Chapter 2

REVIEW OF DISTORTION INVARIANT FILTERS USED FOR PATTERN RECOGNITION

2.1 Introduction

The role of correlation filter is to find out how similar or different the input function is from the reference function. Filters are generally used in conventional VLC geometry and hybrid digital-optical correlators. The most basic correlation filter is the matched filter [5], which is nothing but the conjugate of the Fourier transform of the image. It gives better performance in the recognition process only if there is a complete match between the input image and the reference image used to generate the filter. For example, when an input pattern is presented with a scale or an angular orientation that is different from those of the pattern to which the filter was synthesized, the response of the correct matched filter is reduced, and errors arise in the pattern recognition process. The degree of sensitivity of a matched filter to scale and rotation depends to a large extent on the structure of the pattern to which it is matched [2], as a filter for the character L is obviously much more rotation-sensitive than that for the letter O. Thus, the matched filter performs poorly when the target image appears with different distortions like rotation, scale changes, variation of illumination, contrast, background noise, etc. The shortcomings of the correlators in recognizing scenes with some distortions led to the concept of distortion invariant correlation filters. Many such filters have been reported which result in enhancement in the correlation performance of the correlator with distorted image as the input. The work on certain iterative algorithms has also been reported to produce optimized distortion invariant filters.

The conventional JTC can give better performance in the recognition process only in case of good quality input having the same scale and orientation with respect to the reference image. To improve the performance of JTC as well as to make it distortion invariant, the introduction of filters in JTC technique has also been reported.

In this chapter, various types of simple as well as advanced correlation filters used in VLC and hybrid digital-optical correlators has been reviewed. The iterative algorithms used to synthesize optimized distortion invariant filters have been discussed. The filters used in JTC architecture to improve its performance have also been included.
2.2 Filters used in VLC and hybrid digital-optical correlators

2.2.1 Simple correlation filters

VanderLugt [5] proposed and demonstrated a technique for synthesis of classical matched filter (CMF), which is the basic and simplest filter used in OPR. A linear space invariant filter is said to be matched to a particular signal \( s(x,y) \), if its impulse response \( h(x,y) \) is given by

\[
h(x,y) = s^*(-x,-y) \tag{2.1}
\]

If the FT of a signal \( s(x,y) = |S(u,v)| \exp\{i\phi_s(u,v)\} \), then CMF is given by

\[
H(u,v) = |S(u,v)| \exp\{-i\phi_s(u,v)\} \tag{2.2}
\]

where \(|S(u,v)|\) is the amplitude and \(\phi_s(u,v)\) is the phase of the FT of \( s(x,y) \). Thus \(|S(u,v)|\) forms the amplitude part and \(\exp\{-i\phi_s(u,v)\}\) forms the phase part of CMF.

The filter can be easily synthesized holographically or digitally and can provide the maximum signal-to-noise ratio (SNR) (defined as the ratio of the average output peak value to its standard deviation) in recognizing a known pattern embedded in additive noise. But CMF is highly sensitive to distortions and is light inefficient as its magnitude response at many frequencies is less than one.

The requirement for the enhancement in the performance motivated the development of many variants of CMF. These include phase-only filters, binary phase-only filters, quad phase filters, complex ternary matched filters, inverse filters, etc. which have been discussed below.

2.2.1.1 Phase-only filters

Oppenheim and Lim [18] showed that phase information in a signal is more important than amplitude information. As a result, there has been an abundance of work on phase-only matched filtering [19-31]. The general finding is that the phase information is considerably more important than the amplitude information in preserving the visual intelligibility of the picture. In a practical correlation system, it is also desirable that the overall light utilization should be as high as possible. Horner and Gianino [19] proposed a phase-only filter (POF), which is defined as (from Eq. 2.2)

\[
POF = \exp\{-i\phi_s(u,v)\} \tag{2.3}
\]
POF has higher optical efficiency or Horner efficiency (defined as the ratio of the total light intensity in output to the total light intensity in input) and better discrimination capability than the CMF. It yields much sharper correlation peaks with fewer and smaller sidelobes, which enables it to detect targets better. However, POF is an all-pass filter and thus has low SNR in the correlation plane.

2.2.1.2 Binary phase-only filters

The phase function of a phase-only filter is continuous. Restricting the filter function to binary values has certain advantages. First, a binary process is easier to control, and the two-phase levels can be set very accurately. Second, the diffraction efficiency of a binary phase grating is generally greater than a continuous phase grating. Also, some SLMs like the magneto-optic SLM are capable of accommodating only binary values [32]. This motivated the work on binary phase-only filters (BPOF) [33-43]. The BPOF takes only two values, -1 and +1, where –1 represents $\exp(i\pi)$ and +1 represents $\exp(0i)$. If $F(u,v)$ is the FT of a function $f(x,y)$, then BPOF is given by [34]

$$BPOF = \begin{cases} +1 & \text{if } \text{Im} \left[ F(u,v) \right] \geq 0, \\ -1 & \text{otherwise} \end{cases}$$

(2.4)

where Im is the imaginary part of $F(u,v)$.

Thus, BPOF requires less computation time, less storage and can be represented on high frame-rate binary SLMs. It has significant capability for recognizing multiple objects by allowing multiplexing of many patterns onto the same filter [36]. But BPOF has low SNR and is not phase-canceling as it has only two-phase values, which results in the shift of correlation peak from the origin. Farn and Goodman [41] proposed a theory for optimal BPOF and described a numerical algorithm to design the optimal BPOF for a given pattern, which directly maximizes the SNR.

2.2.1.3 Quad-phase filters

Connelly and Vijaya Kumar [44] analyzed the effects of quantizing the filter phase. They believed that output SNR could be significantly improved by going from two phases to four phases. This led to the concept of quad-phase filters (QPF) [45] in which the real part can be
either +1 or −1 and imaginary part can also be +1 and −1. However, the QPF led to noise sensitivity problems.

2.2.1.4 Complex ternary matched filters

Dickey et al. [46] proposed a complex ternary matched filter, which reduced the noise sensitivity of QPF. The filter accommodates three values −1, 0, and +1. The inclusion of zero value helps in improving the noise suppression capabilities of the filter.

2.2.1.5 Inverse filters

Amplitude information, if properly incorporated with the phase, can be utilized to improve the discrimination efficiency. Mu et al. [47] emphasized this fact by correlating objects with an inverse filter. If the FT of a signal \( s(x,y) = |S(u,v)| \exp\{i\phi_s(u,v)\} \), then the inverse filter is given by [48]

\[
\text{IF} = |S(u,v)|^{-1} \exp\{-i\phi_s(u,v)\}
\]

where \( |S(u,v)| \) is the amplitude and \( \phi_s(u,v) \) is the phase of the FT of \( s(x,y) \).

An inverse filter provides larger discrimination as well as sharper correlation peaks but has poor optical efficiency and operational difficulty as it becomes indeterminate whenever the function passes through zero.

2.2.1.6 Fractional power filters

Vijaya Kumar and Hassebrook [49] introduced fractional power filters (FPFs) to illustrate trade-offs occurring between the noise tolerance and the correlation peak sharpness as we move from CMFs to POFs and eventually to inverse filters (IFs). FPFs provide a systematic procedure to choose the right compromise between the noise tolerance and the correlation peak sharpness. These are defined as

\[
\text{FPF} = \begin{cases} 
|S(u,v)|^p \exp\{-i\phi_s(u,v)\} & \text{if } |S(u,v)| \neq 0 \\
0 & \text{if } |S(u,v)| = 0
\end{cases}
\]

where \( p \) can take on any real value. The CMF, POF, and IF can be obtained by using \( p = +1, 0, \) and −1 respectively.
2.2.1.7 Other simple correlation filters

Several other correlation filters related to matched filter have also been proposed. Some of them include phase-mostly filters (PMFs) [50], phase-with-constrained-magnitude filters (PCMFs) [51], amplitude modulated phase-only filters (AMPOFs) [52], average-amplitude matched filter [53], and amplitude-compensated matched filter [54].

2.2.2 Distortion invariant filters

The simple correlation filters generally rely on a single reference image and are more sensitive to distortions like rotation, scale changes, variation of illumination, contrast, background noise, etc. This distortion sensitivity severely limits the practical application of correlators. One approach to handle images of a particular target with different distortions is to synthesize matched filters for each of the image. However, the enormous storage and processing requirements of this approach makes it impractical. Thus it is required to design robust correlation filters that can overcome the limitations of matched filters. This led to the design of distortion invariant filters [55-57]. These filters can be designed either by using geometric invariance properties e.g. circular harmonic filters, radial harmonic filters or by training a number of images like synthetic discriminant function filter and its variants, maximum average correlation height filters, distance classifier correlation filters, etc. The output of such filters can be used to achieve three objectives (i) recognize distorted versions of the reference pattern, (ii) behave robustly in the presence of noise and clutter, and (iii) yield a high probability of correct recognition while maintaining a low error rate.

2.2.2.1 Circular harmonic filters

Hsu and Arsenault [58] introduced the circular harmonic component (CHC) filter for in-plane rotation invariant pattern recognition. The filter uses the fact that a general 2-D function \( f(x,y) \) when expressed in polar coordinates as \( f(r,\theta) \), is periodic in the variable \( \theta \) with a periodicity \( 2\pi \). Therefore, it is possible to express \( f(r,\theta) \) by Fourier series as follows

\[
 f (r, \theta) = \sum_{m=-\infty}^{\infty} f_m (r) e^{im\theta} 
\]  

(2.7)

where \( f_m (r) \) is the \( m^{th} \) circular harmonic component and is given by
A target rotated by an angle $\alpha$ can be expressed as

$$f(r, \theta + \alpha) = \sum_{m=-\infty}^{\infty} f_m(r) e^{im\theta} e^{i\alpha}$$  \hspace{1cm} (2.9)

If $f(r, \theta)$ and $f(r, \theta + \alpha)$ be the polar coordinate representation of reference, $f_1(x, y)$ (to which the filter is matched) and rotated target, $f_\alpha(x, y)$ (taken as input) respectively, then the correlation output is given as

$$O_\alpha(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_\alpha(\xi, \eta) f_1^*(\xi - x, \eta - y) d\xi d\eta$$ \hspace{1cm} (2.10)

For an arbitrary value of $\alpha$, the value of correlation at the origin i.e. $x = 0$, $y = 0$, is given as

$$O_\alpha(0, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_\alpha(\xi, \eta) f_1^*(\xi, \eta) d\xi d\eta$$ \hspace{1cm} (2.11)

To show the influence of rotation angle $\alpha$ on the center correlation value, the above equation is expressed in polar form as

$$C(\alpha) = \int_{0}^{2\pi} \int_{0}^{\infty} f(r, \theta + \alpha) f_1^*(r, \theta) d\theta$$ \hspace{1cm} (2.12)

Using circular harmonic expansion of $f(r, \theta)$ i.e. (Eq. 2.7), the above expression can be written as

$$C(\alpha) = \int_{0}^{\infty} \left[ \sum_{m=-\infty}^{\infty} f_m^*(r) \int_{0}^{2\pi} f(r, \theta + \alpha) e^{-im\theta} d\theta \right] dr$$ \hspace{1cm} (2.13)

Also,

$$\frac{1}{2\pi} \int_{0}^{2\pi} f(r, \theta + \alpha) e^{-im\theta} d\theta = f_m(r) e^{im\alpha}$$ \hspace{1cm} (2.14)

Thus, Eq. (2.13) can be written as

$$C(\alpha) = 2\pi \sum_{m=-\infty}^{\infty} e^{im\alpha} \int_{0}^{\infty} \left| f_m(r) \right|^2 dr$$ \hspace{1cm} (2.15)

Eq. (2.15) shows that the center correlation value is the summation of the contributions from all the CHCs. If a particular CHC is used in correlation, the correlation peak intensity will be independent of rotation angle $\alpha$, thus yielding invariance to in-plane rotation.
The circular harmonic filter provides complete in-plane rotation invariance but has poor discrimination ability as it ignores much of the object information by using only single CHC. Several methods have been suggested to improve the discrimination capability of CHC-based filters [59-62]. One of the methods involve the formation of a feature vector [59] based on the responses of a bank of filters, each based on a different circular harmonic number. The resulting feature vector is rotationally invariant and can embody robust information related to target shape, supporting good recognition and discrimination. Arsenault et al. [63] performed a study by combining a few of the CHCs which resulted into limited invariance to in-plane rotation.

For the synthesis of CHC filters, proper choice of center for circular harmonic expansion is required to obtain better results. Sheng and Arsenault [64] proposed a method for determining the center for this expansion. This method requires excessive computation time and a human decision for choosing the proper center. Premont and Sheng [65] proposed a new method whose computation becomes two orders of magnitude faster and does not require human intervention for the determination of proper center of CH expansion. Another new criterion for determining the expansion center for CH filter was reported by Garcia-Martinez et al [66].

The work on phase-only and binary phase-only versions of CHC filters has also been investigated [28, 67-68]. Sheng and Arsenault [69] showed the detection of objects in scenes by analyzing correlation outputs at multiple locations. Several other methods based on CHC filters [70-79] have been reported for distortion-invariant pattern recognition.

2.2.2.2 Mellin transform based filters

The Mellin transform [80] based filters are used for achieving scale invariance. Mellin transform of a 1-D function $g(\xi)$ is given as

$$M(s) = \int_0^\infty g(\xi)\xi^{s-1}d\xi$$

where $s$ is a complex variable. Substituting $\xi = e^{-x}$ and taking $s = i2\pi f$, Eq. (2.16) can be written as

$$M(i2\pi f) = \int_{-\infty}^{\infty} g(e^{-x})e^{-i2\pi fx}dx$$

(2.17)

which represents the FT of a function $g(e^{-x})$. Thus Eq. (2.17) shows that it is possible to perform the Mellin transform with an optical FT system provided the input is introduced in a stretched
coordinate system, in which the natural space variable is logarithmically stretched i.e. \( x = -\ln \xi \).

If \( M_a \) represent the Mellin transform of a function \( g(a\xi) \) where \( a \) is the scale factor, then

\[
M_a(i2\pi f) = \int_0^\infty g(a\xi)e^{i2\pi f\xi} d\xi = a^{-i2\pi f} \int_0^\infty g(\xi')e^{i2\pi f\xi'} d\xi'
\]

(2.18)

where \( \xi' = a\xi \). If \( |a|^{-i2\pi f} = 1 \), Eq. (2.18) shows that the magnitude of \( M_a \) is independent of scale factor \( a \). Thus, the basic advantage of Mellin transform is its scale-invariance property. But, the Mellin transform is not shift invariant, which severely limits its applications. Casasent and Psaltis [81-82] and later Sheng and Arsenault [83] demonstrated the implementation of Mellin transform for pattern recognition. They replaced conventional FT by Mellin transform and used the resultant combined Fourier-Mellin transform which was found to be invariant to scale and shift in the input function.

### 2.2.2.3 Radial harmonic filters

Mendlovic et al. [84] introduced Mellin radial harmonic expansion based filter for achieving shift and scale invariance. The reference image is decomposed into a series of radial harmonic components. The shift and scale-invariant correlation is achieved if a single harmonic component is used. The radial harmonic expansion applied to the reference image, \( f(r, \theta) \), in polar coordinates can be written as

\[
f(r, \theta; x_0, y_0) = \sum_{m=-\infty}^{\infty} f_m(\theta; x_0, y_0) r^{i2\pi m} L^{m-1}
\]

(2.19)

with

\[
f_m(\theta; x_0, y_0) = \frac{1}{L} \int_{r_0}^{R} f(r, \theta; x_0, y_0) r^{-i(2\pi L m-1)} rdr
\]

(2.20)

where \( m \) is the harmonic order, \((x_0, y_0)\) is the expansion center, \( R \) is the finite size of the target to be detected (at its maximal scale), and \( r_0 \) is the smallest radius used in defining the expansion. The smallest radius, \( r_0 \), is selected so as to maintain the orthogonality of the decomposition. Thus it should satisfy the condition

\[
\ln R - \ln r_0 = L
\]

(2.21)
By selecting one radial harmonic component from the expansion (Eq. 2.19) with \(x_0 = 0\) and \(y_0 = 0\), the radial harmonic filter (RHF) is given as

\[
h(r, \theta) = f_n(\theta) r^{\frac{2\pi L}{L n - 1}}
\]

(2.22)

If \(f(r/\beta, \theta)\) represents the scaled target, with \(\beta\) being the scale factor, then the correlation output is given as

\[
O_{fh}^{(\beta)}(0,0) = \int \int f(r/\beta, \theta) h^*(r, \theta) r dr d\theta
\]

(2.23)

where \(O_{fh}^{(\beta)}(0,0)\) denotes the correlation at the origin i.e. \(x_0 = 0\) and \(y_0 = 0\).

The scaled target, \(f(r/\beta, \theta)\), can also be decomposed into radial harmonics as

\[
f(r/\beta, \theta) = \sum_{m=-\infty}^{\infty} f_m(\theta) r^{\frac{2\pi L}{L n - 1}} \beta^{1 - i \frac{2\pi m}{L}}
\]

(2.24)

Considering the orthogonality of the radial harmonic expansion and putting Eqs. (2.22 & 2.24) in Eq. (2.23), the correlation output is given as

\[
O_{fh}^{(\beta)}(0,0) = \int \int \left[ \sum_{m=-\infty}^{\infty} f_m(\theta) r^{\frac{2\pi L}{L n - 1}} \beta^{1 - i \frac{2\pi m}{L}} \right] \times \left[ f_n^*(\theta) r^{-\frac{2\pi L}{L n - 1}} \right] d\theta dr
\]

\[
= \beta^{-i \frac{2\pi m}{L}} \int_0^{2\pi} \left[ L f_n(\theta) f_n^*(\theta) \right] d\theta
\]

\[
= \beta e^{-i \frac{2\pi m}{L} \ln \beta} O_{fh}^{(1)}(0,0)
\]

(2.25)

Thus a scale change results only in an additional phase factor in the correlation output with the relative intensity distribution of the correlation output at (0,0) remaining unaffected.

As RHF is formed by selecting a single harmonic component, it contains only a small part of the information from the reference pattern and thus has poor discrimination ability. Also, RHF filter is not fully scale invariant. Several approaches have been proposed to improve the performance of RHF [85-91]. The performance of RHF also depends greatly on proper selection of expansion center and harmonic order. Moya et al. [87] proposed a method for determining the proper expansion center and order for Mellin radial harmonic filters.
2.2.2.4 Synthetic discriminant function filters

Hester and Casasent [92] introduced the basic synthetic discriminant function (SDF) filter (known as equal correlation peak SDF filter or projection SDF filter). The filter uses a linear combination of training images representing possible distortions like scale, rotation etc. in the reference image. In this approach, the filter is designed to yield a specific value at the origin of the correlation plane (also referred to as the ‘correlation peak’) in response to each training image. For example, in a two-class problem, the correlation values at the origin may be set to 1 for training images from one class, say true class, and to 0 for the training images from the other class, say false class. If \( \{ x_n(k,l) \} \) represents a set of \( N \) training images (true class and false class), then the SDF is the linear combination of this set of images and is given as

\[
h(k,l) = \sum_{n=1}^{N} a_n x_n(k,l)
\]  

(2.26)

where \( a_n \) (\( n = 1,2,3 \ldots \ldots N \)) are the coefficients. The correlation between the SDF filter \( h(k,l) \) and any arbitrary training image \( x_i(k,l) \) is given as

\[
h(k,l) \otimes x_i(k,l) = u_i = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(k,l) x_i^*(k,l) dkdl = \sum_{n=1}^{N} a_n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_n(k,l) x_i^*(k,l) dkdl
\]  

(2.27)

where \( u_i \) is either unity or zero depending upon the class to which \( x_i \) belongs. If \( p_{in} \) represents the correlation between \( x_i \) and \( x_n \), then \( u_i \) can be written as

\[
u_i = \sum_{n=1}^{N} a_n p_{in}
\]  

(2.28)

\( N \) linear equations, each similar to Eq. (2.28) but for a different value of \( i \), can be established for \( N \) training images. Such equations can be expressed in a single matrix form given as

\[
P \bar{a} = \bar{u}
\]  

(2.29)

where \( \bar{a} \) and \( \bar{u} \) are column vectors of length \( N \), and \( P \) is an \( N \times N \) matrix of correlations between the training images. From Eq. (2.29), the unknown coefficients \( a_n \) (\( n = 1,2, \ldots N \)) of \( \bar{a} \) can be found using

\[
\bar{a} = P^{-1} \bar{u}
\]  

(2.30)

Vijaya Kumar [93] proposed an efficient approach for designing linear combination filters. Riggins and Butler [94] investigated the feasibility of the implementation of SDF filter
using some of the standard computer-generated hologram methods. Casasent [95] described five different types of SDFs for different pattern recognition problems (intraclass recognition, interclass discrimination, and both intraclass and interclass object identification), which could be formulated using the same general matrix-vector equation (Eq. 2.29). Vijaya Kumar and Pochapsky [96] carried out a theoretical analysis to investigate the effects of training-set sizes used to synthesize the SDF filters. The work on design of phase-only and binary phase-only SDF filters [21-22, 26, 97] was also investigated for their implementation on available SLMs. But it was realized [22] that these filters do not perform well.

The assumption that SDF is a linear combination of training images is useful only when the SDF filters are synthesized in optical laboratory by using multiple exposure techniques. But for the digital synthesis of these filters, there is no need for this restriction. The complete set of solutions obtained by removing this unnecessary restriction was characterized by Bahri and Vijaya Kumar [98]. They worked on the generalization of the basic SDF filter. This led to the introduction of generalized SDF filters.

In generalized SDF filter, the training images \(x_1(k,l), \ldots, x_N(k,l)\), each of size \(d_1 \times d_2\) pixels, are sampled to yield arrays with \(d = d_1 \times d_2\) pixels in it. \(x_1, \ldots, x_N\) represent \(d\)-dimensional column vectors (obtained by scanning a 2-D array from left to right and from top to bottom and placing the resulting sequence of elements into a column vector) and \(h\) the column vector representation of the SDF filter \(h(k,l)\). With these notations Eq. (2.27) can be rewritten as

\[
x_i^T h = u_i
\]

where \(T\) denotes the transpose.

If \(X = [x_1, x_2, \ldots, x_N]\) is a \(d \times N\) matrix with \(N\) training image vectors as its columns and \(u = [u_1, u_2, \ldots, u_N]^T\) is an \(N \times 1\) vector containing the desired peak values for the training images, then \(N\) linear equations can be written as a single matrix-vector equation given as

\[
X^T h = u
\]

Since \(d \gg N\), \(h\) is assumed to be a linear combination of the training images i.e.

\[
h = Xa
\]

where \(a\) is the vector of coefficients \(a_1, a_2, \ldots, a_n\) defined before. The vector \(a\) is determined by putting Eq. (2.33) into Eq. (2.32) which yields

\[
X^T Xa = u \Rightarrow a = (X^T X)^{-1} u
\]
Substituting the value of $a$ from Eq. (2.34) into Eq. (2.33), the generalized projection SDF filter is given as

$$h = X (X^T X)^{-1} u$$  \hspace{1cm} (2.35)

Reid et al. [99] demonstrated the experimental verification of modified SDF filters for rotation invariance. Wang et al. [100] proposed a synthetic discriminant amplitude-compensated filter for optical pattern recognition. Brasher and Kinser [101] described the fractional-power SDFs. Shang et al. [102] proposed the design of SDF for use in a hybrid digital-optical correlator. Jamal-Aldin et al. [103] reported a SDF filter employing nonlinear space domain preprocessing on bandpass-filtered images.

The projection SDF filters suffer from certain problems. They do not consider any input noise and thus are not optimized for noise tolerance. These filters control only one value in the correlation output (the value at the origin) and thus sidelobes (which usually occur) are often much larger than this value at the origin, making the location of the correlation peak difficult. To improve the performance of SDF filters, many researchers extended the work on projection SDF filter, which resulted into its variants that are discussed below.

(i) **Minimum variance synthetic discriminant function filter**

The performance of a filter can be severely affected by the presence of noise and clutter. Thus it is important to characterize the behavior of the filter in the presence of noise and clutter. Vijaya Kumar [104] introduced minimum variance SDF (MVSDF) filter, which maximizes the noise tolerance of the SDFs. If the training vector $x_i$ is corrupted by the additive noise vector $v$, then the filter’s output response is given by (using Eq. 2.31)

$$(x_i + v)^T h = x_i^T h + v^T h = u_i + \delta$$  \hspace{1cm} (2.36)

Thus, the desired output $u_i$ is corrupted by the noise component $\delta$. The MVSDF filter is designed in such a way that the variance of $\delta$ in the output, caused by the input noise, is minimized. The input noise is assumed to be zero-mean, and to obtain an optimum filter, a parameter known as output noise variance (ONV) is considered which is given by

$$\text{ONV} = E \{ \delta^2 \} = E \{ (v^T h)^2 \} = E \{ h^T v v^T h \}$$

$$= h^T E \{ vv^T \} h = h^T Ch$$  \hspace{1cm} (2.37)
where \( C = \mathbb{E}\{vv^T\} \) is the input noise covariance matrix of size \( d \times d \). Minimizing ONV while satisfying the peak constraints on the training images in Eq. (2.32) leads to MVSDF filter, which is given as

\[
h = C^{-1}X(X^TC^{-1}X)^{-1}u \quad \text{(2.38)}
\]

From Eq. (2.38), it can be seen that the projection SDF filter is a special case, which is obtained if the noise is white (i.e. \( C \) is the identity matrix). Thus, the projection SDF filter is the optimum filter for recognizing the training images in the presence of additive white noise only and not for the other noise. Vijaya Kumar et al. [105] considered the case of Markov noise and used its special properties for designing the MVSDF filter. Sudharsanan et al. [106,107] reported SDF with reduced noise variance and sharp correlation structure and also suggested a method for the selection of optimum output correlation values in SDF design.

The MVSDF filters provide maximum robustness to noise but one drawback in using the MVSDF is that it is difficult to estimate and computationally difficult to invert the covariance matrix \( C \). Also, the MVSDF filter, like the projection SDF filter, controls only a single value in the correlation plane. Therefore, its output may contain large sidelobes, which can cause errors and false alarms.

(ii) Minimum average correlation energy filter

The projection SDF filters and MVSDF filters control only one value in the correlation plane. This makes it difficult to locate the correlation peak in the output, which may contain large sidelobes. Mahalanobis et al. [108] proposed minimum average correlation energy (MACE) filter, capable of suppressing sidelobes and producing sharp and distinct correlation peak in the output. The filter is based on the fact that it minimizes the energy (which includes the sidelobes) in the correlation plane. It considers a parameter known as the average correlation energy (ACE). If \( x_i \) (the columns of the matrix ‘\( X \’ \) defined earlier) are vector representations of the FT of the training images and \( h \) is the vector representation of \( H(k,l) \), the filter in Fourier domain, then ACE for \( N \) training images in the frequency domain is given as

\[
\text{ACE} = h^*Dh \quad \text{(2.39)}
\]

where \(^+\) denotes the conjugate transpose and

\[
D = \frac{1}{N} \sum_{i=1}^{N} X_i^*X_i \quad \text{(2.40)}
\]
is a \( d \times d \) diagonal matrix, and \( X_i \) is a diagonal matrix with the elements of \( x_i \) along its main diagonal. Since ACE is represented in Fourier domain, the constraints on the correlation peak in Eq. (2.32) can also be expressed in frequency domain as

\[
X^*h = u
\]  
(2.41)

The MACE filter minimizes ACE in Eq. (2.39) subject to the hard constraints in Eq. (2.41) and is given as

\[
h = D^{-1}X(X^*D^{-1}X)^{-1}u
\]  
(2.42)

Thus, MACE filters control the entire correlation plane and generally produce very sharp correlation peaks. They are the first SDF-type filters to be formulated in the frequency domain and have also been shown to be effective for finding training images in background and clutter. But these filters have two main drawbacks. Firstly, there is no noise tolerance built into these filters. Secondly, these filters are excessively sensitive to intraclass variations i.e. they may not provide high correlation peak values in response to non-training images from the desired class. Casasent et al. [109] proposed Gaussian MACE filters to reduce the sensitivity of the MACE filters to intraclass variations. Mahalanobis and Casasent [110] evaluated the performance of MACE filters. The use of CHCs in the design of MACE filter has also been reported to achieve sharp peaks and rotation invariance [111-115]. Some other variations of MACE filter design include the input-noise considerations [116], minimum squared error SDF filter for controlling the shape of resulting correlation outputs [117] and non-linear MACE filters [118].

### 2.2.2.5 Maximum average correlation height filter

The SDF filters are designed by imposing linear constraints on the training images to yield prespecified correlation values (1 for true class and 0 for false class) at the origin of the correlation plane. Such constraints do not explicitly control the filter’s ability to generalize over the entire domain of the training images. Also, these constraints do not hold for the non-training images, which always yield different values from those specified and achieved for the training images. Thus, an approach that provides distortion tolerance without imposing hard constraints on the design of the filter was proposed. These filters can be designed to offer good performance in the presence of noise and background clutter while maintaining relatively sharp correlation peaks for easy detection of the output.
Mahalanobis et al. [119] introduced an unconstrained correlation filter, known as maximum average correlation height (MACH) filter, based on a statistical approach. It uses a training set of \( N \) images, each of true class and false class, with each image of size \( d_1 \times d_2 \) containing \( d = d_1 d_2 \) pixels. Let \( x_i \) be a \( d \times 1 \) vector obtained by lexicographically reordering the 2-D FT \( (X_i) \) of the \( i^{th} \) training image \( (x_i) \) of true class, and \( h \) be the \( d \times 1 \) filter vector in the Fourier domain. The FT of the correlation plane produced in response to the \( i^{th} \) training image can be given as

\[
g_i = X_i h
\]

(2.43)

where \( X_i \) represents the \( d \times d \) diagonal matrix with the elements of \( x_i \) along the main diagonal. The deviation in the shape of correlation plane with an ideal or reference shape vector \( f \) is quantified by a metric, average squared error (ASE), and is defined as

\[
\text{ASE} = \frac{1}{N} \sum_{i=1}^{N} (g_i - f)^* (g_i - f)
\]

(2.44)

where \( ^+ \) denotes the conjugate transpose. An optimum shape, \( f_{opt} \), obtained by setting the gradient of ASE with respect to \( f \) as zero is given as

\[
f_{opt} = \frac{1}{N} \sum_{i=1}^{N} g_i = \overline{g}
\]

(2.45)

where

\[
\overline{g} = \frac{1}{N} \sum_{i=1}^{N} g_i = \frac{1}{N} \sum_{i=1}^{N} X_i h = \overline{X}h
\]

(2.46)

is the average correlation plane and

\[
\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i
\]

(2.47)

is the average training image expressed as a diagonal matrix. Using \( f = \overline{g} \) in Eq. (2.44), we obtain a metric known as average similarity measure (ASM) defined as

\[
\text{ASM} = h^* \left[ \frac{1}{N} \sum_{i=1}^{N} (X_i - \overline{X})^* (X_i - \overline{X}) \right] h
\]

\[
= h^* S_h h
\]

(2.48)
with
\[ S_x = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X})^* (X_i - \bar{X}) \] (2.49)
and * denotes the conjugate operation. \( S_x \) is a diagonal matrix in the frequency domain measuring the similarity of the training images to the average training image. Thus, ASM is the metric for distortion since it represents the average deviation of the correlation planes from the mean correlation shape. Minimizing the ASM improves the stability of the filter’s output in response to distorted input images.

In addition to being distortion-tolerant, a correlation filter must yield large peak values to facilitate detection of the target and to locate its position. For it, unlike SDFs, the filter’s average response to the training images is maximized without imposing any hard constraints on the peak values. Instead, the correlation peak value is desired to be as large as possible. The correlation peak value of the average correlation plane or the average correlation height (ACH) is given as
\[ ACH = \text{max} \left| g(0,0) \right| = h^* \bar{x} \] (2.50)
To make \( ACH \) large while reducing the ASM, the filter is designed to maximize the criteria
\[ J(h) = \frac{2 \text{ACH}^2}{\text{ASM}} = \frac{h^* xx^* h}{h^* S_x h} \] (2.51)
which is referred to as ACH criterion. Thus, the filter is called a maximum average correlation height filter that maximizes the relative height of the average correlation peak with respect to the expected distortions.

Another metric, called the average correlation energy (ACE), is minimized to reject the false class. If \( y_i \) be a \( d \times 1 \) vector that represents the FT of the \( i \)th training image \( (y_i) \) of false class, then ACE is given as
\[ ACE = h^* \left( \frac{1}{N} \sum_{i=1}^{N} Y_i Y_i^* \right) h = h^* D_y h \] (2.52)
where \( Y_i \) is a diagonal matrix containing the elements of \( y_i \), and
\[ D_y = \frac{1}{N} \sum_{i=1}^{N} Y_i Y_i^* \] (2.53)
is the diagonal matrix containing the average power spectrum of the false class. The modified performance criterion, which maximizes \( ACH \) and minimizes \( \text{ASM} \) and \( \text{ACE} \) is given as
From Eq. (2.54), the optimum maximum average correlation height (MACH) filter is given as

$$J'(h) = \frac{h^*xx^*h}{h^*(S_x + D_y)h}$$

(2.54)

where $c$ is a normalizing scale factor. The MACH filter offers improved distortion tolerance, yields sharp correlation peaks and is computationally simple. Thus, the MACH filter performs better than the previous SDF filters. Mahalanobis and Vijaya Kumar [120] optimized the MACH filter for detection of targets in noise. Alkanhal et al. [121] improved the false alarm capabilities of MACH filter. Singh and Vijaya Kumar [122] studied the performance of the variant of MACH filter known as extended MACH (EMACH). Nevel and Mahalanobis [123] did the comparative study of MACH filter variants using ladar. The use of MACH filter in tracking has also been reported [124-126]. Recently, Bone et al. [127] combined the MACH filter with the log-polar mapping technique to detect targets in background scenes with any kind of geometrical distortions. Goyal et al. [128,129] demonstrated the concept of a wavelet-modified MACH (WaveMACH) filter for rotation invariance. Aran et al. [130,131] reported log-polar transform-based WaveMACH filter for distortion-invariant target recognition.

### 2.2.2.6 Optimal trade-off filters

Refregier [132] proposed that the noise variance and correlation energy could be linearly combined into a single performance criterion to obtain a filter that yields low output noise variance as well as sharp correlation peaks. These filters were referred to as optimal trade-off (OT) SDF filters. Another important development in trading-off the noise sensitivity to peak sharpness is the minimum noise and average correlation energy (MINACE) filter developed by Casasent [133-136].

As the unconstrained MACH filter performs better than the conventional SDF filters with hard constraints, a technique for synthesizing unconstrained OT correlation filters was introduced by Vijaya Kumar et al. [137]. The correlation filter design techniques use several performance parameters such as ONV, ACE, ASM, and ACH. An ideal filter should provide a good balance between these performance criteria i.e. it should minimize ONV, ACE and ASM while maximize ACH. Such a filter could be obtained using an optimal trade-off approach. In this approach, the filter is optimized with respect to one criterion and the rest of the performance
criteria are held constant. Thus, to determine the OT filter $h$, the following energy function, which is a weighted sum of the four performance measures, is minimized

$$E(h) = \alpha \text{(ONV)} + \beta \text{(ACE)} + \gamma \text{(ASM)} - \delta \text{(ACH)}$$

$$= \alpha h^\top Ch + \beta h^\top Dh + \gamma h^\top S_h h - \delta |h^\top \bar{x}|$$

(2.56)

where $\alpha, \beta, \gamma,$ and $\delta$ are non-negative and $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1$. The ACH in $E(h)$ appears with a negative sign as it is to be maximized whereas the other three criteria should be minimized. $E(h)$ in Eq. (2.56) can also be written as

$$E(h) = h^\top Ph - \delta |h^\top \bar{x}|$$

(2.57)

where $P = \alpha C + \beta D + \gamma S_h$ ($\alpha$, $\beta$, and $\gamma$ being the optimal trade-off parameters) is a diagonal matrix with positive diagonal elements. Minimizing $E(h)$ leads to an unconstrained OT filter which is given as

$$h = (\delta/2) P^{-1} \bar{x}$$

(2.58)

with $\delta/2$ as a constant. This provides the best balance between the filter’s tolerance to noise, sharpness of the correlation peak, and distortions.

**2.2.2.7 Distance classifier correlation filters**

Mahalanobis et al. [138] introduced the distance classifier as a new approach of using correlation filters which leads to quadratic decision boundaries. The primary motivation for this was to improve distortion tolerance as well as to reduce the false alarm or error rates. The distortion tolerance is addressed in terms of similarity measures and discrimination is addressed by optimizing the filter to maximally separate the classes. The distance classifier correlation filters (DCCFs) are a generalization of MACE type filters and essentially compare full shapes under a transform which is optimum for simultaneously separating the classes and making them as compact as possible. Thus, for a given unknown input, its distance to all the class centers in the transformed domain is computed. The input is assigned to the class to which the distance from the target is the smallest.

The filter uses a set of $N$ training images, each of say two classes, $C_x$ and $C_y$, and that each image containing $d = d_1 \times d_2$ pixels. Let $X_i$ and $Y_i$ be the $d \times d$ diagonal matrices, which are equivalent representations of the $d \times 1$ image vectors $x_i$ and $y_i$ of classes $C_x$ and $C_y$ respectively. The mean image of the respective classes be denoted by the diagonal matrices $M_x$ and $M_y$ which
may otherwise be represented by the vectors \( \mathbf{m}_x \) and \( \mathbf{m}_y \). Let \( \mathbf{h} \) be the \( d \times 1 \) filter vector, which is the equivalent representation of the diagonal transform matrix \( \mathbf{H} \). The distance classifier uses this global transform \( \mathbf{H} \) to separate the classes maximally while making them as compact as possible. Let \( \mathbf{Z} \) be a diagonal matrix representing the unknown input image vector \( \mathbf{z} \).

The class means \( \mathbf{M}_x \) and \( \mathbf{M}_y \) using the training images is given as

\[
\mathbf{M}_x = \frac{1}{N} \sum_{i=1}^{N} \mathbf{X}_i \\
\mathbf{M}_y = \frac{1}{N} \sum_{i=1}^{N} \mathbf{Y}_i
\]

(2.59)

To separate the classes as much as possible, the filter is required to maximize the distance between the correlation peaks produced in response to the mean images of each class. Thus, the distance between the class means i.e. class separation is given as

\[
\text{Class separation} = |\mathbf{h}^*\mathbf{m}_x - \mathbf{h}^*\mathbf{m}_y|^2 \\
= \mathbf{h}^*(\mathbf{m}_x - \mathbf{m}_y)(\mathbf{m}_x - \mathbf{m}_y)^*\mathbf{h}
\]

(2.60)

As the distance to the true class is to be minimized, the average true class distance is estimated from the training images with respect to their respective class means as

Average true class distance

\[
= \frac{1}{N} \sum_{i=1}^{N} \left[ \mathbf{X}_i^*\mathbf{h} - \mathbf{M}_x^*\mathbf{h} \right] \left[ \mathbf{X}_i^*\mathbf{h} - \mathbf{M}_x^*\mathbf{h} \right] + \frac{1}{N} \sum_{i=1}^{N} \left[ \mathbf{Y}_i^*\mathbf{h} - \mathbf{M}_y^*\mathbf{h} \right] \left[ \mathbf{Y}_i^*\mathbf{h} - \mathbf{M}_y^*\mathbf{h} \right]
\]

\[
= \mathbf{h}^* \left[ \frac{1}{N} \sum_{i=1}^{N} (\mathbf{X}_i - \mathbf{M}_x)^*(\mathbf{X}_i - \mathbf{M}_x) + \frac{1}{N} \sum_{i=1}^{N} (\mathbf{Y}_i - \mathbf{M}_y)^*(\mathbf{Y}_i - \mathbf{M}_y) \right] \mathbf{h}
\]

(2.61)

where

\[
\mathbf{S} = \left[ \frac{1}{N} \sum_{i=1}^{N} (\mathbf{X}_i - \mathbf{M}_x)^*(\mathbf{X}_i - \mathbf{M}_x) + \frac{1}{N} \sum_{i=1}^{N} (\mathbf{Y}_i - \mathbf{M}_y)^*(\mathbf{Y}_i - \mathbf{M}_y) \right]
\]

(2.62)

is the diagonal similarity matrix and \( \mathbf{h}^*\mathbf{S}\mathbf{h} \) is known as the average similarity measure (ASM) as it is a measure of the similarity of the training images of a class to the mean. It can also be considered as a measure of the compactness of the class, hence minimizing ASM makes the classes compact.
To minimize ASM while maximizing the separation between the mean correlation peaks, the ratio $J(h)$ given in Eq. (2.63) is to be maximized

$$J(h) = \frac{h^*(m_x - m_x)(m_y - m_y)^* h}{h^* Sh}$$ (2.63)

The solution which maximizes the ratio $J(h)$ leads to the distance classifier correlation filter (DCCF) $h$ given as

$$h = S^{-1}(m_x - m_y)$$ (2.64)

Mahalanobis et al. [139] used DCCF for multi-class target recognition. Mahalanobis et al. [140] also studied the effect of noise on DCCF designed to recognize synthetic aperture radar (SAR) images and found DCCF to tolerate distortions and recognize targets in the presence of noise and clutter. Carlson et al. [141] proposed an optimal trade-off DCCF for SAR automatic target recognition (ATR). Kozaitis and Thangwaritorn [142] reported optimal trade-off and distance classifier circular filters for rotation-invariance. Alkanhal and Vijaya Kumar [143] worked on polynomial DCCF for pattern recognition. Bal et al. [144] used Fukunaga-Koontz transform and DCCF to improve target detection. The use of polynomial DCCFs along-with MACH filters has also been reported for pattern recognition and tracking [124-126].

### 2.2.2.8 Wavelet matched filters

An optical continuous wavelet transform (WT) [145-147] is inherently shift invariant and therefore is particularly useful for pattern recognition applications. It is a multiscale local operation that offers a way of matching on a local basis, which requires a small number of expansion coefficients and is superior to Fourier analysis for local feature extraction. Roberge and Sheng [148] described the wavelet matched filter (WMF), which is an application of the WT to OPR. The WMF performs the WT to enhance significant features of the images and the correlation between two WT coefficients in a single step. It improves the discrimination capability of the conventional matched filter against unknown inputs. WMF is a bandpass filter and is more robust to noise compared with the phase-only matched filter, which is a high-pass filter. It is defined in the Fourier plane as the product of the matched spatial filter $H^*(u,v)$, and the square modulus of the wavelet transform filter $|\psi(a_x,a_y,u,v)|^2$, $a_x$ and $a_y$ being the scale of the wavelet in $x$ and $y$ direction, i.e. $H^*(u,v) |\psi(a_x,a_y,u,v)|^2$. The correlation of the WMF for an input $f(x,y)$ with the FT $F(u,v)$ is given as
\[
\begin{align*}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u,v) \psi^*(a_x u, a_y v) H^*(u,v) \psi(a_x u, a_y v) \exp[i(xu + yv)] dudv \nonumber \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a_x, a_y, x', y') W_h(a_x, a_y, x' - x, y' - y) dx' dy' \quad (2.65)
\end{align*}
\]

where \( W_f \) denotes \( \text{WT}(f) \). Thus, Eq. (2.65) represents the correlation between \( W_f \ast W_h \).

As the WT and the matched filtering are linear operations, a composite WMF can be designed using a linear combination of the wavelets and a linear combination of the matched filters [149]. If \( \{x_n(k,l)\} \), where \( n = 1,2,3, \ldots \ldots N \), be a set of training images with the desired output values \( c_n \), then the composite WMF in the Fourier domain is given as

\[
G(u,v) = \phi(u,v) \left| \psi(a_x u, a_y v) \right|^2 \quad (2.66)
\]

where \( \phi(u,v) \) is a composite filter. If the training images \( x_n(k,l) \) are the scaled or rotated versions of the same image, the composite WMF may be invariant to scale or rotation changes. Thus, the wavelet filters are useful for distortion-invariant pattern recognition [150-157].

### 2.2.2.9 Other distortion invariant filters

Apart from the frequently used filters discussed above, a few other filters for distortion invariance have also been reported. These include lock and tumbler filters [158-161], quadratic filters [162-167], Kallman filters [35,168], Wiener filters [169-173], nonlinear filters [174-182], morphological filters [91,183-187], etc. Lock and tumbler filters use the concept of multiple filters for rotation-invariance that produce the same constant correlation output and one needs a systematic way of figuring out as to how many correlation filters are needed. They are computationally complex and are difficult to implement. Quadratic filters too use the idea of integrating multiple-filter output intensities on the same output detector array and are computationally intensive which has limited their use in many applications. Kallman filters need intensive computation as they optimize \( N^2 \) number of variables (\( N \times N \) being the size of each training image) and also do not account for possible uncertainties e.g. noise in the input. Wiener filters though effective in detecting a target in cluttered backgrounds produce a very low light efficiency [169]. The nonlinear filtering method [182] deals with the recognition of targets under limited out-of-plane rotations while invariant to different illumination conditions. Morphological filter, a nonlinear filter, is relatively less used than linear filters since an optimized method to design morphological filters was unavailable [184].
2.2.3 Iterative optimization algorithms for filter generation

A feasible procedure for the implementation of efficient pattern recognition is through the synthesis of optimized filters. The implementation of an optimization problem starts by the definition of a goal, which is to design a filter \( h(i,j) \) that performs the desired task under a given set of criteria. The optimization is implemented by minimizing some distance function \( d(h) \), by means of iterations, to reach the goal which is the functional of the filter function \( h \). There are terms like the cost function, figure of merit, energy function, and fitness function analogous to the distance function that are used in different optimization algorithms. The distance function represents some generalized distance from the present iteration of \( h(i,j) \) to the desired final solution of the problem or to a previous iteration, thus optimizing with respect to a prescribed rule. Some of the frequently used optimization procedures include hill-climbing procedure, simulated annealing, and genetic algorithm.

2.2.3.1 Hill-climbing procedure

The hill-climbing (HC) procedure is also known as a direct binary search. \( h^{(t)}(i,j) \) denotes the \( t \)-th iteration of the filter and \( d^{(t)} \) the distance function that is calculated for that iteration. After calculating the distance function of the \( t \)-th iteration, a change is induced over one element of \( h \) to obtain \((t + 1)\)-th iteration which changes \( d^{(t)} \) by an amount

\[
\Delta d^{(t+1)} = d[h^{(t+1)}(i,j)] - d[h^{(t)}(i,j)] \tag{2.67}
\]

This new function \( h^{(t+1)}(i,j) \) is accepted if \( \Delta d^{(t+1)} \leq 0 \), otherwise it is rejected. The procedure is repeated for the next element of \( h \) and so on until a desired minimum is achieved.

2.2.3.2 Simulated annealing

The simulated annealing (SA) algorithm is an iterative algorithm for solving complex optimization problems. In it, the distance function is a non-negative energy function that is to be minimized. The energy depends on the set of free variables, which are to be manipulated. In analogy to a metallurgic annealing process, in which the global minimum energy or the ground state of the physical system is reached by cooling the system from high temperature to low temperature, the SA algorithm is capable of reaching a global optimum by reducing a temperature parameter from an initially high value and perturbing the system variables with the change in temperature.
A random change in the elements of $h^{(t)}$ is induced at the $t$-th iteration to obtain the $(t+1)$-th iteration of the function $h$. This changes the energy function $d^{(t)}$ by an amount $\Delta d^{(t+1)}$ similar to Eq. (2.67). If $\Delta d^{(t+1)} < 0$, the new function $h$ is unconditionally accepted. However, for $\Delta d^{(t+1)} \geq 0$, the function may be conditionally accepted, based on the acceptance probability given as

$$P = \exp \left[ - \frac{\Delta d^{(t+1)}}{T} \right]$$

where $T$ is the temperature parameter. The process is repeated starting from the new function $h^{(t+1)}$ and gradually decreasing the temperature and is terminated when an adequately low energy is obtained. An unsuitable perturbation may cause a large deviation in the error function, which may trap the solution in a nonphysical state. Thus, the reduction in temperature parameter $T$ in Eq. (2.68) or the cooling rate must be chosen properly for best results. Munshi et al. [188] used the SA algorithm to synthesize rotation-invariant and distortion-tolerant filter.

### 2.2.3.3 Genetic algorithms

Genetic algorithms (GA) are search algorithms based on the mechanics of natural selection and natural genetics. There are three important features required to use GA: (i) a chromosomal representation of solutions to the problem, (ii) an evaluation function that gives the fitness of the population, which is in principle a distance function, and (iii) genetic operators to produce new structures from old ones: reproduction, crossover, and mutation. These algorithms efficiently exploit historical information to speculate on new search points with expected improved performance. As compared to other procedures of optimization, GAs show a much faster convergence to the optimal solution.

The new structure or generation is produced using the three operators: (i) reproduction, (ii) crossover, and (iii) mutation. Reproduction is the process in which individuals are selected as parents for contributing one or more offspring for the next generation. Crossover operator produces two offspring from the parents by exchanging some of the characteristics of the parents. This helps in exchanging good characteristics between individuals. Mutation is the occasional (with small probability) random alteration of a characteristic of an individual. The other two operators may sometimes lose a potentially useful characteristic so the mutation operator protects against such an irrecoverable loss. A well-known problem with the GA
algorithm is the premature convergence; therefore a proper mutation rate is important for satisfactory convergence.

Mahlab et al. [189] used GA for optical pattern recognition. Gupta et al. [190] reported the use of GA for the synthesis of optimum composite wavelet matched filter for out-of-plane rotation invariance.

2.3 Filters used in JTC

A conventional JTC gives better performance in the recognition process only if the input image has the same scale and orientation with respect to the reference image. Also, a classical JTC produces broad correlation peaks along-with strong zero order dc. The introduction of filters in JTC has been reported to improve its performance in terms of correlation peak intensity, correlation width, discrimination sensitivity as well as to make it distortion invariant. These include fringe-adjusted filter, SDF-based fringe-adjusted filter, and wavelet filter. There are also certain performance enhancing JTC architectures like binary JTC, differential JTC, and fractional JTC, which have been reported.

2.3.1 Fringe-adjusted filter

Alam and Karim [191] proposed a fringe-adjusted filter based JTC where the joint power spectrum (JPS) is multiplied by the fringe-adjusted filter (FAF) before inverse FT is applied to yield the correlation output. The FAF is given as

\[ H_{\text{fad}}(u, v) = \frac{B(u, v)}{A(u, v) + |R(u, v)|^2} \]  

(2.69)

where \( A(u, v) \) and \( B(u, v) \) are constants and \( |R(u, v)| \) is the amplitude of the FT of reference image. Thus FAF is a real valued function as it involves only the intensity and is suitable for optical implementation. Also, the computation of FAF involves only the reference image, so it can be computed prior to the correlation operation. Therefore, the inclusion of the filter does not have any detrimental effect on the processing speed of the system.

The FAF based JTC yields better results than classical JTC for input scenes involving single as well as multiple targets and also provides improved correlation discrimination [192]. In it, the amplitude matching is used more effectively to produce sharper and larger correlation
peak intensity. It does not cater to distortions like scale, rotation etc. but may be used to attenuate the noise that is present in the input scene by the proper selection of the constant $A(u,v)$. A fractional power FAF based JTC has also been suggested by Alam [193]. Wang et al. [194] proposed a modified FAF based JTC which performs better with noisy input images as compared to FAF-based JTC. Alam and Rahman [195] reported class-associative multiple target detection using fringe-adjusted JTC (FJTC). Alam and Horache [196,197] demonstrated an optoelectronic implementation of FJTC and used binary random phase mask to remove dc term and false alarms in FJTC. Haider et al. [198] demonstrated an enhanced class-associative generalized FJTC for multiple target detection. Wang et al. [199] reported a morphological FJTC.

### 2.3.2 SDF-based fringe-adjusted filter

Alam et al. [200,201] applied the SDF concept to a fringe-adjusted filter based JTC for recognizing distorted objects. The fringe-adjusted JTC reference image is synthesized from the training images. The spatial SDF $r(x,y)$ is constructed using $N$ training images, $r_0$, $r_1$, $r_2$……$r_N$, containing the desired distortion invariant features. This $r(x,y)$ is a linear combination of the training images and is given by

$$r(x, y) = \sum_{n=0}^{N} a_n r_n(x, y)$$  \hspace{1cm} (2.70)

where $a_n$ represents the coefficients which are evaluated using an iterative procedure [159]. Also,

$$R(u, v) = \sum_{n=1}^{N} a_n R_n(u,v)$$  \hspace{1cm} (2.71)

where $R(u,v)$ and $R_n(u,v)$ are the FTs of spatial SDF $r(x,y)$ and the $n^{th}$ training image $r_n(x,y)$ respectively. The SDF-based fringe-adjusted filter is given by

$$H_{saf} = \frac{B(u, v)}{A(u, v) + \left| \sum_{n=1}^{N} a_n R_n(u, v) \right|^2}$$  \hspace{1cm} (2.72)

where $A(u,v)$ and $B(u,v)$ are constants. The $H_{saf}$ is multiplied by JPS before applying inverse FT to obtain the correlation output.

The SDF based fringe-adjusted JTC yields sharp correlation peaks and provides distortion-invariance for the image from the training set. Chen et al. [202] investigated a distortion-invariant joint transform correlation, based on SDF and fractional power FJTC.
Horache et al. [203] demonstrated an optical implementation of SDF based fringe-adjusted joint transform correlation for invariant pattern recognition using both binary and gray level images. El-Saba et al. [204] presented a novel SDF formulated from the Laplacian-enhanced training images for the rotation and scale invariant target detection using FJTC.

2.3.3 Wavelet filter

WT has been drawing increasing attention in the optical pattern recognition community because of its attractive multiresolution, denoising, and feature extraction capabilities. Lu et al. [205] introduced the application of WT to JTC. It is a powerful tool for time (space)-frequency analysis of signals. By expanding a signal into a family of functions that are the dilations and translations of a basic wavelet, a set of wavelet representations of the signal at given frequency bands is obtained. Lu et al. used the input to the JTC by selecting appropriate bandpass filters, which are the representations of input functions at given frequency bands. The resultant correlation functions yielded significantly improved recognition of similar characters.

Ahmed et al. [206,207] investigated a wavelet-based JTC for rotation-invariant pattern recognition. The wavelet features are extracted, using an optimal set of filter parameters, at different resolution from a set of rotationally distorted training images to formulate a composite reference image. The technique resulted into highly robust and discriminating rotation-invariant detection.

Wang et al. [208] reported a joint wavelet-transform correlator for image feature extraction. Li et al. [209] demonstrated object recognition with wavelet-transform-based JTC. Kozaitis and Getbehead [210] reported JTC for wavelet feature extraction. Tripathi et al. [211,212] implemented Hartley transform on conventional and chirp-encoded JTC and also reported wavelet-feature based FJTC for noise immunity and pattern discrimination. Tripathi and Singh [213] introduced and demonstrated the use of wavelet-filter-based processing in JTC for achieving high-quality target discrimination. The JPS is processed by a series of wavelet filters and then Fourier transformed to yield the correlation output. Choice of a suitable wavelet filter yields sharp correlation output for the matched input patterns and a highly reduced correlation for the nearly similar non-target inputs. Alam and Chain [214] reported efficient multiple target recognition using a joint wavelet transform processor. Zhang et al. [215] demonstrated the
optical implementation of a photorefractive JTC using wavelet filters. An optical wavelet subband filtering for JTC has been recently reported [216].

2.3.4 Performance enhancing JTC architectures

2.3.4.1 Binary JTC

Javidi and Kuo [217] proposed a binary JTC (BJTC) in which the JPS is first binarized according to a threshold value before taking its inverse FT. Thus, if \( |F(u,v)|^2 \) represents the JPS of a function \( f(x,y) \), then the binarized JPS is given as

\[
|F(u,v)|^2 = \begin{cases} 
+1 & \text{if } |F(u,v)|^2 = T_f, \\
-1 & \text{otherwise} 
\end{cases}
\]  

(2.73)

where \( T_f \) is the JPS binarisation threshold which can be defined as

\[
T_f = \text{median}[|F(u,v)|^2] 
\]  

(2.74)

A BJTC has been shown to be superior to classical JTC in terms of both correlation peak intensity and discrimination sensitivity [217,218]. Yu et al. [219] discussed the effect of fringe binarization on the performance of multi-object JTC. Davis et al. [220] studied the effect of sampling and binarization on the output of a JTC. Hahn and Flannery [221] discussed the basic design considerations of a BJTC. Wang and Javidi [222] studied multi-object detection using a BJTC with different thresholding techniques. Grycewicz [223] described local windowing in the Fourier plane to propose a BJTC for multi-target detection. Osugi et al. [224] studied the quantization and truncation condition of Fourier power spectrum in a BJTC for good performance. Pati and Singh [225] carried out the experimental and simulation studies on the performance of binary and gray-valued joint transform correlators under poor illumination conditions and non-overlapping background noise.

2.3.4.2 Differential JTC

Zhong et al. [226] proposed a differential JTC (DJTC) in which the differential operation is applied to the JPS. The differential operation of the JPS is equivalent to the operation of putting a mask with a parabolic transmittance on the output plane. This mask has a zero transmittance of the \( dc \) term which is located at the center of the output plane, and a higher transmittance for the two correlation terms that are away from the center. The differential processing of the JPS also redistributes the light energy, which is an advantage to using a mask
in the output plane that absorbs energy. Thus, as differential operation is a high pass operation and can extract abrupt changes in an image, differential processing improves the contrast of JPS, which results into lower \( dc \) peak and higher sharp correlation peak, thereby significantly improving the discrimination ability of a JTC. Pati and Singh [227] studied the illumination sensitivity of JTCs using differential processing. Nishchal et al. [228] demonstrated the benefits of binary differential JTC employing a ferroelectric liquid crystal SLM.

2.3.4.3 Fractional JTC

Lohmann and Mendlovic [229] combined the concepts of JTC and the fractional Fourier transform (FRT) [230] to yield fractional JTC. FRT is a generalization of the conventional FT. The FT in JTC is replaced by FRT in fractional JTC i.e. the JPS and the correlation output is taken not exactly at the Fourier plane but near it, which results into improved correlation output. On the basis of the FRT operation, the classical correlation operation has been generalized to fractional correlation (FC). Kuo and Luo [231] generalized the architecture of JTC to achieve the joint fractional Fourier transform correlator for obtaining FC. Kuo et al. [232] proposed a quasi-joint fractional Fourier transform correlator. Tripathi et al. [233] evaluated the performance of joint FRT correlator in multiobject recognition. Zhu et al. [234] reported nonlinear joint fractional transform correlator. Song et al. [235] added phase masks at the input plane of joint FRT correlator to enhance the intensities of the fractional correlation peaks. Jin et al. [236] reported a joint extended FRT correlator.

2.4 Performance measures used in correlation

Signal-to-noise ratio (SNR) is an important performance metric for designing correlation filters as it quantifies the filter’s noise sensitivity. In addition to this, a sharp correlation peak is also desired in the correlation output so that it is easily discriminated from the background. Vijaya Kumar and Hassebrook [49] described several performance measures used in correlation. These include SNR, peak-to-sidelobe ratio, peak-to-correlation energy, Horner efficiency, and discriminability.
2.4.1 Signal-to-noise ratio

SNR is a measure of performance when the signal is corrupted with noise. Higher SNR values indicate better noise tolerance and lower average probability of error in detection. The SNR can be defined as [1]

$$\text{SNR} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} H(u)S(u) du \right)^2$$  \hspace{1cm} (2.75)

where $H(u)$ is the filter function, $S(u)$ is the FT of the signal $s(x)$ and $P_n(u)$ is the noise power spectral density.

2.4.2 Peak-to-sidelobe ratio

The peak-to-sidelobe ratio (PSR) is a measure of peak sharpness. It is defined as the ratio of the correlation peak to the standard deviation of correlation values in a region that is centered on the peak, but excluding a small region around the peak and is given as [57]

$$\text{PSR} = \frac{\text{peak} - \text{mean}}{\text{std}}$$  \hspace{1cm} (2.76)

where $\text{std}$ is the standard deviation of the sidelobe or annular region. Large values of PSR are obtained for sharply peaked correlations and small values for broad correlations.

2.4.3 Peak-to-correlation energy

The peak-to-correlation energy (PCE) is another measure of peak sharpness. It is defined as the ratio of correlation peak intensity value (square of the magnitude of correlation peak) to the total energy of correlation plane and is given as [49]

$$\text{PCE} = \frac{\left| y(0) \right|^2}{\int_{-\infty}^{\infty} \left| y(x) \right|^2 dx}$$  \hspace{1cm} (2.77)

where $\left| y(0) \right|$ denotes the peak magnitude. For sharp correlation peaks, the correlation plane energy is much smaller than $\left| y(0) \right|^2$ and hence PCE will be large whereas for broad correlation peaks, the PCE approaches to zero.
2.4.4 Horner efficiency

In optical correlation, it is important to make sure that the correlation plane gets much of the input light so that the detectors in the correlation plane can respond accurately and quickly. To quantify this, Horner introduced a light efficiency measure often termed as Horner efficiency, which is defined as [19]

\[
\eta = \frac{\text{total light intensity in output}}{\text{total light intensity in input}}.
\]  

(2.78)

2.4.5 Discriminability

An important attribute of correlation filters is their ability to discriminate one class of signal from another. Two non-noisy signals will differ in frequency domain at least at one frequency, and that single frequency can be used to discriminate the two signals without error. But the presence of random noise degrades this procedure. A good measure of the ability of a filter to discriminate between two signals in the presence of noise is the Fisher ratio (FR) defined as [49]

\[
\text{FR} = \frac{\left| E\{y_1(0)\} - E\{y_2(0)\} \right|^2}{\left[ \text{var}\{y_1(0)\} + \text{var}\{y_2(0)\} \right]/2}
\]

(2.79)

where \( y_1(0) \) and \( y_2(0) \) denote the output values at the origin when \( s_1(x) \) and \( s_2(x) \) are the inputs respectively, \( E\{,\} \) is the average and \( \text{var}\{,\} \) is the variance.