Chapter 1

Introduction

1.1 Introduction

In recent years, the techniques of face recognition have become an active area of research for its potential biometric interest. Majority of systems whose primary interest is face recognition and understanding of face images emphasize on the analysis of the spatial representation of the images i.e. the intensity values of the images. While there has been varying and significant levels of performance achieved through the use of spatial 2-D image data, the use of a frequency domain representation sometimes achieves better performance for the face recognition tasks. The use of the Fourier and other frequency domain transforms allow to quickly and easily obtain raw frequency data which are significantly more discriminating (after appropriate data manipulation) than the raw spatial data, from which it is derived. One can further increase the discrimination ability through additional and specific feature extraction algorithms intended for use in the frequency domain. In majority cases, correlation filters[1] are used to achieve desired performances.

Correlation is a robust and general technique for pattern recognition. Ever since

---

the first use of the optical correlator for implementing matched spatial filters by Vand-
derLugt [2], researchers have been trying to develop better filters for the recognition of shapes, objects and faces. Such filters are popularly referred to as correlation filters [1] since they are designed for implementation in frequency plane correlators. Frequency domain face recognition techniques are executed by cross-correlating the Fourier transform of test face image with a synthesized template or filter, generated from Fourier transform of training face images. The processing results in a correlation output via inverse Fourier transform. An ideal correlation filter for face recognition would yield sharp correlation peak for a perfect match of the correlation filter with test face image present in the database. Such a test face is generally labeled as authentic face. On the other hand if no such peak is found in the correlation plane the corresponding face images are labeled as impostor. Fig.(1.1) shows the nature of typical correlation plane in response to authentic and impostor face images. The
correlation output is searched for peak and the relative height of this peak is analyzed to determine whether the test face is recognized or not. Fig.(1.2) describes pictorially how the frequency domain correlation technique is carried out for face recognition using correlation filter. As shown in Fig.(1.2) the information of $N$ number of training images from $k$th face class ($k \in C$), out of total $C$ number of face classes for a given
Figure 1.2: Basic frequency domain correlation technique for face recognition

database, are Fourier transformed to form the design input for \( k \)th correlation filter. In ideal case a correlation peak with high value of peak to sidelobe ratio (PSR)\(^3\) is obtained, when any Fourier transformed test face image of \( k \)th class is correlated with \( k \)th correlation filter.

PSR \(^4\) is measured from the correlation plane image. A rectangular region (say \( 20 \times 20 \) pixels) centered at the peak is extracted and used to compute PSR. A \( 5 \times 5 \) rectangular region centered at the peak is masked out and the remaining annular region, shown in Fig.(1.3), defined as the side lobe region, is used to compute the mean and standard deviation of the side lobes. The peak-to-side lobe ratio (PSR) is then calculated as,

\[
PSR = \frac{\text{peak-mean}}{\text{standard deviation}} \quad (1.1)
\]

The major correlation peak is below the threshold PSR value, if the test face image belongs to other class i.e. say, \( j \)th class where, \( j \in C \) and \( j \neq k \). Evidently, the
Figure 1.3:  
Pictorial representation of PSR metric evaluation from correlation plane output.

performance of the system in terms of recognition rate depends on the design of the correlation filter.

The process given in Fig.(1.2) can be mathematically summarized. Let $X$ and $H$ denote the 2D discrete Fourier transforms (DFTs) of 2D image $X$ and 2D filter $H$ in spatial domain respectively, and let $X_i$ is the $i$th Fourier transformed test image of dimension $d_1 \times d_2$. The correlation output $G_i$ in space domain in response to $i$th image for the filter $H$ can then be expressed as the inverse 2D DFT of frequency domain conjugate product as,

$$G_i = FFT^{-1}[X_i \circ H^*], \quad G_i \in \mathbb{R}^{d_1 \times d_2} \quad (1.2)$$

where, $\circ$ represents the element wise array multiplication, $^*$ stands for complex conjugate operation, and FFT is an efficient algorithm to perform DFT.

The approach of using correlation filters offers several advantages. 1) it has built-in shift invariance, 2) correlation filters are based on integration operation and thus offer graceful degradation of any impairment to the test face image, 3)correlation filters can be designed to exhibit attributes such as noise tolerance and high ability.
for discrimination and 4) finally, design of correlation filter is derived from closed form expressions and thus physically realizable.

1.2 A brief review on correlation filters

Development of correlation filters can be broadly categorized into two different classes: 1) linear constrained correlation filters and 2) linear unconstrained correlation filters. Constrained correlation filters are designed by specifying the output of filters for each training image. For \( N \) training images, this results in \( N \) constraints, which are typically much less than the number of free parameters, called the dimensionality of the filter. For this reason, many of these designs optimize some filter performance criterion while satisfying \( N \) constraints. General form of a constrained linear filter \( h \) is given by,

\[
h = \bar{Q}^{-1}A(A^+\bar{Q}^{-1}A)^{-1}u
\]  

(1.3)

where, \( A \) is a matrix whose \( N \) columns are \( N \) frequency-domain training images \( (x_i) \) in vector form, \( \bar{Q} \) is a diagonal matrix, and \( u \) is an \( N \times 1 \) vector for the specified correlation output values for each training image.

Special cases of \( \bar{Q} \) result in well-known filter designs. These cases are listed in Table(1.1).

<table>
<thead>
<tr>
<th>Filter Type</th>
<th>Value of ( \bar{Q} ) from Eq.(1.3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECPSDF[5]</td>
<td>( \bar{Q} = \bar{I} ) (Identity matrix)</td>
</tr>
<tr>
<td>MVSDF[6]</td>
<td>( \bar{Q} = \bar{O} )</td>
</tr>
<tr>
<td>MACE[7]</td>
<td>( \bar{Q} = \bar{D} ), ( \bar{D} = \sum_{i=1}^{N} \bar{D}_i ), where, ( \bar{D}_i = \bar{X}_i\bar{X}_i^* )</td>
</tr>
<tr>
<td>OTSDF[8]</td>
<td>( \bar{Q} = \alpha\bar{O} + \sqrt{1 - \alpha^2}\bar{D} )</td>
</tr>
<tr>
<td>MINACE[9]</td>
<td>( \bar{Q} = \max(\alpha\bar{O}, \sqrt{1 - \alpha^2}\bar{D}_1, \ldots, \sqrt{1 - \alpha^2}\bar{D}_N) )</td>
</tr>
</tbody>
</table>
In Table(1.1), if $\bar{Q}$ is replaced by $\bar{I}$ (identity matrix), the design equation reduces to ECP-SDF filter. The drawback of the ECP-SDF is that it cannot tolerate significant input noise. To achieve robustness to noise, minimum variance synthetic discriminant function (MVSDF) filter is introduced [6]. Design equation of MVSDF is obtained by replacing $\bar{Q}$ by $\bar{O}$ where $\bar{O}$ is a diagonal matrix containing the power spectral density of the noise. Thus, MVSDF minimizes the correlation output noise variance (ONV) while satisfying the correlation peak amplitude constraints. MVSDF controls only one point in the correlation map like the ECP-SDF, and the variance of the noise matrix must be known beforehand in order to design the filter. However, if the latter is known exactly, MVSDF is impractical because it requires inverting a large noise covariance matrix [10, 1].

Minimum average correlation energy (MACE) filter was an attempt to control the entire correlation plane, where reduced correlation function levels are reduced at all points except at the origin of the correlation plane and a sharp correlation peak is obtained [7]. It has been shown that the operation is equivalent to minimizing the energy of the correlation function while satisfying intensity constraints at the origin. Closed form solution of MACE filter is obtained by replacing $\bar{Q}$ with $\bar{D}$, as shown in Table(1.1), where $\bar{D}_i$ is the power spectrum of the $i$th training image, and $\bar{D}$ contains the average training power spectrum, and in $\bar{X}_i = diag\{x_i\}$. However, MACE filter often suffers from two major drawbacks. Firstly, there is again no built-in immunity to noise. Second, the MACE filter is often excessively sensitive to intra-class variations. Nevertheless, this filter establishes the utility of frequency domain design approach for pattern recognition.

The optimal trade-off synthetic discriminant function (OTSDF) filter[8] includes a trade-off parameter $\alpha$ that allows the user to emphasize low output noise variance (ONV) ($\alpha$ closer to 1) or low average correlation energy (ACE) ($\alpha$ closer to 0). Setting $\alpha = 1$ yields MVSDF having minimum ONV but this usually exhibits broad
correlation peak. In contrast, setting $\alpha = 0$, yields MACE filter, which has minimum ACE and produces sharp peak. However, MACE filter is highly sensitive to noise and distortion.

The minimum noise and correlation energy (MINACE) filter [9] achieves an alternative compromise between these two extremes by using an envelope equal to or greater than the noise in power spectra of training image at each frequency. It may be noted that the trade-off parameter $\alpha$ appearing in the MINACE formulation in Table(1.1) is not a part of the traditional MINACE filter design as reported in [9]; rather, the value of $\bar{O}$ is varied directly, since the input noise level is typically unknown. This difference is merely semantic; in practice, the same effect is achieved by varying either $\bar{O}$ or $\alpha$. In both OTSDF and the MINACE filter designs, a single parameter $\alpha$ simultaneously accomplishes both these goals, because both the input noise level and the trade-off can be effected by scaling $\bar{O}$ relative to $\bar{D}$.

Studies have shown that hard constraints on correlation values at the origin are not only unnecessary but can be counter-productive [11]. Relaxing or removing such constraints might lead to larger filter solution space. Also, the matrix inversion in the constrained design may be ill-conditioned, when highly similar training images are included. For these reasons, several unconstrained linear filter designs have been proposed. These designs maximize some measure of the average output on true-class training images while minimizing other criteria such as ONV and ACE. The maximum average correlation height (MACH) filter[10] is one such design, which achieves distortion tolerance by maximizing the similarity of the shapes of true-class correlation outputs over the training images. This maximization is realized by minimizing a dissimilarity metric known as the average similarity measure (ASM) for true class images. Design equation of MACH filter is given in Table(1.2) where $\bar{S}$ represents the measure of ASM.

Replacing $\bar{S}$ by $\bar{D}$ results in unconstrained MACE (UMACE) filter[10] solution,
Table 1.2: Unconstrained linear filter designs

<table>
<thead>
<tr>
<th>Filter Type</th>
<th>Filter $h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACH[10]</td>
<td>$S^{-1}m$</td>
</tr>
<tr>
<td>UMACE[10]</td>
<td>$D^{-1}m$</td>
</tr>
<tr>
<td>UOTSDF[12]</td>
<td>${\alpha \hat{O} + \beta \hat{D}}^{-1}m$</td>
</tr>
<tr>
<td>OTMACH[13]</td>
<td>${\alpha \hat{O} + \beta \hat{D} + \gamma \hat{S}}^{-1}m$</td>
</tr>
<tr>
<td>EMACH[14]</td>
<td>Dominant eigenvector of ${\alpha \hat{I} + (1 - \alpha^2)^{1/2} \hat{S}}^{-1} \hat{C}^d$</td>
</tr>
<tr>
<td>EEMACH[15]</td>
<td>Dominant eigenvector of ${\alpha \hat{I} + (1 - \alpha^2)^{1/2} \hat{S}}^{-1} \hat{C}^d$</td>
</tr>
</tbody>
</table>

given in Table(1.2). The unconstrained OTSDF (UOTSDF) filter [12] is a similar design that minimizes a trade-off between true-class ACE and ONV (as in the OTSDF design). The optimal trade off approach in introduced in [13] by relating correlation plane metrics which resulted in OTMACH filter as given in Table(1.2), where, $\alpha$, $\beta$ and $\gamma$ are the non-negative optimal trade off (OT) parameters.

In addition to OTMACH filter different variations of MACH filters were proposed. In [14] an extended MACH (EMACH) filter design is addressed by reducing the dependence on the average training image $m$. A tunable parameter $\beta$ (given in Table(1.2)) is used to control this reduction. Two new metrics are used in the design: (1) the all-image correlation height (AICH), which takes into account of the filter output on $m$ as well as on individual training images, and (2) a modified average similarity measure (MASM), which measures the average dissimilarity to the optimal output shape. This optimal shape reduces the dependence on $m$ as realized by new AICH metric. EMACH filter design also includes ONV criterion to help in maintaining noise tolerance. A trade-off parameter $\alpha$, given in Table(1.2), is used to control the relative importance of the ONV and MASM criteria, where higher values of $\alpha$ correspond to greater emphasis on ONV and vice versa. If the covariance matrix $\hat{C}^\beta$ is approximated by only its dominant eigenvectors, the eigenvalues yield a new matrix $\hat{C}^d$. The result-
ing filter solution is referred to as the eigen-extended MACH (EEMACH) filter [15].

A linear correlation output is an array of scalar output values from a linear discriminant applied to the input image at every shift. While limited in capability by their linear nature, linear correlation filters have the important advantage of efficient frequency domain computation. In addition to linear correlation filters several nonlinear correlation filters were developed. Design equations of some of the filters are given in Table (1.3). Special cases of nonlinear discriminant functions have been proposed in which some attractive computational properties of linear filters are retained by specialized implementation schemes.

Nonlinear correlation filters are designed in two ways. One class of design, termed as quadratic correlation filters (QCFs), is obtained by solving for a quadratic discriminant function in \(d\)-dimensional space, where \(d\) is the number of pixels in the image. This quadratic discriminant can then be efficiently implemented as a set of linear filters via eigen decomposition. Several methods were proposed for solving the diagonal matrix in QCF design such as 1) subspace quadratic synthetic discriminant functions (SSQSDFs) [16], 2) Rayleigh quotient quadratic correlation filters (RQQCFs) [17], 3) minimum variance quadratic synthetic discriminant functions (MVQSDFs) [16] and 4) QCFs based on the Fukunaga Koontz transform [17].

In contrast, in the second type of design, termed as polynomial correlation filters (PCFs), are sets of linear filters applied to multichannel input images, whose outputs are subsequently summed to form a single output. Two variants of PCF are proposed, 1) Constrained PCF (CPCF) [28] where CPCF design minimizes a weighted sum of ONV and ACE analogous to the OTSDF design, using the trade-off parameter \(\alpha\) in a similar manner, 2) Unconstrained PCF (UPCF) [29], where UPCF design maximizes the average correlation height (ACH), while minimizing a weighted sum of ONV and ACE.

In addition to above mentioned correlation filters some other correlation filters are
1.2. Chapter 1: A brief review on correlation filters

Table 1.3: Some advanced correlation filters

<table>
<thead>
<tr>
<th>Filter</th>
<th>Design Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMACH[18, 14]</td>
<td>$h = {\delta \Omega + \alpha \bar{O} + \beta \bar{D} + \gamma \bar{S}}^{-1}m$ where, $\Omega$ is $d^2 \times N$ matrix with rank $N$.</td>
</tr>
<tr>
<td>WaveMACH[19]</td>
<td>$h = {\bar{S}^{-1}m}</td>
</tr>
<tr>
<td>Log-WaveMACH[20]</td>
<td>$h = {\bar{S}^{-1}m}</td>
</tr>
<tr>
<td>ARCF[21]</td>
<td>$h = (\bar{D} + \epsilon \bar{I})^{-1}\overline{X}[\overline{X}^+(\bar{D} + \epsilon \bar{I})^{-1}\overline{X}]u$ where $\epsilon = 0$ indicates MACE filter and $\epsilon = \infty$ represents SDF filter.</td>
</tr>
<tr>
<td>CMACE[22]</td>
<td>$h = V^{-1}A{A^+V^{-1}A}^{-1}u$ in feature space</td>
</tr>
<tr>
<td>ActionMACH[23]</td>
<td>3D version of MACH filter.</td>
</tr>
<tr>
<td>ASEF[24]</td>
<td>$H = \frac{G_i}{X_i}$, where $G_i = FFT{exp{(m-m_i)^2+(n-n_i)^2}}$, is transformed Gaussian function at target location$(m_i, n_i)$ and $\sigma \triangleq$ standard deviation.</td>
</tr>
<tr>
<td>MOSSE[25]</td>
<td>$H = \sum_i \frac{G_i}{X_i X_i^*}$, for single training image</td>
</tr>
<tr>
<td>MMCF[26]</td>
<td>$h = {(I - \lambda)\left(\bar{D} - \bar{A}A^+\right)}^{-1/2}\bar{A}\alpha$ where $\bar{A} = [\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_N]$,</td>
</tr>
<tr>
<td></td>
<td>and, $\bar{x} \triangleq {(I - \lambda)\left(\bar{D} - \bar{A}A^+\right)}^{-1/2}x_i$ and $\alpha[26]$ is evaluated by sequential minimum optimization[27] technique.</td>
</tr>
</tbody>
</table>

Table (1.3) includes design equations of generalized MACH(GMACH) filter, wavelet modified MACH (WaveMACH) filter, log-transformed WaveMACH(Log-WaveMACH) filter, Action MACH filter, average exact synthetic function (ASEF) filter, correentropy MACE (CMACE) filter, adaptive robust correlation filter (ARCF), minimum output sum of squared error (MOSSE) filter, maximum margin correlation filter (MMCF). Detailed discussions on each of them are beyond the scope of the present work.
1.3 Mathematical background of some extensively used correlation filters

1.3.1 SDF filter design

Traditionally in the design of SDF-type correlation filters, linear constraints are imposed on the training images to yield a known value at specific locations in the correlation plane. The classical SDF\cite{5, 30} filter is designed as a two-class problem, where the correlation values at the origin is set to 1 (may be selected to other values for multi-class problem) for training images from one class, generally authentic or true class, and to 0 for training images from other class or false class. SDF can be formulated by a single matrix-vector equation denoted as,

\[ A^+ h = u \]  

(1.4)

where, \( A = [x_1, x_2, ..., x_N] \) is a \( d \times N \) matrix with \( N \) training Fourier transformed vectors as its columns, and \( u = [u(1), u(2), ..., u(N)]^T \) is an \( N \times 1 \) vector containing the desired peak values at the origin of correlation plane for the desired class, and \( d \) is the total number of pixels present in one image. \( h \) is the desired filter of size \( d \times 1 \) and the superscript + indicates the complex conjugate transpose.

Considering that the conventional SDF is matched to a composite image \( h \) given by,

\[ h = A c \]  

(1.5)

where the coefficient vector \( c = [c(1), c(2), ..., c(N)]^T \) of the linear combination are chosen to satisfy the deterministic constraints indicated in Eq.(1.4).

Thus, from Eq.(1.5) and Eq.(1.4) the SDF filter \( h_{SDF} \) can be formulated by substituting \( c \) as,

\[ h_{SDF} = A (A^+ A)^{-1} u \]  

(1.6)
1.3.2 MACE filter design

MACE filter is designed to ensure sharp correlation peak and to allow easy detection in the full correlation plane as well as to control the correlation peak value. To achieve good detection, it is necessary to reduce the levels of correlation function at all points except at the origin of the correlation plane. Specifically, the value of the correlation function must have an user specified value at the origin but the value is free to vary elsewhere. This is equivalent to minimizing the energy of the correlation function while satisfying intensity constraints at the origin. The correlation peak amplitude constraint for MACE filter is same as that considered in case of SDF filter given in Eq.(1.4). The correlation plane in response to $x_i$ for the MACE filter $h$ can be expressed in matrix-vector form as,

$$g_i = \bar{X}_i^* h$$  \hspace{1cm} (1.7)

where $\bar{X}_i$ represents a $d \times d$ diagonal matrix containing $i$th training vector $x_i$ along its diagonal.

Hence the energy of the $i$th correlation plane can be formulated as,

$$|g_i|^2 = |\bar{X}_i^* h|^2$$  \hspace{1cm} (1.8)

where, $\bar{D}_i = \bar{X}_i \bar{X}_i^*$ is a $d \times d$ diagonal matrix containing power spectrum correspond to $x_i$.

For all $i = 1, 2, \cdots, N$ the ACE is given by,

$$\text{ACE} = \frac{1}{N} \sum_{i=1}^{N} |g_i|^2 = \frac{1}{N} \sum_{i=1}^{N} h^+ \bar{D}_i h = h^+ \bar{D} h$$  \hspace{1cm} (1.9)

where, $\bar{D}$ represents $d \times d$ diagonal matrix containing average power spectrum along
its diagonal and is given by,

$$\bar{D} = \frac{1}{N} \sum_{i=1}^{N} \bar{D}_i = \frac{1}{N} \sum_{i=1}^{N} \bar{X}_i \bar{X}_i^*$$  \hspace{1cm} (1.10)

Therefore to synthesize MACE filter, attempt is made to minimize ACE given in Eq. (1.9) while meeting the linear constraints in Eq. (1.4). The solution to this problem can be founded by using the method of Lagrange multipliers. Similar derivation of the constrained optimization problem can be found in [7, 6]. The optimum solution of Eq. (1.9) is obtained as,

$$h_{MACE} = \frac{1}{N} \sum_{i=1}^{N} \bar{D}_i = \frac{1}{N} \sum_{i=1}^{N} \bar{X}_i \bar{X}_i^*$$  \hspace{1cm} (1.11)

1.3.3 MVSDF filter design

MVSDF minimizes the correlation output noise variance (ONV) in $h^+ \bar{O} h$, where $\bar{O}$ is the diagonal matrix whose diagonal entries are the noise power spectral density while satisfying the constraints of correlation peak amplitude. The solution of MVSDF is expressed as,

$$h_{MVSDF} = \bar{O}^{-1} A (A^+ \bar{O}^{-1} A)^{-1} u$$  \hspace{1cm} (1.12)

1.3.4 Optimal trade-off (OTF) filter design

Due to the minimization criteria of ACE of MACE filter, sharp correlation peak is possible by suppressing the side lobes as this filter emphasizes the high frequency components. However, MACE filter can result in poor intra-class recognition of images which are not included in the training set. Moreover, MACE filter is often excessively sensitive to noise as there is no in-built immunity to noise. To get a sharp correlation peak with suppressed noise, MACE filter is combined with MVSDF. The technique resulted in the design of an optimal trade-off function (OTF) [8]. The optimum
solution of OTF is given by,

\[
h_{\text{OTF}} = \bar{T}^{-1}A(A+\bar{T}^{-1}A)^{-1}u \tag{1.13}
\]

where \( \bar{T} = \alpha \bar{D} + \sqrt{1 - \alpha^2} \bar{O} \), \( 0 \leq \alpha \leq 1 \). \( \alpha \) is used as controlling trade-off parameter, i.e. for \( \alpha = 0 \) leads to MVSDF and \( \alpha = 1 \) leads to MACE filter.

### 1.3.5 MACH filter design

Another approach of designing correlation filter is to remove the hard constraint at the correlation plane and hence in general, these type of filters are termed as unconstrained type correlation filters. The unconstrained correlation filter offers improved distortion tolerance as during the design phase of such filters the training images are not treated as deterministic representations of the image, but as samples of a class whose characteristic parameters are used in encoding the filter[10]. In order to achieve this, an optimal shape of correlation plane \( f \) (in vector form of dimension \( d \times 1 \) ) is required and the deviation of \( i \)th correlation plane in Eq.(1.7) from the ideal shape vector \( f \) will be minimized. This deviation can be quantified in terms of average squared error (ASE) as,

\[
\text{ASE} = \frac{1}{N} \sum_{i=1}^{N} |g_i - f|^2 \tag{1.14}
\]

Minimizing ASE by setting \( \nabla_f(\text{ASE}) = 0 \) the optimum shape vector is obtained as

\[
f_{\text{opt}} = \frac{1}{N} \sum_{i=1}^{N} g_i = \frac{1}{N} \sum_{i=1}^{N} \bar{X}_i^*h = \bar{M}^*h \tag{1.15}
\]

where, \( \bar{M} = \frac{1}{N} \sum_{i=1}^{N} \bar{X}_i \). Eq.(1.15) represents the average correlation plane and \( \bar{M} \) is the average training image expressed in diagonal form. The average correlation plane \( \bar{M}^*h \) offers minimum ASE out of all possible reference shape and hence least distortion in squared error sense is achieved.
The average similarity measure (ASM) is obtained from Eq. (1.14) by substituting \( f = f_{\text{opt}} = \bar{M}^* h \) and \( g_i = X_i^* h \) as,

\[
\text{ASM} = \frac{1}{N} \sum_{i=1}^{N} |\bar{X}_i^* h - \bar{M}^* h|^2 = h^* \bar{S} h
\]

(1.16)

where,

\[
\bar{S} = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{M})(X_i - \bar{M})^*
\]

(1.17)

is a \( d \times d \) diagonal matrix measuring the similarity of the training images to the class mean in the frequency domain.

Another filter design criteria is formulated by maximizing the correlation peak intensity of average correlation plane instead of specifying values at the correlation planes for each training images. The peak intensity of the average correlation plane is written as,

\[
|g(0, 0)|^2 = |m^* h|^2 = h^* mm^* h
\]

(1.18)

where,

\[
m = \frac{1}{N} \sum_{i=1}^{N} x_i
\]

(1.19)

represents the mean vector corresponding to training vectors \( x_i \) for all \( i = 1, 2, \ldots, N \).

The behavior of the average correlation plane is explicitly optimized by minimizing ASM and maximizing peak value. Hence the criteria to be optimized to improve distortion tolerance is given by,

\[
J(h) = \frac{h^* mm^* h}{h^* \bar{S} h}
\]

(1.20)

where \( J(h) \) is called the Rayleigh quotient.

The filter of interest \( h \) maximizes this criterion and thus is called a maximum average correlation height (MACH) filter. The MACH filter maximizes the relative
1.3. Chapter 1: Mathematical background of correlation filter

height of average correlation peak with respect to expected distortions. Since \( J(h) \) in Eq.(1.20) results in small denominator, the filter \( h \) reduces ASM given in Eq.(1.16). The optimum filter is found by setting the gradient of \( J(h) \) with respect to \( h \) to zero and is given by [10]

\[
h_{\text{MACH}} = \bar{S}^{-1} m
\]  

(1.21)

where, \( h_{\text{MACH}} \) is the desired MACH filter, the transformed class dependent mean image.

1.3.6 UMACE filter design

Replacing \( \bar{S} \) by \( \bar{D} \) in Eq.(1.21), the closed form solution of UMACE filter is obtained as,

\[
h_{\text{UMACE}} = \bar{D}^{-1} m
\]  

(1.22)

1.3.7 OTMACH filter design

It has been shown in [13, 10] that MACH filter and its other variants, most notably optimal trade off MACH (OTMACH) filter is very powerful correlation filter algorithm. In practice, other performance measures like ACE, ONV are also considered to balance the system performance for different application scenario. Optimal trade off approach is introduced in [13] by relating correlation plane metrics such as ONV, ACE, ASM and ACH. The performance of OTMACH filter is improved by minimizing the energy function \( E(h) \) of the correlation filter \( h \), given by,

\[
E(h) = \alpha(ONV) + \beta(ACE) + \gamma(ASM) - \delta(ACH) \\
= \alpha h^+ \bar{O} h + \beta h^+ \bar{D} h + \gamma h^+ \bar{S} h - \delta |m^+ h|^2
\]  

(1.23, 1.24)
These considerations lead to the expression for OTMACH filter as,

$$h_{\text{OTMACH}} = \frac{m}{\alpha O + \beta D + \gamma S} \quad (1.25)$$

where, $\alpha$, $\beta$ and $\gamma$ are the nonnegative optimal trade off (OT) parameters.

## 1.4 Physical requirements in designing correlation filters

In physical term, the correlation plane is treated as a new linearly transformed image generated by the filter in response to input image. Therefore, not only the correlation peak but also the entire correlation plane needs to be tailored for better performance and hence ONV, ACE, ASM and ACH have to be properly tuned with the help of parameters $\alpha, \beta, \gamma$ and $\delta$. In general, for face recognition applications, the values of non-negative constants $\alpha, \beta$ and $\gamma$ are chosen to tailor the filter’s performance under noise and variations in illumination conditions and distortions in face images. The value of $\delta$ in minimizing energy function in Eq.(1.24) modifies the peak height at the correlation plane to ensure good correlation and therefore must dominate the other performance criteria. Minimization of ACE is required since the low value of ACE emphasizes the high frequency components of images. The control of tradeoff parameters is possible in OTMACH filter [4], which exhibits significantly better recognition performance than other filters. Easy detection of the correlation peak, better distortion tolerance and the ability to suppress the clutter noise are the three basic criteria those are fulfilled by using OTMACH. On the other hand, OTSDF filter includes a tradeoff parameter that takes high value of $\alpha$ close to 1 and $\beta(=\sqrt{1-\alpha^2})$ close to 0, so as to emphasizes on high value of ONV and low value of ACE. Similarly, MVSDF filter is designed for minimum ONV but usually exhibits broad correlation peaks. Setting $\alpha = 0$, a MACE filter is designed. Though MACE produces a sharp peak, yet it is highly sensitive to noise and distortion and therefore its usefulness for ro-
bust face authentication in presence of variations in illumination condition is limited [31, 32]. Clutter rejection can be achieved by reducing the dependence on average training images by including ONV criteria. A tunable parameter $\beta$ is used to control the performance.

1.5 A review of correlation filter applications in face recognition

Human face recognition has been studied for more than twenty years[33, 34, 35, 36]. Unfortunately developing a computational model of face recognition is quite difficult, because faces are complex and produce multi-dimensional visual stimuli. Therefore, face recognition is a very high level computer vision task. Based on the basic procedure adopted, a face detection method can be classified as (a) holistic or image based and (b) methods based on facial feature extraction. In the simplest version of the holistic approaches, the image is represented as a 2D array of intensity values and recognition is performed by direct correlation comparisons between the input face and all the other faces in the database. Major sub-areas under image based methods can be categorized as statistical methods[37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48], subspace methods[49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59] and AI approaches where neural networks and machine learning techniques[60, 61, 62, 63, 64] are used to recognize faces. Feature extraction and computationally efficient matching require the application of efficient clustering and classification algorithms to work on standard face databases. Feature-based approaches first process the input image to identify and extract (and measure) distinctive facial features such as the eyes, mouth, nose, etc., as well as other fiducial marks, and then compute the geometric relationships among those facial points, thus reducing the input facial image to a vector of geometric features. Standard statistical pattern recognition techniques are then employed to match faces using these measurements. Sub-areas under feature extraction methods
are skin color and texture based segmentation[65], deformable template matching[66],
snake models[67], feature searching and constellation analysis and other several feature
based face recognition techniques can be found in[68, 69, 70, 71]. The list is not
complete as some techniques use a mixture of stated categories.

In majority cases of face recognition techniques, the analysis and processing are
carried out on the spatial representation of the face image i.e., the intensity values of
the face image. However, in this work attempts have been made to evolve efficient
recognition techniques in frequency domain, as recent advancements in the techniques
of using of frequency domain for face recognition have shown promising results. There-
fore further review in space domain analysis is not carried out. The Table(1.4) gives
a short snap of major work done in space domain face recognition techniques.

**Table 1.4: Categorization of still face recognition techniques**

<table>
<thead>
<tr>
<th>Approach</th>
<th>Representative work</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Holistic methods</strong></td>
<td></td>
</tr>
<tr>
<td>Principal component analysis(PCA)</td>
<td>Direct application of PCA</td>
</tr>
<tr>
<td></td>
<td>Eigenfaces [38, 39, 40]</td>
</tr>
<tr>
<td>Probabilistic eigenfaces</td>
<td>Two-class problem with</td>
</tr>
<tr>
<td></td>
<td>probability measure[47]</td>
</tr>
<tr>
<td>Fisherfaces</td>
<td>Fisher Linear Discriminant</td>
</tr>
<tr>
<td></td>
<td>on eigenspace[44, 72]</td>
</tr>
<tr>
<td>Laplacianfaces</td>
<td>Face recognition by Laplacianfaces[55]</td>
</tr>
<tr>
<td>2D PCA</td>
<td>Two dimensional PCA for</td>
</tr>
<tr>
<td></td>
<td>face recognition[73]</td>
</tr>
<tr>
<td><strong>Feature-based methods</strong></td>
<td></td>
</tr>
<tr>
<td>Pure geometry methods</td>
<td>Earlier method [74]; recent</td>
</tr>
<tr>
<td></td>
<td>methods[75, 76]</td>
</tr>
<tr>
<td>Dynamic link architecture</td>
<td>Graph matching method[77]</td>
</tr>
<tr>
<td>Hidden Markov model(HMM)</td>
<td>HMM methods [78, 79]</td>
</tr>
<tr>
<td><strong>Other representations</strong></td>
<td></td>
</tr>
<tr>
<td>Feature lines</td>
<td>Point-to-line distance based[80]</td>
</tr>
<tr>
<td>Evolution pursuit</td>
<td>Enhanced GA learning [62]</td>
</tr>
<tr>
<td>Independent Component Analysis(ICA)</td>
<td>ICA-based feature analysis [81]</td>
</tr>
<tr>
<td>Convolution Neural Network(CNN)</td>
<td>Self organizing Map learning based</td>
</tr>
<tr>
<td></td>
<td>CNN methods [82]</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>Face recognition using SVM[83, 84]</td>
</tr>
</tbody>
</table>
A review of correlation filter applications in face recognition

Different constrains on the face images such as illumination variations, occlusion, expression variations yielded many types of correlation filters. In most of the cases, MACE, UMACE and their different phase extensions are used for verification purpose. Several noticeable works in face recognition using correlation filters can be found in [85, 86, 87, 32, 88, 89, 90, 91, 92, 93].

In [94], MACE filter is synthesized with some training images and applied over AMP facial expression database [95] where overall 0.1% equal error rate (EER) is achieved. A comparative performance of MACE filter and individual eigenface subspace method (IESM) for face recognition in terms of margin of separations is presented in [96]. In [97] an efficient method of designing the MACE filter is proposed where, the complexity of filters is reduced without sacrificing the system performance. Therefore even on limited resource platforms, the algorithm can perform face localization and recognition. An idea of incrementally updating the unconstrained filters for limited memory devices is also successfully proposed in [98] with incremental updating of single training image one at a time. This updating method iteratively selects which of the captured images, during the enrollment stage. Boosting the performance of MACE filter in illumination invariant face recognition using logarithmic transformation is proposed in [99]. It has been shown here that using this transformation the MACE filter gives better discrimination between authentic and impostor PSRs. An approach to encrypting MACE filter is reported in [100]. It has been shown here that an arbitrary random convolution kernel can be used. This helps to guard against the types of attacks where the attacker might try to intercept the decrypted filter during the verification stage.

In [101], a principal components analysis (PCA) is run on the phase spectrum of the training images in the Fourier domain allowing the phase information as a spanning linear subspace. The primary advantage of using a subspace to represent the target instead of a single filter is that it represents a larger set of target variations.
This results is higher PSRs than conventional MACE filter.

A successful combination of support vector machine (SVM) with advanced correlation filter to produce maximum margin SVM correlation filter is proposed in [102]. It gives more control over the relationship of peaks to sidelobes in the training correlation planes. In addition, it inherently minimizes the sensitivity to additive white noise by minimizing filter energy subject to the existence of a margin.

Illumination invariant face recognition and impostor rejection using different minimum noise and correlation energy (MINACE) filter algorithms is proposed in [103]. Two different MINACE filter formulations i) spectral envelope and ii) additive spectrum and two different correlation plane metrics i) peak and ii) peak-to-correlation plane energy ratio (PCER) were used to create face recognition systems that function with illumination variations. Good performance scores were presented for both face verification and identification on PIE database.

A different approach of using correlation filter is suggested in [104], where a quaternion array is developed from wavelet decomposition and used in synthesizing the correlation filter. By using the quaternion correlation filter to model the inter-sub-band characteristics as well as the intra sub-band characteristics a decomposed representation is developed. The numerical experiments on PIE data set shows that the proposed method achieve improvement when trained by a single near frontal lighting mug-shot image and tested on unknown, variable lighting face images. Redundant class-dependence feature analysis (CFA) method for face recognition using correlation filters is proposed in [105]. In this method correlation filters are designed, one for each subject in the generic training set to get a bank of correlation filters. All these filters are used for feature extraction. The nearest neighbor rule is applied to decide on the class label for the test image.

Face class code (FCC) based approach using correlation filter and support vector machine (SVM) is proposed in [106]. This method is used as binary classifiers for
face recognition when the number of the classes is large. FCC is combined with error control codes and better recognition results under variable illumination conditions. Template matching method of correlation filter \([107, 108]\) are used for facial feature extraction, where the cosine distance is measured from a similarity score. This method is successfully experimented over FRGC2.0 dataset. In \([109]\), it has been shown that kernel correlation feature analysis (KCFA) has good representation and discrimination ability for unseen datasets and produces better verification and identification rates on PIE, FERET and AR dataset.

In general, two dimensional(2D) correlation feature analysis (2D-CFA) cannot be used for vectors and \(N^{th} (N \geq 3)\) order tensors. This limitation is overcome by Yan et.al. in \([110]\), where a generalized method of analysis is proposed by using the image data as tensors. The improved recognition rate is obtained by tensor based method in comparison to traditional 2D-CFA for standard face databases. An 1D-CFA is proposed \([111]\) in low dimensional subspace (PCA) instead of 2D-CFA, where peak height is minimized subject to linear constraint. Another research in this area is called correntropy MACE (CMACE) filters \([22]\). In this case, the kernel function is limited to a Gaussian kernel. When combined with the fast Gaussian transform, the technique allows fast approximation of the full correlation output while retaining the increased representational power. In \([22]\) it is shown that, though slow during operation, better face recognition rate is achieved with CMACE filters.

In \([112]\) a comparative study of some recent advanced correlation filters is made to test recognition performance in different situations involving variations in facial expression, illumination conditions and head pose. It demonstrates that it is possible to obtain illumination invariance without using any training images for this purpose. The correlation filter classifiers also has greater robustness and accuracy than traditional appearance-based methods (such as PCA). It has also reported that the phase extended unconstrained MACE filter is the best choice for facial matching.
Adaptive and robust correlation filters (ARCF) is proposed in [21] and describe their usefulness for reliable face authentication using recognition-by-parts strategies. ARCF provide information that involves both appearance and location. The cluster and strength of the ARCF correlation peaks indicate the confidence of the face authentication made, if any. The adaptive aspect of ARCF comes from their derivation using both training and test data, similar to transduction, while the robust aspect benefits from the correlation peak optimization to decrease their sensitivity to noise and distortions.

An approach for face verification using local binary pattern (LBP) operators and optical correlation filters can be found in [113]. LBP is operated on training images to form local binary pattern-unconstrained minimum average correlation energy (LBP-UMACE) filters as an optical correlation filter to enhance recognition rates and reduce error rates simultaneously. Better performance of LBP-UMACE compared with UMACE filters are demonstrated.

In [114], the original idea based on the unconstrained optimal trade-off quaternion filter (UOTQF) is extended and two additional different correlation filters in quaternionic domain are evaluated, i) a phase only quaternion filter (POQF) and ii) separable trade-off quaternion filter (STOQF). Three different quaternion-based correlation filters are designed and conjugated with four face feature extraction methods. Advantage of synthesis correlation filters in quaternionic domain is only one face image of a person is needed for training. Combination of quaternionic representation with a quaternion-based correlation filter confirms good discriminating and illumination invariant properties and an improvement in face recognition accuracy is obtained.

As an extended version of SVM and correlation filter more generalized approach is presented in [26] where maximum margin correlation filter (MMCF) is proposed. It combines the generalization capability of SVM and localization capabilities of correlation filters. MMCF is successfully implemented in different object recognition and
face classification problems.

1.6 Performance of correlation filters in face recognition

In this section some comparative performances of general purpose correlation filters are made for face recognition task under various facial expressions and varying lighting conditions. Correlation filters like MACH, UMACE, OTMACH, quad phase UMACE (QPUMACE) and phase extended UMACE (PEUMACE) are synthesized and tested over several databases\(^1\) including AMP, Cropped YaleB, PIE and AR face datasets. The performance evaluation of different unconstrained correlation filters has been made by setting different optimal trade-off parameters in Eq.(1.25). UMACE filter is designed with the help of the equation, 

\[
h_{\text{UMACE}} = \overline{D}^{-1}m, (\alpha = 0, \beta = 1, \gamma = 0).
\]

Similarly MACH and OTMACH are designed using 

\[
h_{\text{MACH}} = \overline{S}^{-1}m, (\alpha = 0, \beta = 0, \gamma = 1)
\]

and 

\[
h_{\text{OTMACH}} = (\alpha \overline{O} + \beta \overline{D} + \gamma \overline{S})^{-1}m, (\alpha = 0.2, \beta = 0.5, \gamma = 0.3)
\]

respectively, where values of \(\alpha, \beta, \gamma\) are chosen empirically. PEUMACE is obtained as the full phase extension of \(H_{\text{UMACE}}\) and is given by,

\[
H_{\text{PEUMACE}} = e^{j\angle H_{\text{UMACE}}}
\]

(1.26)

In designing QPUMACE filter each element in the filter array will take on \(\pm 1\) for the real component. The imaginary component \(\pm j\) is calculated in the following manner:

\[
H_{\text{QPUMACE}} = \begin{cases} 
+1 & \Re\{H_{\text{UMACE}}(u, v)\} \geq 0 \\
-1 & \Re\{H_{\text{UMACE}}(u, v)\} < 0 \\
+j & \Im\{H_{\text{UMACE}}(u, v)\} \geq 0 \\
-j & \Im\{H_{\text{UMACE}}(u, v)\} < 0 
\end{cases}
\]

\(^1\)Detail database description is given in Appendix-A
1.6. Chapter 1: Performance of correlation filters in face recognition

PSR values are used to test verification accuracy of the above correlation filters for each database.

1.6.1 Test results

During performance analysis of correlation techniques for face recognition two types of tests are performed. (1) Identification test - where the class is labeled based on the filter that scores a relative maximum PSR and (2) Verification test - where an authentic test face images must achieve a score above a preset threshold. Face recognition performance of correlation filters is measured based on verification approach and it is maintained for every experiments on AMP, YaleB, PIE and AR face databases.

1.6.2 Performance evaluation using PSR values

All unconstrained filters are synthesized with same number of training images (1,21,41) from Person-1 of AMP database and tested over the whole database. Fig(1.4) shows the performance of different filters in terms of PSR values. The separation margin between authentic and impostor face image is calculated by subtracting the minimum PSR value of the authentic class and the maximum PSR value among the impostor classes. The largest distance of separation (DoS) is achieved for UMACE filter compared to others. UMACE filter is also tested where the training images (1,21,41) from Person-2 are used. From Fig.(1.5) it is observed that the results degrades as reduced DoS is found for UMACE. Hence the phase extension of UMACE is considered since phase contains more information than magnitude in an image. It is interesting to observe that while the full phase instead of quad phase of UMACE is considered, DoS is increased indicating better performance for face recognition task.

1. the full phase extension of test image is also taken during correlation
1.6. **Chapter 1**: Performance of correlation filters in face recognition

**Figure 1.4**: PSR performance of different unconstrained filter when tested over AMP database and synthesized with Person 1.

**Figure 1.5**: PSR performance of UMACE and its full-phase extended variation. Filters are tested over whole AMP database when synthesized with Person 2.

### 1.6.3 Performance evaluation in terms of %RR and %FAR

Table(1.5) summarizes the % mean recognition rate with corresponding % false acceptance rate (FAR) while the correlation filter’s performance are tested on whole
AMP database. In Table(1.5) the preset threshold is taken as 7. Table(1.6) shows the %mean recognition rate at zero FAR.

**Table 1.5: The performance of different filters in face recognition on AMP facial expression database.**

<table>
<thead>
<tr>
<th>Training images</th>
<th>MACH %rec,%far</th>
<th>UMACE %rec,%far</th>
<th>OTMACH %rec,%far</th>
<th>QPUMACE %rec,%far</th>
<th>PEUMACE %rec,%far</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,2,1,4,1</td>
<td>90.12,0.92,69</td>
<td>99.89,1.76</td>
<td>99.89,2.89</td>
<td>99.37,0.46</td>
<td>100,1.16,9</td>
</tr>
<tr>
<td>3,2,2,8</td>
<td>90.12,0.92,6</td>
<td>99.58,2.48</td>
<td>99.58,3.17</td>
<td>99.16,0.87</td>
<td>99.37,1.89</td>
</tr>
<tr>
<td>46,50,55</td>
<td>64.13,0</td>
<td>99.68,1.69</td>
<td>99.79,2.61</td>
<td>98.75,0.74</td>
<td>99.27,1.75</td>
</tr>
</tbody>
</table>

**Table 1.6: The %mean recognition rate on AMP database obtained by different filters when FAR = 0. % mean recognition rate increases when number of training images increased.**

<table>
<thead>
<tr>
<th>Training images</th>
<th>MACH %rec</th>
<th>UMACE %rec</th>
<th>OTMACH %rec</th>
<th>QPUMACE %rec</th>
<th>PEUMACE %rec</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,2,1,4,1</td>
<td>53.43</td>
<td>97.50</td>
<td>97.61</td>
<td>96.04</td>
<td>97.92</td>
</tr>
<tr>
<td>3,2,2,8</td>
<td>62.68</td>
<td>98.75</td>
<td>99.06</td>
<td>98.12</td>
<td>98.24</td>
</tr>
<tr>
<td>46,50,55</td>
<td>64.13</td>
<td>92.72</td>
<td>92.203</td>
<td>87.00</td>
<td>89.91</td>
</tr>
<tr>
<td>1,2,3,10,7,4</td>
<td>90.85</td>
<td>98.64</td>
<td>98.75</td>
<td>98.33</td>
<td>98.44</td>
</tr>
</tbody>
</table>

In case of performance evaluation of correlation filters on Cropped YaleB database, each filter is synthesized with each subset images and correlated over the whole database. Out of 38 persons a subset of 10 persons are taken for performance evaluation. Hence $10 \times 9 \times 64 = 5760$ number of impostor scores (PSR) and $10 \times 64 = 640$ authentic scores for YaleB are obtained. While testing with PIE database $65 \times 64 \times 21 = 87360$ number of impostor scores and $65 \times 21 = 1365$ authentic scores for each filter are obtained. From the PSR distribution %mean recognition rate are evaluated according to the verification method and the corresponding %FAR are recorded. Table(1.7) summarizes the %mean recognition rate of correlation filters while tested over YaleB database. It is observed from the Table(1.7) that the recogni-

---

1. subset description of images of YaleB database is given in Appendix-A.
tation results are greatly affected according to the choice of different lighting directions. Overall performance of the correlation filters shows that best recognition accuracy is obtained when subset-4 and subset-5 are chosen for training purpose. It is due to the fact these training set has images with wide variation of lighting or in other words these images have illumination distributed evenly over the camera’s visual field.

Table 1.7: The %mean recognition rate along with the %FAR obtained by different filters while the threshold is fixed at 10 (with no illumination compensation).

<table>
<thead>
<tr>
<th>Filters</th>
<th>UMACE</th>
<th>QPUMACE</th>
<th>PEUMACE</th>
<th>OTMACH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%rec,%far</td>
<td>%rec,%far</td>
<td>%rec,%far</td>
<td>%rec,%far</td>
</tr>
<tr>
<td>Subset-1</td>
<td>69.53,0.1042</td>
<td>71.1,0.0174</td>
<td>76.4,0.1215</td>
<td>89.22,16.99</td>
</tr>
<tr>
<td>Subset-2</td>
<td>73.28,0.2431</td>
<td>69.37,0.086</td>
<td>73.9,0.257</td>
<td>89.06,21.2</td>
</tr>
<tr>
<td>Subset-3</td>
<td>87.34,0.43</td>
<td>82.96,0.086</td>
<td>87.5,0.2431</td>
<td>94.68,10.42</td>
</tr>
<tr>
<td>Subset-4</td>
<td>92.81,1.42</td>
<td>87.65,0.26</td>
<td>91.56,0.78</td>
<td>97.65,12.17</td>
</tr>
<tr>
<td>Subset-5</td>
<td>92.18,2.06</td>
<td>73.9,0.69</td>
<td>82.03,1.54</td>
<td>98.59,18.92</td>
</tr>
</tbody>
</table>

It is also observed that OTMACH filter provides better % RR comparing to others but from this result it cannot be concluded that OTMACH gives the best performance as % FAR is very high. Hence by considering both % RR and % FAR it may be noted that the performance of PEUMACE in case of subset-4 only is slightly better than the other filters.

1.6.4 Performance evaluation by receiver operating characteristic (ROC) curves

Another way of observing the performance is by plotting receiver operating characteristics (ROC) curves. The performance of correlation filters can be characterized in terms of the probabilities of correct detection \( P_D \) and probability of false alarm \( P_{FA} \). In general low detection thresholds improve the probability of correct recognition, while large thresholds decrease false alarm probabilities by rejecting erroneous peaks (or specifically PSRs). The relationship of \( P_D \) and \( P_{FA} \) with threshold PSR
can be represented by ROCs. ROCs are calculated with increasing PSRs as threshold. When comparing ROC-curves of different tests, curves for better performance lie closer to the top left corner and the worst case performance is indicated by a diagonal line. The diagonal line represents $P_D = P_{FA}$. The curves nearer to the diagonal line represents the worst detection performance. Fig(1.6) shows that in general the

![ROC plots](image)

**Figure 1.6:** ROC plots of different correlation filters for different subsets. The performance of PEUMACE filter is also observed for different subset training.

average filter performance is best when subset-4 is used as a training set. This can be explained as the subset-4 includes the training images having wide illumination variation comparing to others. Hence any face image that lies in the convex hull of these training images should be perfectly recognized. The correlation filters are also designed with randomly taken 4 and 5 face images from an authentic individual and tested over whole AR face database. Fig(1.7) shows the filter performances on AR faces in terms of ROC curves.
1.7 Chapter 1: Motivation for the present work

Although several advanced correlation filters are discussed in the review works regarding correlation filters and their applications, this thesis considers some of the standard filters like MACH, UMACE, full phase extended UMACE (PEUMACE), quad phase UMACE (QPUMACE), UOTSDF, OTMACH filters. The main objective is to improve the performance of these standard correlation filters for face recognition in the light of test results obtained so far.

It is observed from Table(1.5) and Table(1.6) almost 99% recognition accuracy can be achieved by UMACE filter and its phase extended variations. But the performance of these filters degrades when face images from YaleB database are used. From Fig.(1.6) it is observed that very poor recognition accuracy obtained at $P_{FA} = 0$. This fact is again verified with Table(1.7). % RR with greater than 90% is achieved for subset-4 and subset-5 but simultaneously % FAR is increased. It is due to the fact the correlation filters are not capable of handling the unseen images with drastic lighting variations. Again from Fig(1.7) it may be noted, that the correlation filters do not give an acceptable amount of recognition accuracy (about 70% recognition rate at 0% FAR, while 5 training images are used for synthesis) on AR database. This is because the UMACE type filters are very much sensitive to intra-class variation which results
misclassification i.e. false acceptance rate (FAR) and false rejection rate (FRR) are increased. Hence some modifications are needed on the correlation filters either by preprocessing the raw image data in space domain or in frequency domain so that reduced FAR and FRR could be achieved.

In the above study the trade off parameters of OTMACH filter are chosen with some fixed value. This makes no guarantee that with these values the OTMACH filter gives the optimal performances. Hence the parameter optimization is needed depending on the nature of applications. Including some information of false class images during the synthesis of a correlation filter can provide better intra-class distortion tolerance. To achieve reduced FAR and FRR class compactness is one of the obvious choices. Again the training image selection to synthesize these filters may not be optimum as designed filter may not have all information of illumination variations. To include large illumination variations, high number of correlation filters are required which affects negatively on signal to noise ratio. Hence a systematic investigation is needed to overcome the problem of intra-class variations and to develop a dynamic correlation filtering technique which reduces the requirement of increased number of correlation filters. Further investigations are also needed to improve upon the robustness of the system due to occlusions, illuminations variations and expression variations of the test face image where false acceptance and false rejection may likely to occur.

These facts and test results obtained from standard correlation filters have established the need for further study with an objective of improving the performance of correlation filters suitable for face recognition under several constraints. The modifications may include post and preprocessing either in space or in frequency domain and optimization of parameters of the filters. Further face recognition form the video images is also an important area where correlation based techniques need to be evaluated.
1.8 Organization of the work

Chapter-2: An improved strategy for face recognition using correlation filter under varying lighting conditions and occlusion is proposed. Some modifications have been made on UMACE filter to reduce the intra-class distortion sensitivity. Spatial domain preprocessing is carried out by two convolution kernels for edge enhancement of face images. These convolution kernels are obtained by training a generalized regression neural network using enhanced face features obtained by principal component analysis.

Chapter-3: A preprocessing is evolved in the frequency domain instead of spatial domain. The continuous wavelet transform (CWT) is used for enhancement of edge features in face image which improves discrimination and association of the image during recognition. The modified UMACE filter solution is combined with Mexican hat wavelet for noise tolerant face recognition under poor lighting condition of faces.

Chapter-4: To reduce the misclassification rate during face recognition, the information of both authentic face class and impostor face class images are considered during synthesis of the correlation filter. In addition, further consideration has been made with respect to overall class compactness. An optimized preferential correlation filter is proposed for multi-class face recognition. The optimization of tradeoff parameters considered to design the preferential filter of both constrained and unconstrained type, is carried out by particle swarm technique.

Chapter-5: Projection based correlation filters are developed by class specific subspace analysis. Two types of subspace analysis has been performed. Initially 1D subspace is proposed. In 1D subspace analysis, one dimensional data that results

---

2 This work is published in Proceedings of the Second International Conference on Intelligent Human Computer Interaction, Springer, 2009, IIIT Allahabad, India
4 This work is published in Proceedings of International Conference on Trends in Optics and Photonics, Dec 7-9, 2011, Kolkata, India
5 This work is published in Optics and Laser Technology, Elsevier, vol. 50, pp. 33-42, 2013
6 This work is published in Lecture Notes in Computer Science, Springer, 7143, pp-338-345, 2012
Chapter 1: Organization of the work

from 2D image by lexicographical ordering, creates an extremely large covariance matrix and hence the chance of proper analysis with large number of samples becomes computationally inefficient and difficult. This problem is overcome by 2D-subspace analysis to synthesize the correlation filters for face recognition under different illumination conditions\(^7\).

Chapter-6: A frequency domain nonlinear correlation technique for performance improvement in face recognition under varying lighting conditions is proposed. This technique is based on phase correlation between optimum projecting image correlation filter and optimum reconstructed image correlation filter in class specific subspace operation. Development of nonlinear correlation filters is carried out by exploiting the point wise nonlinearities of image pixels. The optimization is achieved by minimizing the energy at the correlation plane while maximizing the correlation peak\(^8\).

Chapter-7: Simultaneous detection and verification of single individual face from video is proposed by a combination of unconstrained video filter and distance classifier correlation filter\(^9\). Detection of face in video is performed by the synthesized video filter which is a modified version of three dimensional unconstrained optimal tradeoff filter. The probable location of face is identified according to the location of high correlation peak. The region of interest around the correlation peak is detected and the face part is extracted and fed to distance classifier correlation filter (DCCF) for verification. Simulation results on ViDTIMIT video database indicate the effectiveness of the proposed method for simultaneous face detection and verification in real time system.

\(^7\)This work is published in *Opik (International Journal for Light and Electron Optics)*, Elsevier, vol. 124, issue 17, pp. 3173-3179, 2013

\(^8\)This work is accepted in *Pattern Recognition Letters*, Elsevier, http://dx.doi.org/10.1016/j.patrec.2013.10.012