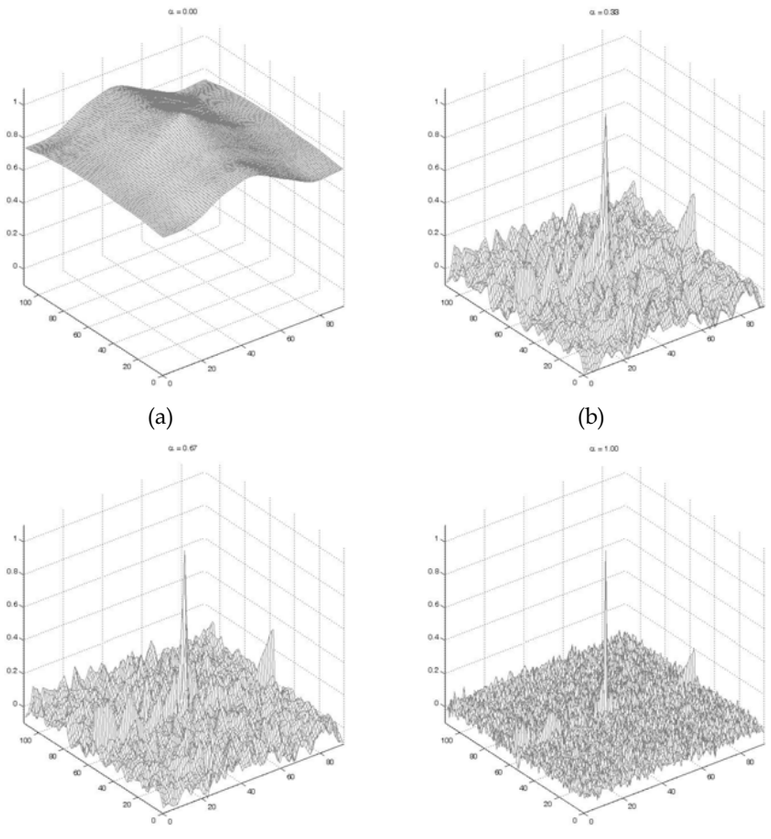
where α and β are non-negative constants that can be varied to achieve a desired amount of performance with respect to noise and peak sharpness while all other variables retain their definitions from previous sections. In order better constrain the relationship between the tradeoff between noise tolerance and peak sharpness we constrain the relationship between α and β with the following:

$$\alpha^2 + \beta^2 = 1 \tag{52}$$

This constraint is a result of the quadratic nature of the filter solution (Vijaya Kumar, 1994) Allowing us to rewrite Eq. (51) as function of α alone in the following manner:

$$\mathbf{h} = \left(\alpha\mathbf{D} + \left(\sqrt{1-\alpha^2}\right)\mathbf{C}\right)^{-1}\mathbf{X}\left[\mathbf{X}^{+}\left(\alpha\mathbf{D} + \left(\sqrt{1-\alpha^2}\right)\mathbf{C}\right)^{-1}\mathbf{X}\right]^{-1}\mathbf{u} \tag{53}$$

Since α is non-negative, we can vary the amount of noise tolerance and peak sharpness in the filter by varying α between 0 and 1. If we were to set α to 0, then Eq. (53) reduces to the MVSDF filter solution in Eq. (38) while and α of 1 yields the MACE filter solution of Eq. (35). By choosing values of α in this range we are essentially creating a filter which is a weighted combination of the MACE and MVSDF filters. Typically α is set close to 1 in order to maintain sharp peaks while incorporating a small degree of noise tolerance. Most experiments concerning the effect of α on filter performance have supported this notion.



(a) (b)

(c) (d)

Figure 10. Mesh plots of correlation planes produced from application of OTSDF filter expecting AWGN of SNR 20 dB to a training image with varying values of α (a) $\alpha = 0$ (b) $\alpha = 0.33$ (c) $\alpha = 0.67$ (d) $\alpha = 1.00$

However, one must take into account the type and degree of noise the filter is being designed to accommodate for. In many applications *Additive White Gaussian Noise* (AWGN) is the standard form of noise for which depending on its magnitude or equivalently its SNR can be negligible. However, when the magnitude of the noise is non-negligible we can observe the effect of α paremeter upon the filter design and any subsequent correlation planes. Figure 10 demonstrates this aspect of the OTSDF filter by presenting the correlations of one of the training images with four OTSDF filters each designed with different values of α and expecting AWGN of SNR 20 dB. The most noticeable change between the correlation planes is the relative strength of the sidelobes throughout the correlation plane.

Since the MACH filter is often thought of as the unconstrained version of the MACE filter, we call the MACH filter's OT filter the *Unconstrained Optimal Tradeoff Synthetic Discriminant Function* (UOTSDF) filter (Vijaya Kumar, 1994). The solution to the UOTSDF filter is

$$\mathbf{h} = \left(\alpha\mathbf{D} + \beta\mathbf{C} + \gamma\mathbf{S}\right)^{-1}\mathbf{m} \tag{54}$$

where α, β, and γ are tradeoff parameters for ACE, ONV, and ASM respectively while all other variables retain their previous definitions. Though there exists a quadratic relationship between these parameters we often choose to fix at least one parameter while optimizing the others with respect to performance.

4.8 Performance Measures

When considering the use of correlation in pattern recognition and in particular face recognition applications it becomes necessary to define a metric by which to quantify the "goodness" or "correctness" of a correlation. A simple and sometimes effective way to quantize a match is to take the largest value in a correlation plane and threshold it to yield a match or no-match decision. This approach works well when there is relatively small variation in the data such that the variance in the value of the correlation peak is small. This assumption is of course is an idealization and, with particular focus on face recognition, a poor one. The value of the correlation peak will vary in the presence of intensity changes and noise in non-negligible amounts and as such a strict threshold cannot be expected to be a reliable performance measure.

When considering such ACFs as the MACE and OTSDF filters a more appropriate measure is that of peak sharpness since these filters are designed to suppress the sidelobes adjacent to peaks. This relationship can be quantized by the *Peak-to-Sidelobe Ratio* (PSR) which for a particular peak is defined as

$$\mathrm{PSR} = \frac{\left(\text{peak value}\right) - \mu_{\text{area}}}{\sigma_{\text{area}}} \tag{55}$$

where μ_{area} and σ_{area} are the mean and standard deviation respectively of some small area or neighborhood around but not including the peak.

Similarly the MACH and UOTSDF filters are designed to maximize the value of the peak relative to the rest of correlation plane also. Thus a similar but alternate performance measure would be one that measured the magnitude of difference between the peak and the rest of the correlation plane. Using the metric of *Peak-to-Correlation Energy* (PCE) we can quantify this difference as

$$\mathrm{PCE} = \frac{\left(\text{peak value}\right) - \mu_{\text{plane}}}{\sigma_{\text{plane}}} \tag{56}$$

where μ_{plane} and σ_{plane} are the mean and standard deviation respectively of the entire correlation plane excluding the peak.

Both PSR and PCE can be used with any ACF but the optimal measure often depends on the application. In most situations where the resolution or size of the target is relatively constant as is the case with many face recognition applications, PSR is a sufficient measure. On the other hand, algorithms that use multi-resolution techniques might benefit more from PCE. Regardless, both measures still require a threshold to determine a match or no-match decision although in contrast with a strict threshold on the peak value alone, a threshold on PSR or PCE values is far more normalized and predictable.

5. Face Recognition Using Advanced Correlation Filters

5.1 Face to Sketch Correlation

One of the primary issues in many face recognition systems is that of illumination variation. An innumerable number of changing factors determine the exact nature of the illumination a face may be subject to at any given time. As such, the span of illumination variation is vast and often of non-negligible magnitude. In order for a face recognition system to objectively make the claim that it is capable of unrestricted field deployment it must be able to compensate for any type of illumination variation. One approach to this issue is to re-train the recognition system each time it is presented with a new environment or situation where the illumination has varied from previously known conditions. This can be costly both in terms of time and money and most of the time, this is not feasible or possible to capture all possible lighting conditions (especially when outdoors) so as a result this is not done in practice. Another approach is to incorporate some sort of illumination-preprocessing algorithm in order to compensate for varying illuminations. This method is much preferred to the former due to its hopefully broader and more effective application. Nonetheless, deriving such a preprocessing stage is in itself challenging given the degree of illumination compensation one is attempting to achieve. One of the more novel approaches to this problem involves using eigenanalysis and ACFs to reconstruct and recognize images respectively using a different representation of the face (Li et al., 2006). The relative uniqueness of this approach can be traced to the fact that it utilizes both traditional facial images coupled with corresponding facial sketches thsat are similar to those found in law enforcement.

Consider the field of law enforcement applications where one of the most commonly used tools is that of a police sketch which is used to to help identify suspect criminals. Although visual surveillance equipment is present in many everyday environments, they are often of low quality and are not optimal for enrolling police sketches. To this end, the role of the witness becomes exceptionally important as a source of more reliable evidence. The police sketch allows the witness' recollection of a suspect to manifest itself as a piece of visual evidence. Nonetheless the usefulness of the previously mentioned surveillance equipment should not be discounted. In many high security locations, continuous video surveillance is present and provides us with some record of people who have passed through those monitored locations. However, due to factors such as time of day and physical setup of the surveillance equipment, the exact lighting conditions which illuminate the faces of the passing people can vary. The issue now becomes one of recognizing the suspect in the surveillance data using the witness' police sketch as the template.

This kind of question can be categorized as robust face recognition for illumination variation. However this application is different from the normal face recognition scenario; where the enrollment gallery image is a real face image, as in this case the enrollment gallery image for finding the suspect from surveillance video is a police sketch of the suspect's face. Of course there are strong similarities and high correlations between the real face images and the corresponding police sketch image. If one can capture the correlation between these two representations of same person, and describe those correlations in a useful mathematical form, then it will be very useful for finding a solution to this problem.

In literature, there are two main types of approaches proposed for this kind of face-sketch dual space modeling problem. Both methodologies utilize eigenanalysis to form a basis for representing the face-sketch dual space, similar to how eigenanalysis is used in PCA and *Eigenface* applications. However, these two methodologies differ in the way of how they

form the eigen-subspace and how they capture the correlation between the two subspaces (i.e., face and sketch subspace).

The first approach tries to construct the PCA subspace for both the face and sketch images separately, by transforming all the training data (which are real face images) into corresponding sketch images, and then perform classification in sketch space (Tang et al., 2002), this approach may face issues in practice as images with illumination variation will generate sketches with artifacts. The second approach to this problem tries to reconstruct the original face image from the given sketch image using a hybrid-eigenspace representation, and then perform classification using ACFs (Li et al., 2006). In the next few paragraphs, we will look at more details of this algorithm. The key idea us that a face recognition system, which takes surveillance footage, typically is subject to variable and unknown illumination artifacts and will not be able to synthesize good sketch images (corresponding to the illumination distorted face image) on the fly due illumination variation artifacts. These artifacts will be enhanced and significantly affect the resulting automatically generated sketch image. which will in turn ultimately affect recognition performance. In our proposed approach we reconstruct what the person 'looks-like' from the sketch image and then use this reconstructed image as the gallery enrollment image in the face recognition system which can recognize the person under the presence of illumination variations (as demonstrated with example experiments on the PIE database).

We can use three stages to describe this algorithm: the training stage, the synthesis stage, and the recognition stage. Assume the training data has N face images and their corresponding N sketch images. We denote the i^{th} face image as $\mathbf{f}_i$, and the i^{th} sketch image as $\mathbf{s}_i$ where $i = 1, 2, \ldots, N$. By appending each face image with its corresponding sketch counterpart, we can form a new subspace, which is called "hybrid-subspace" in (Li et al., 2006). Then we can describe all of the training data in the following matrix form:

$$\mathbf{D_h} = \left[\frac{\mathbf{D_f}}{\mathbf{D_s}}\right] = \left[\frac{\mathbf{f}_1 \ldots \mathbf{f}_m}{\mathbf{s}_1 \ldots \mathbf{s}_m}\right] \tag{57}$$

where each $\mathbf{f}_i$ and $\mathbf{s}_i$ are column vectors, and $\mathbf{D_f}$ consists of the face data matrix, and $\mathbf{D_s}$ as the corresponding sketch data matrix. Our next step is to derive an orthonormal basis that represents our combined face data. Therefore, standard PCA is performed on the hybrid data matrix $\mathbf{D_h}$. We first remove the mean of the data by computing $\boldsymbol{\mu}_f$ and $\boldsymbol{\mu}_s$:

$$\boldsymbol{\mu}_\mathbf{f} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{f}_i \tag{58}$$

$$\boldsymbol{\mu}_\mathbf{s} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{s}_i \tag{59}$$

$$\mathbf{X} = \left[\frac{\mathbf{X_f}}{\mathbf{X_s}}\right] = \left[\frac{\mathbf{f}_1 - \boldsymbol{\mu}_\mathbf{f} \ldots \mathbf{f}_\mathbf{m} - \boldsymbol{\mu}_\mathbf{f}}{\mathbf{s}_1 - \boldsymbol{\mu}_\mathbf{f} \ldots \mathbf{s}_\mathbf{m} - \boldsymbol{\mu}_\mathbf{f}}\right] \tag{60}$$

Then the covariance matrix $\boldsymbol{\Sigma}$ is defined as:

$$\boldsymbol{\Sigma} = \frac{1}{N}\mathbf{X}\mathbf{X}^\mathbf{T} \tag{61}$$

Once we have $\boldsymbol{\Sigma}$, we perform eigenanalysis to derive its eigenvectors and eigenvalues:

$$\boldsymbol{\Sigma\Omega_h = \Lambda_h\Omega_h} \tag{62}$$

where $\boldsymbol{\Omega_h} = [\boldsymbol{\omega}_1 \ldots \boldsymbol{\omega}_m]$ such that $\boldsymbol{\omega}_i$ is the i^{th} eigenvector of $\boldsymbol{\Sigma}$, and $\boldsymbol{\Lambda_h} = diag(\boldsymbol{\lambda}_1 \ldots \boldsymbol{\lambda}_m)$ such that $\boldsymbol{\lambda}_i$ represents the i^{th} eigenvalue.

Because every single column in $\mathbf{D_h}$ contains a face image and its corresponding sketch images, each $\boldsymbol{\omega}_i$ can be interpreted as consisting of two components as follows:

$$\boldsymbol{\omega}_i = \begin{bmatrix} \boldsymbol{\omega}_{\mathbf{sf},i} \\ \boldsymbol{\omega}_{\mathbf{ss},i} \end{bmatrix} \tag{63}$$

We can call $\boldsymbol{\omega}_{\mathbf{sf},i}$ as "pseudo-eigenface" and $\boldsymbol{\omega}_{\mathbf{ss},i}$ as "pseudo-eigensketch", because they represent the variations in face and sketch subspace, respectively. The reason we add "pseudo" in front of "eigenface" and "eigensketch" is because the orthogonality is no longer preserved when the $\boldsymbol{\omega}_i$ vector is partitioned into two parts. The set of $\boldsymbol{\omega}_i$ vectors form an orthogonal basis, however neither $\boldsymbol{\omega}_{\mathbf{sf}}$ or $\boldsymbol{\omega}_{\mathbf{ss}}$ do. Therefore, one should not use the standard projection method to compute the projection coefficients, as one would do in standard PCA case. Instead, one should use "pseudo-inverse" method to derive projection coefficients, and this is exactly what is proposed in hybrid-subspace method.

Given a probe sketch image $\mathbf{s_p}$, the pseudo-inverse procedure is performed to find the optimal projection coefficient:

$$\mathbf{P} = (\boldsymbol{\omega}_{\mathbf{ss}}^{\mathrm{T}}\boldsymbol{\omega}_{\mathbf{ss}})^{-1}\boldsymbol{\omega}_{\mathbf{ss}}\mathbf{s_p} \tag{64}$$

By using this projection coefficient in the subspace spanned by $\boldsymbol{\omega}_{\mathbf{sf}}$, one can reconstruct the face image in pseudo-eigenface subspace, as described in following equation:

$$\mathbf{I}_{\text{reconstructed}} = \boldsymbol{\omega}_{\mathbf{sf}}^{\mathrm{T}}\mathbf{P} \tag{65}$$

Hence, a new face image is hallucinated from the given sketch image. A few examples of the face images and their corresponding sketch images, probe sketch images and the reconstructed (hallucinated) face images are shown in Figure 11. We can see that the reconstructed face images preserved most of the characteristics from the original face images, which exhibit the effectiveness of the hybrid-subspace method. However, there are also some discrepancies between the original face images and the reconstructed ones. The differences are mostly the level of intensity around the forehead and cheek. This is because from a sketch it is not possible to extract the color of the face, that is meta-data which is given by the victim and can easily be added to this model.

We have shown in previous sections that ACFs have significant illumination tolerance which shows that when test images have different level of illumination than training images, ACFs can successfully achieve high recognition rate without the need to re-train the classifiers. Therefore, ACFs are one of the best candidates of the possible pattern recognition classifiers used in this application. The performance of ACF is reported to be significantly much higher, when compared to traditional approach of nearest-neighbor method (1-NNM), with exactly the same reconstruction steps used in face reconstruction (hallucination) stage, as shown in Figure 12.

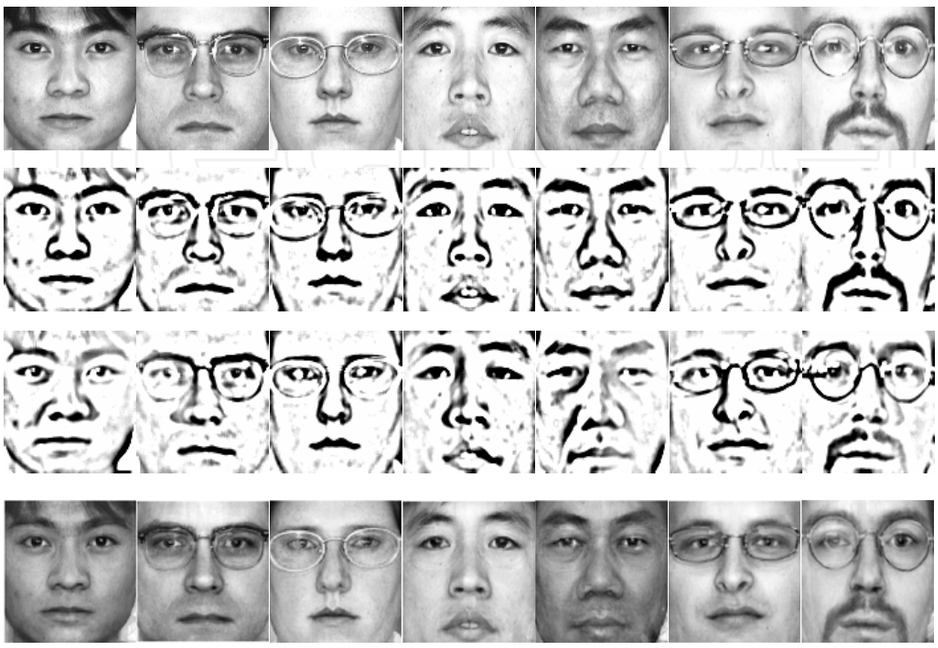
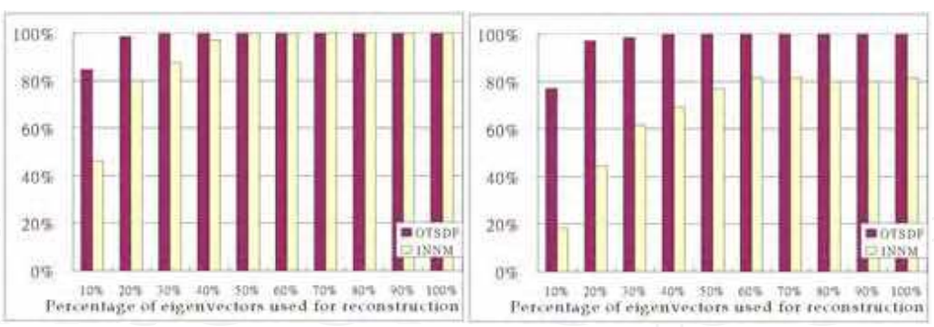


Figure 11. Examples from CMU PIE database (first row) example face images in training database (second row) corresponding sketch images with respect to the first row (third row) given probe sketch images (fourth row) reconstructed (hallucinated) face images based on the hybrid-subspace approach



(a) (b)

Figure 12. Experimental rank-1 identification rate results of hybrid-subspace method (a) Results from CMU PIE Light database: OTSDF and 1-NNM, using 8th and 11th image of all the subjects to train hybrid-subspace while using sketch of 20th image, and testing against all images (b) Results from CMU PIE No-Light database: OTSDF and 1-NNM, using 7th and 10th image of all the subjects to train hybrid-subspace while using sketch of 19th image, and testing against all images

In summary, for face recognition problems in face-sketch dual subspace, hybrid-subspace method combined with ACF has been proved as a good direction. It captures the correlation between face and sketch subspaces by form a hybrid subspace and train pseudo-eigen basis from it. It can successfully reconstruct the original face image and by then performing classification using ACF, one can overcome difficulties resulted from the illumination variation and still achieve high recognition results.

5.2 Empirical Mode Decomposition Preprocessing and ACFs

Amongst our latest research that utilizes ACFs makes use of the powerful signal processing tool of *Empirical Mode Decomposition* (EMD) (Huang et al., 1998). Relatively new to the field of face recognition, EMD is traditionally applied to 1-D signal processing problems where the goal is to isolate underlying trends and details in data. Fundamentally this is the goal of illumination preprocessing where the underlying trend is the neutral illumination.

Pioneered as a signal processing technique for adaptive representation of non-stationary signals as sums of zero-mean AM and FM components, EMD has been successfully employed in multiple applications not directly related to facial recognition. EMD's ability for adaptive representation of signals allows for controlled reconstruction of signals. Though it is considered a very powerful tool, it is fundamentally an empirical algorithm as opposed to theory wherein lays the potential for multiple and varying interpretations. However, we present here only the most basic flavor of EMD from which all other variations of EMD are derived from. A more thorough development and description of EMD is presented in other works (Flandrin et al., 2003) as compared to the one detailed in Table 2.

1.)	Identify all local minima and maxima of $x(t)$
2.)	Interpolate between all minima to yield an envelope $e_{\min}(t)$. Similarly, interpolate between all maxima to yield an envelope $e_{\max}(t)$
3.)	Compute the mean envelope $m(t) = \left(e_{\min}(t) + e_{\max}(t)\right)/2$
4.)	Compute the detail $d(t) = x(t) - m(t)$
5.)	If $\left\|m(t)\right\| < \varepsilon$. If not repeat steps 1-4 with $d(t)$ as the input signal $x(t)$. If so, $d(t)$ is an *Intrinsic Mode Function* (IMF)
6.)	Calculate residual $r(t) = x(t) - d(t)$
7.)	Go back to step 1 with $r(t)$ as the input signal $x(t)$
8.)	Repeat until input signal no longer has any extrema

Table 2. Basic EMD algorithm

Although the described algorithm implies the use of 1-D data, there are variants of EMD specifically created for use with 2-D data (Damerval et al., 2005) such as facial images. In the interest of conciseness, we will not thoroughly develop the EMD algorithm but instead emphasize the end result of applying EMD to a signal. Essentially EMD decomposes an input signal into a set of *Intrinsic Mode Functions* (IMFs) from which the original input signal can be recovered via the simple summation of said IMFs. In this sense, the IMFs that are the

result of application of EMD to a signal can be thought of as a series of basis signals for the input signal. Using EMD as a preprocessing tool, we can decompose facial images into their IMFs or basis images of which a few will contain the majority of illumination effects. Reconstruction of the original facial image sans these illumination-variant IMFs will yield a more illumination-neutral image from which more accurate recognition can be performed (Bhagavatula & Savvides, 2007).

(a) (b)

Figure 13. Result of EMD preprocessing on an image taken from the PIE No-Lights face database (a) Prior to EMD preprocessing (b) After EMD preprocessing

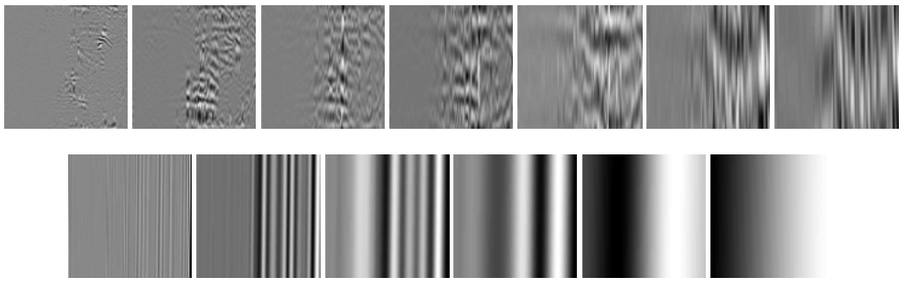


Figure 14. IMFs created from applying EMD to the face image in Figure 13 (a)

As Figure 13 demonstrates, EMD preprocessing is capable of removing cast shadow effects while retaining the majority of useful information. Although the image in Figure 13 (b) appears discolored, it is a far better image to perform face recognition on than the original image presented in Figure 13 (a). To further illustrate this point we present in Figure 14 the IMFs decomposed from the image in Figure 13 (a). Taking note of the last IMF, we can clearly see the overall effect of the cast shadow in this IMF and can intuitively appreciate the effect of reconstructing the facial image minus this particular IMF. We show in Figure 15 the average performance of ACFs prior to and after EMD preprocessing on the Carnegie Mellon University Pose-Illumination-Expression (CMU PIE) No-Lights face database (Sims et al, 2003). Our results indicate that although ACFs perform exceedingly well even under illumination-variant conditions, their performance does benefit from some illumination normalization as is provided by EMD preprocessing. These results not only underscore the power of ACFs but also that of EMD which as signal decomposition tool which effectively yields AM and FM components of a signal is also a frequency domain processing technique. With both these algorithms available to us, we are capable of achieving significantly accurate face recognition in illumination-variant conditions.

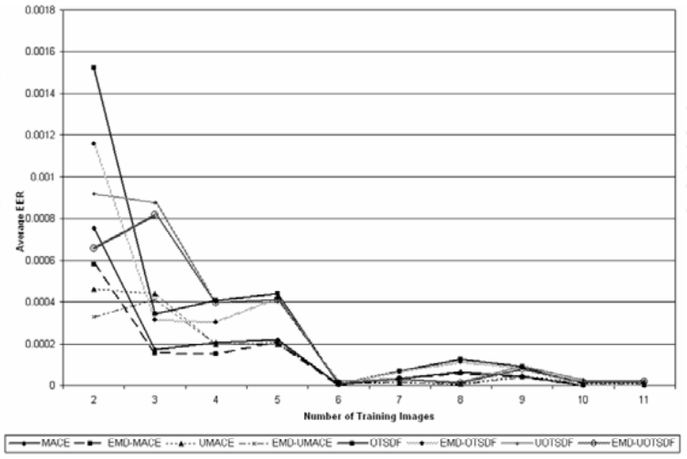


Figure 15. Average EERs comparing performance of ACFs prior to and after EMD preprocessing

6. Conclusions and Future Work

We have shown through the course of this chapter that the Fourier or frequency domain of facial data contains significantly more useful information when processed than its spatial counterpart. The simple coupling of standard algorithms such as *Eigenfaces* and *Fisherfaces* with frequency domain representation of phase and magnitude spectrums, can result in noticeable improvements in performance as we have shown for pose and illumination tolerance. Evolving our intuition about the frequency domain leads us to the group of algorithms collectively referred as ACFs. Primarily originating from frequency domain interpretations of data, ACFs allow for significant discriminative ability while providing other attractive qualities such as shift invariance, noise tolerance, and graceful degredation. As the presented results indicate, ACFs are capable of performing highly accurate face recognition in varying and challenging circumstances. In particular, the presented work also demonstrates the compatibility of ACFs with other algorithms allowing them to be easily integrated into most face recognition systems.

Frequency domain related algorithms, particularly ACFs, still hold much potential in advancing the area of face recognition and biometrics in general. Our proposed future work spans the broad horizon of face recognition including but not limited to improved general face recognition, large scale applications, improved illumination tolerance, hardware implementations, and privacy issues. The last area mentioned holds great significance in today's digital world. Although biometrics are gaining popularity as a reliable and secure method of authentication and identification, they are as susceptible to loss as typical ciphers or passwords. Represented as digital data, a biometric template can be stolen and as an almost unique identifier of a person cannot be replaced. To this end, cancellable biometrics are being developed to allow re-usability and re-issuement of biometrics using encryption type methods and performing the recognition in the encrypted domain. ACFs easily integrate into the scheme of cancellable biometrics (Jain & Uludag, 2003; Savvides et al.,

2004c, 2004e; Ratha et al., 2006). This research is amongst the most pressing as widespread acceptance of face recognition is contingent on allaying these privacy concerns.

Face recognition research and technology has made significant progress over the last decade. Advances in recognition algorithms have enabled some headway into commercially viable systems. However, performance is still considered lacking with respect to the need for reliable and accurate identification. Our research into frequency domain algorithms is but one of many approaches to this problem. However, unlike other approaches, ours' is relatively unique and offers a great potential for improvement with the designed distortion tolerance and shift-invariance. We intend to continue with our research in frequency domain face recognition exploiting and analyzing all aspects of the frequency content of useful for identifying human faces.

7. References

Bhagavatula, R. & Savvides, M. (2005a) Eigen and Fisher-Fourier spectra for shift invariant pose-tolerant face recognition. *Proceedings of International Conference on Advances in Pattern Recognition, 2005*, pp. II-351 - II-359, Bath (UK), Aug. 2005, Springer

Bhagavatula, R. & Savvides, M. (2005b) PCA vs. automatically pruned wavelet-packet PCA for illumination tolerant face recognition. *Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies, 2005*, pp. 69 - 74, Buffalo, NY (USA), Oct. 2005, IEEE

Bhagavatula, R. & Savvides, M. (2007) Empirical mode decomposition for removal of specular reflections and cast shadow effects. *Proceedings of SPIE Defense and Security Symposium, 2007*, (pending publication), Orlando, FL (USA), April 2007, SPIE

Belhumeur, P.N.; Hespanha, J.P. & Kreigman, D.J. (1997) Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, July 1997, pp. 711 - 720

Chen, T.; Hsu, Y.J.; Liu, X. & Whang, W. (2002) Principle component analysis and its variants for biometrics. *Proceedings of IEEE International Conference on Image Processing, 2002*, pp. I-61 - I-64, Rochester, NY (USA), Sept. 2002, IEEE

Chellappa, R.; Wilson, C.L. & Sirohey, S. (1995) Human and machine recognition of faces: a survey. *Proceedings of the IEEE*, Vol. 83, No. 5, May 1995, pp. 705 - 741

Damerval, C.; Meignen, S. & Perrier, V. (2005) A fast algorithm for bidimensional EMD. *IEEE Signal Processing Letters*, Vol. 12, No. 10, Oct. 2005, pp. 701 - 704

Flandrin, P.; Goncalves, P. & Rilling, G. (2003) On empirical mode decomposition and its algorithms. *Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, 2003*, Grod-Trieste (Italy), June 2003, IEEE

Fisher, R. (1936). The use of multiple measures in taxonomic problems. *Ann. Eugenics*, Vol. 7, 1936, pp. 179-188

Grudin, M.A. (2000) On internal representation in face recognition systems. *Pattern Recognition*, Vol. 33, No. 7, 2000, pp. 1161 - 1177

Hayes, M.H; Lim, J.S & Oppenheim, A.V (1980) Signal reconstruction from phase or magnitude. *IEEE Transactions on Acoustics and Signal Processing*, Vol. 28, Dec 1980, pp 672 - 680

Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.L.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung C.C., & Liu, H.H. (1998) The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings Royal Society London*, Vol. 454, 1998, pp. 903 - 995

Heo, J.; Savvides, M. & Vijaya Kumar, B.V.K. (2005) Performance evaluation of face recognition using visual and thermal imagery with advanced correlation filters. *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition, 2005*, pp. III-9 - III-14, San Diego, CA (USA), June 2005, IEEE

Heo, J.; Savvides, M.; Abiantun, R.; Xie, C. & Vijaya Kumar, B.V.K. (2006) Face recognition with kernel correlation filters on a large scale database. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2006,* pp. II-181 - II-184, Toulouse (France), May 2006, IEEE

Hester, C. & Casasent, D. (1980) Multivariant technique for multiclass pattern recognition. *Applied Optics*, Vol. 19, No. 11, June 1980, pp.1758 - 1761

Jain, A.K. & Uludag, U. (2003) Hiding biometric data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 11, Nov. 2003, pp. 1494 - 1498

Li, Y.; Savvides, M. & Vijaya Kumar, B.V.K. (2006). Illumination tolerant face recognition using a novel face from sketch synthesis approach and advanced correlation filters. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2006,* pp. II-357 - II-360, Toulouse (France), May 2006, IEEE

Mahalanobis, A.; Vijaya Kumar, B.V.K. & Casasent, D. (1987). Minimum average correlation energy filters. *Applied Optics*, Vol. 26, Issue 17, Sept., 1987, pp. 3633 - 3640

Mahalanobis, A.; Vijaya Kumar, B.V.K.; Song, S.; Sims, S.R.F. & Epperson, J.F. (1994). Unconstrained correlation filters. *Applied Optics*, Vol. 33, No. 17, June 1994, pp. 3751 - 3759

Ng, C.K.; Savvides, M. & Khosla, P.K. (2005) Real-time face verification on a cell-phone using advanced correlation filters. *Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies, 2005*, pp. 57 - 62, Buffalo, NY (USA), Oct. 2005, IEEE

Oppenheim, A.V & Lim, J.S. (1981). The importance of phase in signals. *Proceedings of the IEEE* Vol. 69, No. 5, May 1981, pp 529 - 541

Phillips, P.J.; Moon, H.; Rizvi, S.A. & Rauss P.J. (2000) The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, Oct. 2000, pp. 1090 - 1104

Phillips, P.J.; Grother, P.; Micheals, R.; Blackburn, D.M.; Tabassi, E. & Bone, M. (2003) Face recognition vendor test 2002. *Proceedings of IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2003*, pp. 44, Nice (France), Oct. 2003, IEEE

Phillips, P.J.; Flynn, P.J.; Scruggs, T.; Bowyer, K.W.; Chang, J.; Hoffman, K.; Marques, J.; Min, J. & Worek, W. (2005) Overview of the face recognition grand challenge. *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2005*, pp. I-947 - I-954, San Diego, CA (USA), June 2005, IEEE

Ratha, N.; Connell, J.; Bolle, R.N. & Chikkerur, S. (2006) Cancelable biometrics: a case study in fingerprints. *Proceedings of IEEE International Conference on Pattern Recognition, 2006*, pp. IV-370 - IV-373, Hong Kong (China), Aug. 2006, IEEE

Savvides, M.; Vijaya Kumar, B.V.K. & Khosla P.K. (2002) Two-class minimax distance transform correlation filter. *Applied Optics*, Vol. 31, No. 32, Nov. 2002, pp. 6829 - 6840

Savvides, M. & Vijaya Kumar (2003a) Illumination normalization using logarithm transforms for face authentication. *Proceedings of International Conference on Advances in Pattern Recognition, 2003*, pp. 549 - 556, Guildford (UK), June 2003, Springer

Savvides, M. & Vijaya Kumar (2003b) Quad phase minimum average correlation energy filters for reduced memory illumination tolerant face authentication. *Proceedings of International Conference on Advances in Pattern Recognition, 2003*, pp. 19 - 26, Guildford (UK), June 2003, Springer

Savvides, M.; Venkataramani & Vijaya Kumar, B.V.K. (2003c) Incremental updating of advanced correlation filters for biometric authentication systems. *Proceedings of IEEE International Conference on Multimedia and Expo, 2003*, pp. III-229 - III-232, Baltimore, MD (USA), July 2003, IEEE

Savvides, M.; Vijaya Kumar, B.V.K. & Khosla, P.K. (2004a) "Corefaces" - robust shift invariant PCA based correlation filter for illumination tolerant face recognition. *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2004*, pp. II-834 - II-841, Washington, DC (USA), July 2004, IEEE

Savvides, M.; Vijaya Kumar, B.V.K. & Khosla, P.K. (2004b) Eigenphases vs. Eigenfaces. *Proceedings of IEEE International Conference on Pattern Recognition, 2004*, pp. III-810 - III-813, Cambridge (UK), Aug. 2004, IEEE

Savvides, M. & Vijaya Kumar, B.V.K. (2004c) Cancellable biometric filters for face recognition. *Proceedings of IEEE International Conference on Pattern Recognition, 2004*, pp. III-922 - III-925, Cambridge (UK), Aug. 2004, IEEE

Savvides, M.; Vijaya Kumar, B.V.K. & Khosla P.K. (2004d) Robust shift invariant biometric identification from partial face images. *Proceedings of SPIE Defense and Security Symposium, 2004*, pp. 124 - 135, Orlando, FL (USA), Aug. 2004, SPIE

Savvides, M.; Vijaya Kumar, B.V.K. & Khosla, P.K. (2004e) Authentication invariant cancellable correlation filters for illumination tolerant face recognition. *Proceedings of SPIE Defense and Security Symposium, 2004*, pp. 156 - 163, Orlando, FL (USA), Aug. 2004, SPIE

Savvides, M.; Vijaya Kumar, B.V.K. & Khosla P.K. (2004f) Illumination tolerant face recognition using advanced correlation filters trained from a single image. *Presented at the Biometrics Consortium, 2004*, Crystal City, VA (USA), 2004

Savvides, M.; Heo, J.; Abiantun, R.; Xie, C. & Vijaya Kumar, B.V.K. (2006a) Class dependent kernel discrete cosine transform features for enhanced holistic face recognition in FRGC-II. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2006,* pp. II-185 - II-188, Toulouse (France), May 2006, IEEE

Savvides, M.; Heo, J.; Abiantun, R.; Xie, C. & Vijaya Kumar, B.V.K. (2006b) Partial and holistic face recognition on FRGC-II data using support vector machines. *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2006*, pp. 48 - 53, New York, NY (USA), June 2006, IEEE

Sim, T.; Baker, S. & Bsat, M. (2003) The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 12, Dec. 2003, pp. 1615 - 1618

Swets, D.L. & Weng, J. (1996) Discriminant analysis and eigenspace partition tree for face and object recognition from view. *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 1996*, pp. 192 - 197, Killington, VT (USA), Oct. 1996, IEEE

Tang, X. & Wang X. (2002). Face photo recognition using sketch. *Proceedings of IEEE International Conference on Image Processing, 2002*, pp. I-257 - I-260, Rochester, NY (USA), Sept. 2002, IEEE

Turk, M.A. & Pentland, A.P. (1991) Face recogntion using eigenfaces. *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 1991*, pp. 589 - 591, June 1991, IEEE

Vijaya Kumar, B.V.K. (1986). Minimum variance synthetic discriminant functions. *Journal of the Optical Society of America*, Vol. 3, No. 10, Oct. 1986, pp. 1579 - 1584

Vijaya Kumar, B.V.K. (1992). Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, Vol. 31, No. 23, Aug. 1992, pp. 4773 - 4801

Vijaya Kumar, B.V.K.; Carlson, D. & Mahalanobis, A. (1994). Optimal tradeoff synthetic discriminant function (OTSDF) filters for arbitrary devices. *Optics Letters*, Vol. 19, No. 19, Oct. 1994, pp. 1556 - 1558

Vijaya Kumar, B.V.K.; Mahalanobis, A. & Juday, R. (2005) *Correlation Pattern Recognition*, Cambridge University Press, 13 978-0-521-57103-6, New York, NY (USA)

Vijaya Kumar, B.V.K.; Savvides, M. & Xie, C. (2006) Correlation pattern recognition for face recognition. *Proceedings of the IEEE*, Vol. 94, No. 11, Nov. 2006, pp 1963 - 1976

Yang, J.; Zhang, D.; Frangi, A.F. & Yang, J. (2004) Two-dimensional PCA: a new approach to appearance-based face recognition and recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, Issue 1, Jan. 2004, pp. 131 - 137

Zhao, W.; Chellappa, R. & Krishnaswamy, A. (1998) Discriminant analysis of princpal components for face recognition. *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 1998*, pp 336 - 341, Nara (Japan), April 1998, IEEE

Zhao, W.; Chellappa, R. & Phillips, P.J. (1999) Subspace linear discriminant analysis for face recognition. *Technical Report CAR-TR-914*, 1999

Zhao, W.; Chellappa, R.; Phillips, P.J. & Rosenfield, A. (2003) Face recognition: a literature survey. *Association for Computer Machinery Computing Surveys*, Vol. 35, No. 4, Dec. 2003, pp. 399 - 458

Face Recognition

Edited by Kresimir Delac and Mislav Grgic

ISBN 978-3-902613-03-5
Hard cover, 558 pages
Publisher I-Tech Education and Publishing
Published online 01, July, 2007
Published in print edition July, 2007

This book will serve as a handbook for students, researchers and practitioners in the area of automatic (computer) face recognition and inspire some future research ideas by identifying potential research directions. The book consists of 28 chapters, each focusing on a certain aspect of the problem. Within every chapter the reader will be given an overview of background information on the subject at hand and in many cases a description of the authors' original proposed solution. The chapters in this book are sorted alphabetically, according to the first author's surname. They should give the reader a general idea where the current research efforts are heading, both within the face recognition area itself and in interdisciplinary approaches.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen