



# UNIVERSITAT DE BARCELONA

## Multichannel optical correlators: analysis, evaluation, and development of procedures for pattern recognition

Mario Montes Usategui

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Departament de Física Aplicada i Electrònica  
Universitat de Barcelona

Programa de Micro i Optoelectrònica Física  
1991-1993



# Multichannel optical correlators: analysis, evaluation, and development of procedures for pattern recognition

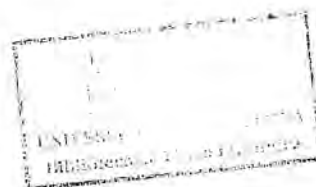
Mario Montes Usategui



BIBLIOTECA DE LA UNIVERSITAT DE BARCELONA



0700450333



Memoria presentada para optar al título de Doctor en Ciencias Físicas.

Realizada bajo la dirección de: Dr. Ignacio Juvells Prades  
Dr. Juan Campos Coloma

"For every  $\omega$ -consistent recursive class  $\aleph$  of *formulas* there exists a recursive *class formula*  $r$  such that neither  $v \text{ Gen } r$  nor  $\text{Neg } (v \text{ Gen } r)$  belongs to  $\text{Flg } (\aleph)$  (where  $v$  is a *free variable in the formula*  $r$ )."

K. Gödel, "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I," Monatshefte für Mathematik und Physik **38**(1), 173-198 (1931).

# Contents

<b>Context, objectives and outline .....</b>	<b>1</b>
<b>Chapter one. Pattern recognition by means of optical techniques: Optical correlators</b>	
Introduction .....	5
1.1 Fourier Transform and correlation .....	7
1.2 Optical correlation .....	10
1.3 Correlation between discrete functions .....	14
1.4 Comparison of digital and optical correlators .....	22
<b>Chapter two. Filter design for VanderLugt correlators</b>	
Introduction .....	31
2.1 Previous considerations: notation, procedures and quality criteria .....	32
2.2 Filter design for a reliable detection. Necessary conditions .....	39
2.3 Design of single-image filters .....	50
2.4 Lack of invariance of single-image filters .....	61
2.5 Design of invariant filters	
2.5.1 Circular harmonic filters .....	64
2.5.2 The general approach: synthetic discriminant functions .....	68
2.6 Equivalence of circular harmonic filters and invariant SDFs .....	85
<b>Chapter three. Limitations of single-filter correlations</b>	
Introduction .....	89
3.1 Correlation as a discriminant function .....	90
3.2 Decision boundaries in an optical correlator .....	93



<b>Chapter four. Sidelobe removal by a multichannel procedure: real case</b>	
Introduction .....	99
4.1 Method .....	100
4.2 The correcting filter .....	103
4.3 Necessary conditions .....	105
4.4 Computer experiments .....	115
4.5 Optical results .....	118
<b>Chapter five. Sidelobe removal by a multichannel procedure: complex case</b>	
Introduction .....	125
5.1 Method .....	126
5.2 Optimal settings for the parameters	
5.2.1 Constraints over the value of constant $c$ .....	131
5.2.2 Maximum eliminable sidelobes with a given number of filters .....	134
5.2.3 Optimal settings for constant $c$ and threshold $\theta$ .....	140
5.3 Algorithm .....	143
<b>Chapter six. Arbitrarily constrained filters</b>	
Introduction .....	147
6.1 The structure of the algorithm .....	149
6.2 The minimum Euclidean distance synthetic discriminant function .....	154
6.3 The algorithm .....	156
6.4 Results .....	161
6.5 Appendix: scaling constant leading to minimum error .....	172
<b>Chapter seven. The alternative: iterative correlators</b>	
Introduction .....	175
7.1 Cellular automata .....	175
7.2 Optical Life .....	176
<b>Conclusions</b> .....	185
<b>Bibliography</b> .....	191

## Context, objectives and outline

*"Fourier's theorem is not only one of the most beautiful results of modern analysis, but it may be said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics"*

Sir William Thomson, Lord Kelvin.

The probably somewhat exaggerated claim of Lord Kelvin -his enthusiasm is understandable if it is taken into account that the Fourier transform was originally conceived to solve the heat equation- is absolutely precise just by changing the term Physics by the more restricted of Optics. A good example of this is the so called Fourier Optics, to which this dissertation belongs.

Within this general frame, optical information processing and pattern recognition, as modern technological applications of the basic principles of Fourier processing, enjoy at present a patent interest. The reasons for this can be found in the singular nature of the methods involved: whereas the rest of the approaches are mainly digital and use electronic devices -computers-, optical pattern recognition is based on optical systems, mostly coherent ones, and utilizes analog procedures. The advantages and drawbacks derived from this singularity are therefore worth studying and part of this work is devoted to do so.

The historic roots of the methods for optical image processing go back to the experiments of E. Abbe in 1893 [Abb93] and A. B. Porter in 1906 [Por06] about how coherent images are formed. Such experiments, surprisingly close to present methods consisted of the spatial filtering of the Fourier transform of a pattern and its subsequent reconstruction. The results showed the power of such procedure for modifying and analyzing the images.

Soon this experiments found a practical application. In 1935, F. Zernike [Zer35] proposes a spatial filtering method that leads to the phase contrast microscopy, which is of great value for observing transparent structures. Later, in the fifties, A. Maréchal and coworkers propose a set of techniques to enhance the quality of photographic images. Among them, the reinforcement of the high spatial frequencies as a means to emphasize the details and a type of inverse filtering to eliminate the defocusing, lie already within the basic procedures used in this work.

By this same time, D. Gabor [Gab48][Gab49][Gab51] discovers holography, a wavefront reconstruction technique -originally intended to, although never used in, microscopy- whose subsequent development by E. N. Leith and J. Upatnieks [Lei62] made possible the VanderLugt filter and as a consequence the optical correlator.

A. VanderLugt -or Vander Lugt-, a researcher of the Radar Laboratory of the University of Michigan proposes an interferometric procedure that enables the recording of complex valued functions. This was the adaptation to the optical frequencies domain of the heterodine methods well known by radar engineers and commonly used in the synthetic aperture systems. Although VanderLugt's method was virtually identical to the off-axis holography -in contrast to the original procedure of Gabor- of Leith and Upatnieks, researchers as well in the Radar Laboratory, its development was independent of the later. This surprising duplicity, the restrictions to the spread of information owing to military reasons should be taken into account, and its coincidence in space and time, were without doubt the result of the exchange of ideas between optics and the theory of signal processing.

The pioneers of this marriage between the two subjects and of the adaptation to images of the signal detection problems were, among others, P. Elias [Eli52], E. L. O'Neill [One56] and L. J. Cutrona [Cut60]. The basic idea, finally possible thanks to the VanderLugt filter, was the detection by means of correlation. This operation, as detailed in the first chapter, is a measure of the similarity degree between two functions and can be optically performed by means of a spatial filtering of Fourier transforms.

It is worth pointing out that this filtering process, formally identical to the mentioned precedents of Abbe, Porter and Maréchal, implies a conceptual

difference. Until then these "optical processors" could be viewed of as imaging systems -the images more or less modified-. As such they handled all patterns the same way emphasizing, for instance, the high frequencies by means of a high-pass filter. Correlators, on the contrary, are systems matched to one image -hence the name of matched filter for VanderLugt's-, in such a way that they respond in a particular way when the input is the image to be detected. Such optical system carries specific information about a given object.

In this way and up until now, optical information processing has been benefited from the theory of signal processing in such a way that, as pointed by Goodman, ".. the merger of the two points of view [optics and signal processing] has become so complete that it is sometimes difficult to judge whether a particular piece of work should be published in an optics journal or an electrical engineering journal" [Goo68].

The object of this thesis, following that tradition, is the analysis of the basic limitations of optical correlation, for which several results well known in other branches of pattern recognition will be adapted. Also, we will develop procedures and alternatives as a solution to some of the difficulties encountered. The development of such objectives in this written work is as follows. The first chapter introduces the basic definitions that will be often used in later chapters. Furthermore, here we intend to show the usefulness of optical pattern recognition through a calculation of the equivalent computational power of an optical correlator.

Chapter two is devoted to survey the techniques of filter design for VanderLugt correlators. Here we cite and analyze more specific -and more recent- precedents than those of this general introduction. The purpose is to unify under a single argumental line works, whose development, owing to the initial immaturity of every subject, has been, ironically, incoherent. This therefore required a selection of the relevant results and an unification effort -as an example most of the mathematical proofs are original- that should be considered part of the results of this thesis. Within the general framework, this chapter illustrates the power and flexibility of modern filters, which suggests that the limitations of optical correlation do not arise from inappropriate designs.

Chapter three analyzes, based on the theory of discriminant functions and on the formalism of decision regions, the structural limitations of single-filter

correlations. Such limitations, quite severe when the inherent shift invariance capabilities of the correlation function are utilized, can be overcome by using multichannel correlators. Chapter four and five, in addition to illustrating this point, introduce practical methods to correct the sidelobes, one of the most common problems of the correlation.

Multichannel correlators face in turn several problems, now arising from their practical implementation, owing to the inadequacy of the current technology of spatial light modulators. Chapter six proposes an algorithm to design filters that constitutes a solution to this difficulty. The procedure offers several advantages with respect to similar algorithms previously proposed.

Finally, in chapter seven it is shown, now borrowing an argumentation from the field of cellular automata, that an optical correlator working in an iterative fashion, can as well solve the intrinsic limitations of single-filter correlators, thus being an alternative to multichannel setups.



## **Chapter one. Pattern recognition by means of optical techniques: Optical correlators**

### **Introduction.**

Artificial image understanding is being vigorously pursued because of its vital importance in such fields as robotics, artificial vision or artificial intelligence, all of them of an immense commercial interest. It has, by this reason, a long history that can be traced back to the first decades of the century when, with the inception of electronic computers, a practical solution seemed feasible. However pattern recognition by artificial means has revealed to be a much more complicated problem than previously thought and a lot of techniques have been proposed and discarded and subsequently rediscovered and discarded again in a sort of dialectic process which still continues.

The apparent ease with which we identify objects no matter the orientation in space, the scale, the illumination conditions, etc, is just that, apparent; the processes involved in the visual cortex are of great complexity. Two data clearly indicate this point. In humans, about 20% of the energy consumed by the body is due to the brain -its weight is only 2% of the total- and more than 60% of the brain is devoted, totally or partially, to vision tasks. The 100 watts of power that this represents are necessary to feed the  $10^{11}$  computing elements, the neurons, each connected to an average of 3000 companions, which constitute the human visual system.

The necessarily more modest hardware used in artificial pattern recognition forces us to restrict the type of problems we can satisfactorily solve. In this sense pattern recognition methods usually deal with a limited number of images in controlled environments. Whether or not a good method for these toy-

problems could be effective for the more general ones of real life is still an unknown.

Within this context, Optics has something to offer because it is the natural technology to deal with images. The way in which optical systems handle the two dimensional information that a given image carries resembles very much that of the neural systems [Rog91]. Its inherent parallelism, that is, the simultaneous processing of a large amount of data, the three dimensional nature of optical systems, the massive interconnection between processing elements with null cross-talk, and finally, its low power requirements, makes optical pattern recognition a strong competitor of the computer oriented approaches. Nevertheless, the great development of digital electronics with its intrinsic flexibility and accuracy is displacing optical technology from several traditional areas in image processing. For applications requiring small processing times, the parallelism of optics still offers advantages over electronics. This is specially the case of real-time pattern recognition where a large amount of data -an 8 bit gray scale image of 1024x1024 pixels contains 1Mbyte of information- must be processed in a very short time, frequently at video rates (25 Hz), which implies computation times of about 40 miliseconds per frame.

For computing a single correlation of a 1024x1024 image at video rate we need a device capable of achieving at least 200 Mflops, as we will show in the next sections, even by using the FFT, a very efficient algorithm for computing Fourier transforms. This is ten times the throughput of off-the-shelf signal processing hardware and can only be attained by means of supercomputers such as the CRAY, whose Y-MP 232 model runs at 600 Mflops on favorable problems [Alm94].

On the other hand, the operations that can be implemented through optical systems are very limited and they are almost exclusively reduced to linear transforms. Among these the easiest to accomplish and the more used in pattern recognition is the Fourier transform [Bra65], which allows to perform correlations between images, the procedure that is analyzed in this study. In this chapter these arguments are developed in detail and a comparison between electronic and optical processors is presented.

## 1.1 Fourier Transform and Correlation.

The Fourier Transform is a linear operation that provides information about the frequency distribution of a function. It is a basic tool in signal processing since the frequency content of a signal is often an important factor in its analysis. For example the elimination of a periodic noise becomes very easy by means of a filtering process in Fourier space [Gon93]. The Fourier Transform of the function  $f(x)$  is defined as:

$$F(u) = \mathfrak{F}[f(x)] = \int_{-\infty}^{+\infty} f(x) e^{-j2\pi ux} dx \quad (1.1)$$

where  $j$  is the imaginary unit and one dimensional notation has been used for convenience. Analogously, the inverse Fourier transform of the function  $F(u)$  can be written as:

$$f(x) = \mathfrak{F}^{-1}[F(u)] = \int_{-\infty}^{+\infty} F(u) e^{j2\pi xu} du \quad (1.2)$$

The basic properties of Fourier transforms that will be frequently used throughout this work are:

a) Inversion equations:

$$\mathfrak{F}^{-1}\{\mathfrak{F}[f(x)]\} = f(x) \quad (1.3)$$

$$\mathfrak{F}\{\mathfrak{F}[f(x)]\} = f(-x) \quad (1.4)$$

Equation 1.4 is specially interesting because it means that we can use the direct rather than the inverse Fourier Transform with the only difference of an inversion of  $f(x)$  with respect to the origin.



b) Parseval's Theorem: which, for our purposes, represents the conservation of the energy between the object and Fourier planes:

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |F(u)|^2 du \quad (1.5)$$

c) Correlation Theorem: The correlation between two functions,  $f(x)$  and  $g(x)$ , defined as:

$$(f \star g)(y) \equiv \int_{-\infty}^{+\infty} f(x)g^*(x-y)dx \quad (1.6)$$

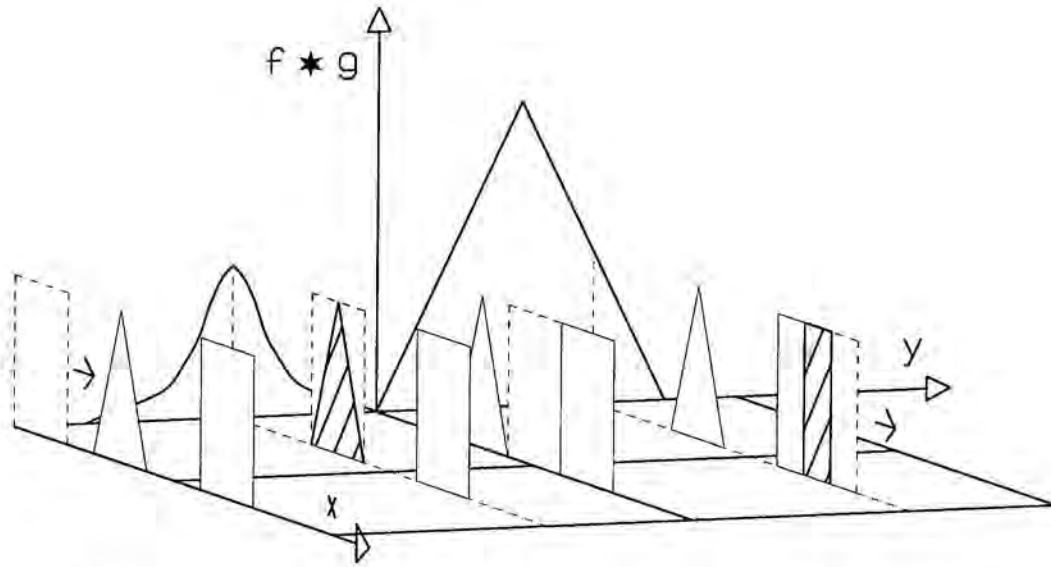
where the superscript \* means complex conjugation, can be expressed as:

$$f \star g = \mathfrak{F}^{-1}[F.G^*] \quad (1.7)$$

where  $F$  and  $G$  are the Fourier Transforms of  $f$  and  $g$  respectively. This property enables a simple implementation of the correlation by means of optical hardware as will be later shown.

The correlation can be utilized to measure the degree of similarity between the two functions being correlated. However, some care should be taken in doing this, since we can obtain arbitrarily large correlations by multiplying one of the functions by a proper constant. This means that the correlation is also a measure of the energy, so to speak, of the functions that are to be compared, and so we must normalize them in some way, to obtain a true similarity criterion. This issue has been considered, for the particular case of images by Kast and Dickey among others [Kast91].

To simplify the discussion, let us assume that we are dealing with real functions. In this case, the complex conjugation has no effect and thus the term  $g^*(x-y)$  reduces to  $g(x-y)$ , which is simply the function  $g(x)$  centered at point  $y$ .



**Figure 1.1-** Graphic representation of the correlation function. The plot shows  $g(x-y)$  for several values of  $y$  as well as the overlapping between  $f(x)$  and  $g(x-y)$ .

Then the product of  $g(x-y)$  and  $f(x)$ , and the sum of all these products are carried out. Intuitively it is clear that a large correlation is obtained if the high values of one of the functions are multiplied by the high values of the other and therefore, when the two functions are similar.

Figure 1.1 is an example of the correlation between an square,  $g(x)$ , of unit area and a function,  $f(x)$ , composed by a triangle and this same square. In this simple case, the correlation at a given point,  $y$ , is merely the area of the intersection of  $f(x)$  and  $g(x-y)$ . As  $g$  scans the X-axis, we obtain a measure of the degree of overlapping between the patterns. As can be observed, the maxima indicate the position of the two figures in  $f(x)$  and the absolute maximum is obtained when the position of both squares coincides.

Although we have used one-dimensional notation the correlation function can be easily generalized to an arbitrary number of dimensions. Two-dimensional correlation is of particular interest because it can be applied to the two-dimensional functions by which images can be represented. Correlation thus offers an effective method to identify and locate images, specially for applications implying the recognition of an object immersed in a large scene with high clutter and noise.

## 1.2 Optical Correlation.

The Fourier transform, which seems hard to evaluate since it implies a complex integral, can be readily obtained in practice, just by letting light pass through a positive lens. In particular if an object, whose transmittance could be represented by the function  $f(x)$ , is placed at the front focal plane of a convergent lens and is illuminated by an on-axis plane wave, the light distribution observed at the back focal plane is a scaled version of the Fourier Transform of  $f(x)$ .

This property of lenses is usually derived using diffraction theory -see for example [Goo68]- but we will give here a self-contained proof which needs only elementary optical principles. Since we assume that both object and lens are infinite, the diffraction integral only describes interference phenomena, whose result can be easily obtained by calculating optical path differences.

Let  $x$  represent the coordinate of a point of an object whose transmittance is  $f(x)$ , placed at the front focal plane of the lens  $L$  (see Fig 1.2). Notice that if it is illuminated by a on-axis plane wave, the phase in the plane immediately after the object only depends on the complex transmittance  $f(x)$ . Let  $x'$  be the coordinate of a point in the back focal plane of the lens. If  $A(x',x)$  represents the complex amplitude observed at  $x'$  due to light coming from  $x$ , we have:

$$A(x',x) = Cte f(x) e^{j\Delta\phi(x')} \quad (1.8)$$

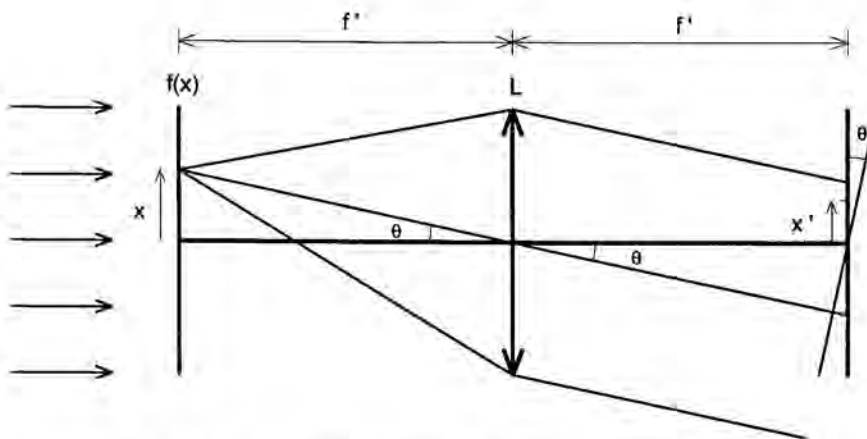


Figure 1.2- Optical setup for obtaining Fourier Transforms.

where  $\Delta\phi(x')$  is the phase associated to the propagation from  $x$  to  $x'$ . Since  $x$  is a point of the focal plane of  $L$ , rays coming from it will become parallel with an angle  $\theta$  with respect to the optical axis after passing through the lens. By applying the Malus-Dupin theorem [Bor75] -which states that the orthogonal surfaces to the light rays have constant phase- to the resulting plane wave centered at  $x'=0$  we have (Fig 1.3):

$$\sin\theta = \frac{\Delta C(x')}{x'} \Leftrightarrow \Delta C(x') = x' \sin\theta \quad (1.9)$$

where  $\Delta C(x')$  is the difference in optical path between the light rays reaching the center and those reaching  $x'$ . The paraxial approximation allows us to write:

$$\sin\theta = \tan\theta = \frac{x}{f'} \Rightarrow \Delta C(x') = \frac{x' x}{f'} \quad (1.10)$$

where  $f'$  is the focal length of  $L$ . Assuming that the origin of phases is at  $x'=0$ , such optical path represents a phase delay of:

$$\Delta\phi(x') = -k\Delta C(x') = -\frac{2\pi}{\lambda} \frac{x' x}{f'} \quad (1.11)$$

where  $k$  is the wavenumber and  $\lambda$  the wavelength of the light. By substituting eq. 1.11 into eq. 1.8 we have:

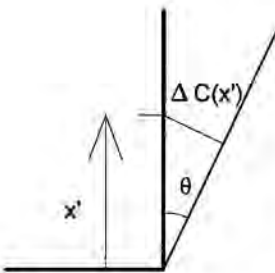


Figure 1.3- Optical path differences.

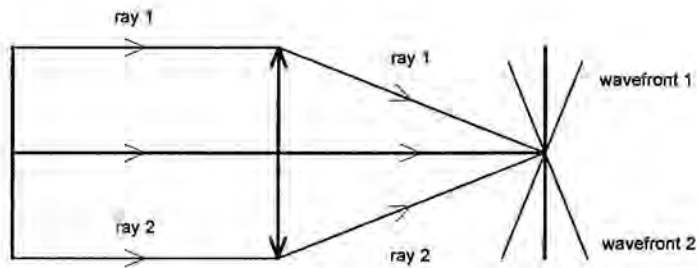


Figure 1.4- The phase at the origin of the back focal plane is the same for all wavefronts.

$$A(x', x) = \text{Cte } f(x) e^{-j \frac{2\pi x'x}{\lambda f'}} \quad (1.12)$$

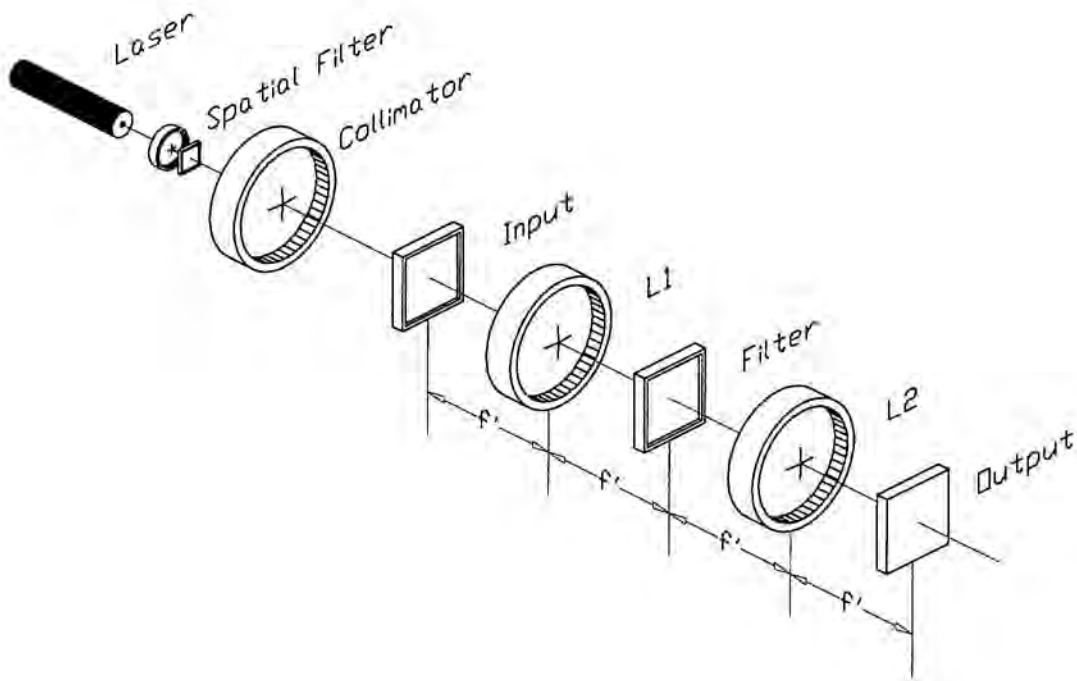
At point  $x'$  we will observe the coherent superposition of the wavefronts emitted by the whole object. The terms of that superposition can all be expressed by means of equation 1.12 with respect to the phase of each wavefront at  $x'=0$ . To carry out the sum we have first to write all these terms with respect to a common origin. However, this is actually not necessary as illustrated in Fig. 1.4. The figure shows the wavefronts originated by two distinct object points and two light rays that reach the origin and belong to those wavefronts. Since these rays can also be considered as part of a plane wave being focused by the lens, applying again the Malus-Dupin theorem we see that the phase at the origin of all plane waves is the same. Therefore, finally, the complex amplitude at point  $x'$  can be written as:

$$A(x', x) = \text{Cte} \int_{-\infty}^{+\infty} f(x) e^{-j \frac{2\pi x'x}{\lambda f'}} dx \quad (1.13)$$

which is the Fourier Transform of  $f(x)$ , where the frequency  $u$  is observed at point  $x'=u\lambda f'$ , *quod erat demonstrandum*.

This property of lenses together with the possibility of obtaining the correlation product by means of Fourier transforms allows us to perform correlations with a reduced optical setup. The basic architecture, known as 4-f VanderLugt correlator, is sketched in figure 1.5. It consists of an illumination system -composed of a low-power laser, a spatial filter to clean the beam and a collimating lens to obtain the on-axis plane wave- two displaying devices, two converging lenses, and a detector. The first display, labeled *Input*, is used to show the scene that presumably contains the target pattern and the second one, labeled *Filter*, encodes the information about the pattern being searched for.

When there is no filter at the focal plane of  $L_1$ , the VanderLugt correlator is just a telecentric imaging system that provides a lateral magnification of -1 as a simple analysis based on geometrical optics shows. However, an alternative



**Figure 1.5-** Schematic diagram of a 4- $f$  VanderLugt correlator.

explanation based on Fourier optics is more enlightening for our purposes. The first lens  $L_1$ , provides the Fourier transform of the input image at its back focal plane. The complex distribution observed at the output plane will be the Fourier transform -due to the lens  $L_2$ - of that observed at the filter plane, so in absence of a filter it is merely a double Fourier transform of the input image. Thus, as equation 1.4 indicates, we will observe the input pattern with a minus one magnification. If the Fourier transform of the input is modified by interposing a filter at the back focal plane of  $L_1$ , the second lens will reconstruct a modified version of the original image; for instance, if we put a mask that blocks the high frequency components, that is a low-pass filter, we will register at the output a blurred version of the input image.

When the transmittance of the filter encodes the properly scaled Fourier transform of a reference object, the output will be the correlation between this reference and the input image. This is so because, after the filter we obtain the frequency-by-frequency multiplication of both Fourier transforms, the incoming and that codified in the filter display. The lens  $L_2$  performs a final Fourier



transform of this product, which according to eqs. 1.4 and 1.7 is, aside of the inversion, the correlation between both patterns.

The design of filters for the VanderLugt correlator is an issue of the utmost importance and will be extensively covered in chapter 2. In practice, the basic setup we have analyzed should be modified in order to get variable-scale Fourier transforms or to shorten the overall length of the system [Clar91][Cra91][Geb91], although for the ends of this work such changes are unimportant. It is worth pointing out that the VanderLugt architecture is not the only possible but there exist different alternatives to carry out the correlation between images by optical means. In particular, the Joint Transform Correlator has been widely studied because it offers some advantages such as a better stability and an easy alignment. On the other hand, the flexibility, basically in the filter design, is lower than that of a VanderLugt correlator although the nonlinear effects recently proposed are rapidly overcoming this difficulty [Jav89]. Finally, the Acousto-Optic correlator based on the direct realization of the correlation integral instead of on the Fourier approach is also a good possibility to build high speed optical pattern recognition systems. The input devices, the acousto-optic cells, are used in communications and spectrum analysis applications, they have been commercially available for many years and thus represent a mature technology. Furthermore they do not require coherent illumination and have a large bandwidth, a fast response and high dynamic range [Psa82][Mol91].

### **1.3 Correlation between discrete functions: aliasing.**

Up to now we have been considering continuous functions to represent the information contained by images and although this provides an elegant formalism, is unadequate for most practical purposes. The limited resolution, that is the limited capability to treat as distinct two close or similar points, of actual devices for capturing, processing and storing information makes the discrete representation a more suitable one. The influence of the discretization of signals is specially important on Fourier processing and thus an analysis of these effects

is needed. Signals will be represented by the array of values,  $\{f(x_k)\}$ , taken by the function  $f(x)$  at points  $x_k=k\Delta x$ , where  $\Delta x$  is called sampling interval.

We will focus our attention on bandlimited signals. Natural images can be considered with wide generality as members of this special class of functions for which the intuitive rule-of-thumb that the smaller the sampling interval, the more accurate the representation of the signal is valid until a limit. A given sampling interval is enough to exactly reconstruct them by using interpolation techniques and thus a smaller  $\Delta x$  does not result in a better description of  $f(x)$ . Bandlimited functions are defined as those whose Fourier transform has a finite region-of-support, being the region-of-support the subset of the domain for which the function has a nonzero value. This result is usually summarized in the

Wittaker-Shannon sampling theorem:

Let  $f(x)$  be a bandlimited function, whose Fourier transform  $F(u)$  is different from zero only inside the interval  $(-B,B)$ . The necessary and sufficient condition for  $f(x)$  to be recoverable from the regularly sampled data  $\{f(k\Delta x)\}$  is:

$$\Delta x \leq \frac{1}{2B} \equiv \frac{1}{v_n} \quad (1.14)$$

where  $v_n=2B$  is known as Nyquist frequency.

A thorough demonstration of the theorem as well as some pointers to relevant bibliography can be found in [Goo68] and we give only a general explanation. The possibility of recovering a given signal  $f(x)$  is based on the fact that when a function is sampled in one domain, it becomes replicated in the other -a discrete function has a periodic Fourier transform and a periodic function has a discrete one- with a periodicity  $\Delta u$  given by:

$$\Delta u = \frac{1}{\Delta x} \quad (1.15)$$

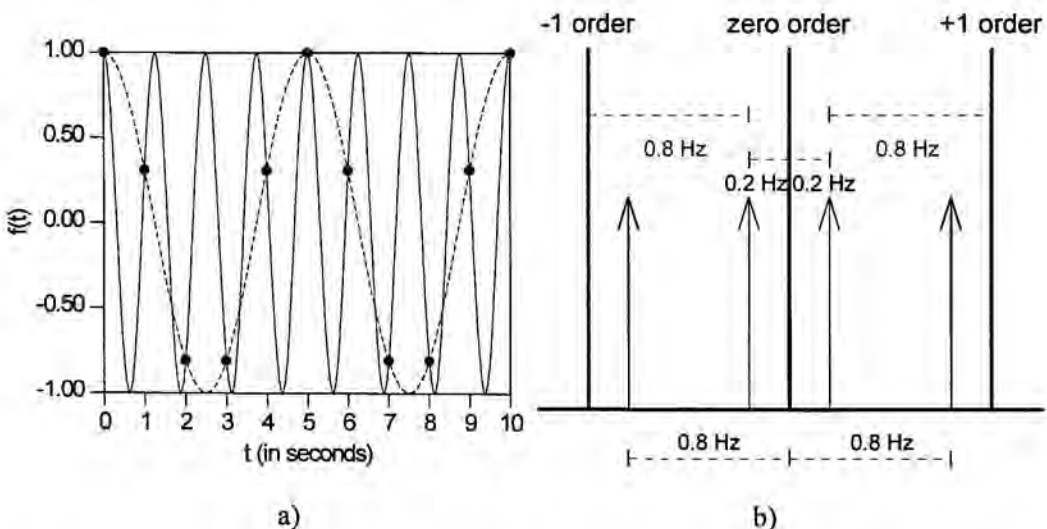


where  $\Delta x$  is the sampling interval. A physical example of this phenomenon is the spectral orders ( $m=0,+1,-1,..$ ) observed at points  $m\lambda f^*/\Delta x$  -notice the scale factor- in the Fraunhofer pattern of a diffraction grating, the discrete equivalent of a constant function.

If the Fourier transform of  $f(x)$  is zero outside the interval  $(-B,B)$ , the condition for the replicas to not overlap each other is thus the one expressed by eq. 1.14. Finally, if one of the terms can be isolated from the rest, the function can be reconstructed simply by performing an inverse Fourier transform.

The Witteraker-Shannon theorem thus establish a lower limit for the frequency at which a function can be digitized without any information being lost in the process: the sampling frequency must be twice the highest frequency contained in the signal. Failing in observing this principle results in errors known under the generic name of aliasing.

Examples of aliasing are abundant in real life: the flickering of computer screens in TV news programs or wheels appearing to spin backwards and rain to fall upwards in ancient movies, to mention only a few. All of them have in common the conversion of high frequencies into unreal low ones, as illustrated in Figure 1.6. Figure 1.6.a) shows a cosine function of frequency 0.8 Hz (solid line) being digitized at a sampling rate of 1 Hz, and the filled circles represent the



**Figure 1.6-** Aliasing due to insufficient sampling. a) time domain; the 0.8 Hertz cosine (continuous line) can not be distinguished from a 0.2 Hertz cosine (dashed) when digitized at 1 Hertz (filled circles). b) frequency domain; overlapping of different orders.

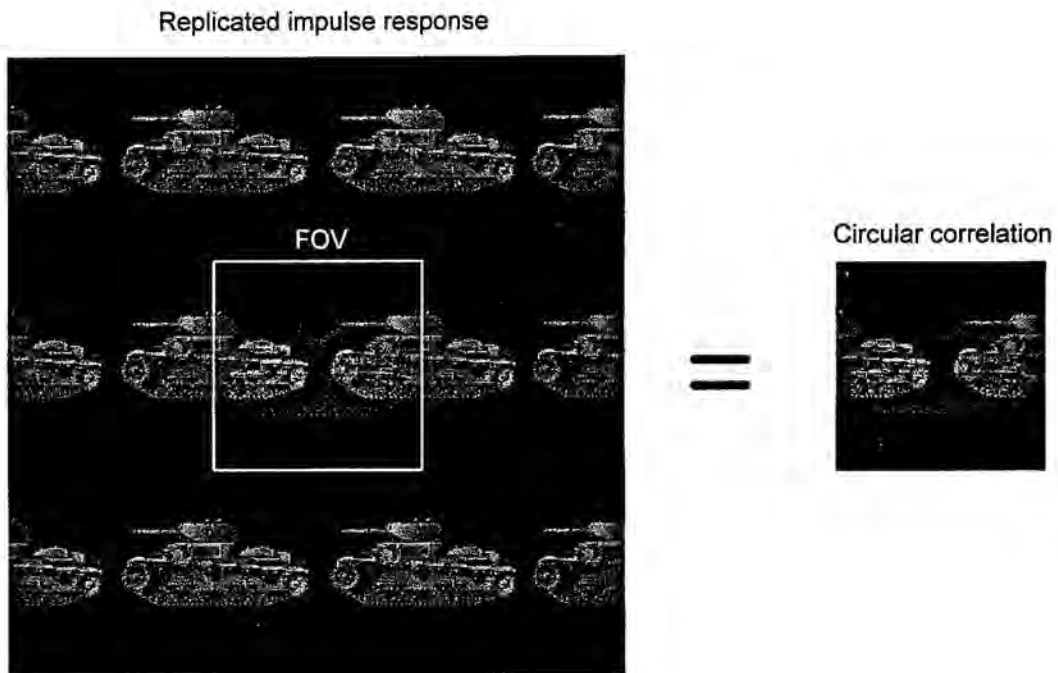
discrete values taken by the function (one per second). Since the Nyquist frequency for this case is  $2 \times 0.8 = 1.6$  Hz, the discrete values do not completely characterize the function, which can be confused with a  $(1 - 0.8)$  Hz = 0.2 Hz cosine. Figure 1.6.b) provides a complementary interpretation from the point of view of Fourier domain. The discrete cosine gives rise to infinite replicas of the Fourier transform of the continuous cosine -a pair of delta functions-. As the sampling was carried out at 1 Hz, their separation is also 1 Hz, which is insufficient to prevent the mixture of adjacent terms. The net result is the appearance of a false 0.2 Hertz cosine.

When the correlation is implemented via Fourier transforms the effect of aliasing has to be carefully considered. The main limitation comes from using a discrete filter as we show in the following. This is valid whenever we deal with discrete images, that is in a digital implementation or in an optical correlator when pixelated modulators are used, although the analysis we are carrying out is focused on this latter case. For the sake of clarity and because the discussion is somewhat lengthy we summarize it here, at the beginning. Our purpose is to show that, owing to sampling effects and no matter the size of the target pattern, the filter must always have the same number of pixels as the whole input scene. The problems raised and a possible solution are also briefly commented.

Let us suppose that we have an input scene of  $N \times N$  pixels that contains a target image of  $Q \times Q$  pixels. The spacing between pixels -the pitch- in the input modulator is  $\Delta s$ . The filter plane modulator consists of a regular lattice of  $T \times T$  pixels of pitch  $\Delta u$ . In these conditions, the filter impulse response -or more correctly, the correlator impulse response-, that is the inverse Fourier transform of the discrete filter, is composed of infinite copies of the intended encoded image appearing at  $x'_m = m\lambda f / \Delta u$ . When the dimensions of the scene are small compared with the period of the replicas, namely when:

$$\Delta x \equiv (N - 1)\Delta s \leq \frac{\lambda f'}{\Delta u} \quad (1.16)$$

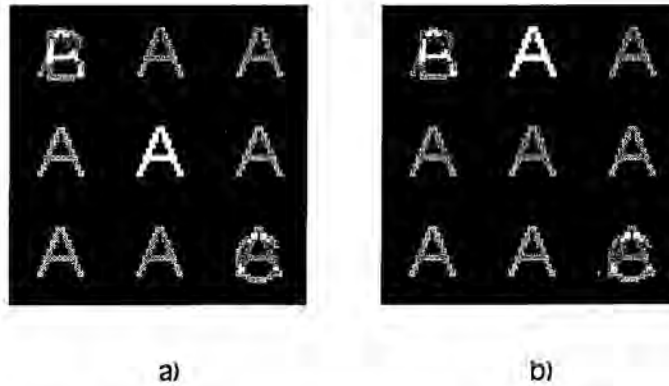
where  $\Delta x$  represents the lateral dimensions of the input modulator, the result is a circular correlation as depicted in Figure 1.7. The picture shows the impulse response of a  $128 \times 128$  filter, encoding the image of an Abrahams tank being



*Figure 1.7- Circular correlation produced by a discrete filter.*

correlated with a Panzer tank, the latter displayed in a darker gray scale. The correlation is being evaluated at point (54,0) as indicated by the displacement of the array of Abrahams. The scene, here only the Panzer on a black background, is limited by the square labeled FOV (field of view). Obviously, the position of the target makes sense only for points belonging to the field of view and thus we are only interested in the correlation on this region. For the region of interest everything happens as if the edges of the scene were glued together, in such a way that when part of the reference image goes out of the field of view, reenters from the opposite side.

Circular correlations retain most of the properties of the aperiodic ones and in particular the capacity for locating and identifying objects. Furthermore, the most powerful filter designs can only be obtained by means of digital computations and implemented through digital holograms, which implicitly assume the circularity of the correlations and thus, in practice, this effect is not disadvantageous. The problem appears when the input scene is large compared with the periodicity of the impulse response, as illustrated in Figure 1.8. In this example, the field of view, formed by the letters A, B and C, spans nine copies of



*Figure 1.8- Same correlation at the origin. a) centered target. b) centered at the order (0,-1).*

the target pattern, the letter A, which yields to a high correlation at the origin of the correlation plane no matter whether the A is actually located at the center or at one of the spectral orders. In this situation there are errors in the determination of the position of the target as well as false peaks. Therefore, to avoid them, the size of the input scene must be at most:

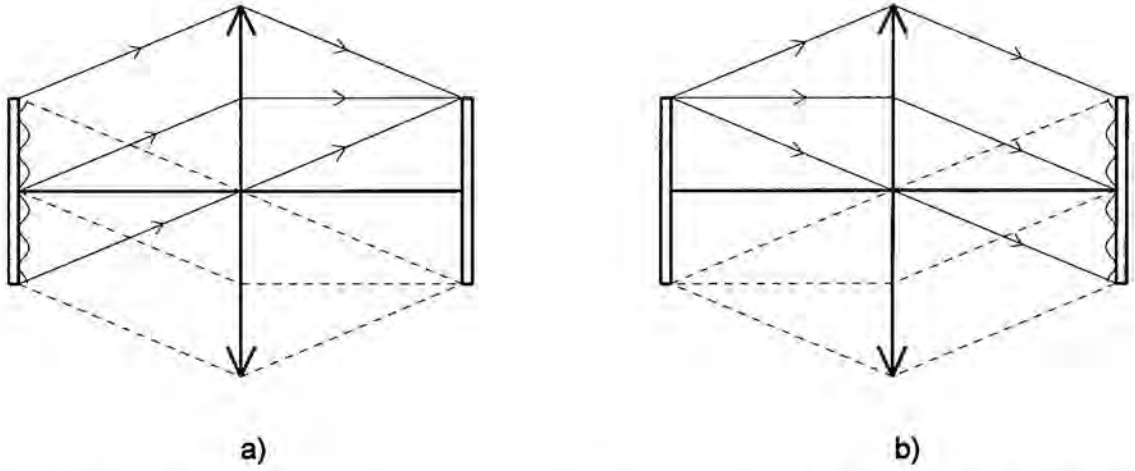
$$\Delta x = \frac{\lambda f'}{\Delta u} \quad (1.17)$$

On the other hand, the factor  $\lambda f'$  can not be arbitrarily set because of the need of matching the scales of both filter and scene, which forces us to use a filter with the same number of pixels as the input modulator. This is so because the maximum spatial frequency contained in the displayed image is:

$$u_{\max} = \frac{N-1}{2\Delta x} \quad (1.18)$$

The plane wave that carries the information of this frequency must be focused at most at the last pixel of the filter modulator, namely at a distance:

$$x' = \frac{T-1}{2} \Delta u \quad (1.19)$$



**Figure 1.9-** Size-frequency relationships between input and filter planes. a) The maximum frequency present in the input image must be focused on the outmost pixel of the filter. b) the outmost pixel of the image gives rise to the maximum frequency contained in the Fourier transform.

from the optical axis (see Figure 1.9.a) so due to the scale of the optical Fourier transform given by Equation 1.13 we have:

$$\frac{N-1}{2\Delta x} \lambda f' = \frac{T-1}{2} \Delta u \quad (1.20)$$

whence

$$\lambda f' = \frac{T-1}{N-1} \Delta u \Delta x \quad (1.21)$$

and thus we have:

$$\frac{\lambda f'}{\Delta u} = \frac{\frac{T-1}{N-1} \Delta u \Delta x}{\Delta u} = \frac{T-1}{N-1} \Delta x \quad (1.22)$$

and substituting into equation 1.17 we finally get  $T=N$ .

It is worth pointing out that we have supposed that the target image is only  $Q \times Q$  pixels. Let us derive the number of pixels needed to accurately represent its Fourier transform.

The maximum frequency contained in the Fourier transform of a centered target is due to the outmost pixels (Figure 1.9.b), namely those at the distance from the optical axis given by:

$$x = \frac{Q-1}{2} \Delta s = \frac{Q-1}{2} \frac{\Delta x}{N-1} \quad (1.23)$$

whence the maximum spatial frequency is:

$$\omega_m = \frac{\frac{Q-1}{2} \frac{\Delta x}{N-1}}{\lambda f'} = \frac{\frac{Q-1}{2} \frac{\Delta x}{N-1}}{\frac{T-1}{N-1} \Delta u \Delta x} = \frac{Q-1}{2(T-1)\Delta u} \quad (1.24)$$

The maximum frequency that can be represented in the filter modulator is:

$$\omega'_m = \frac{1}{2\Delta u} \quad (1.25)$$

and, according to the Witterker-Shannon theorem, for the Fourier transform of the target to be well represented we need:

$$\omega'_m \geq \omega_m \Leftrightarrow \frac{Q-1}{2(T-1)\Delta u} \geq \frac{1}{2\Delta u} \Leftrightarrow T \geq Q \quad (1.26)$$

which means that a Fourier Transform is properly encoded by using the same number of pixels as the object. Therefore, although a filter of  $Q \times Q$  pixels is enough to represent a target of  $Q \times Q$  pixels, the necessity of avoiding the errors



raised by the discrete nature of the Fourier plane modulator, namely the aliasing errors, makes necessary the use of a filter with the same bandwidth as the whole scene. Notice that this makes the Fourier implementation of the correlation inefficient for small targets embedded in a large scene when compared with the direct evaluation of the correlation in object space. This is considered by some authors, [Hua91][Ghe92], as one of the major handicaps for building competitive high speed optical correlators -a large bandwidth modulator needs a large memory and complicated control electronics and is necessarily slow- to address broad-area-search problems. A possible solution is the use of randomly pixelated modulators [Hed92], for which the sampling theorem is less restrictive and which require a bandwidth equal to that of the target image. The price is a lower fill factor and thus a smaller optical efficiency.

#### **1.4 Comparison of digital and optical correlators.**

Optics provides a simple way, although not exempt of some difficulties, to perform the correlation product, as shown in the preceding sections. However, the key question is whether or not this represents advantages enough over digital electronics that can compensate the difficulties of developing a new technology.

Huang *et al.* and Gheen *et al.* carry out a comparison between optical and digital correlators in [Hua91] and [Ghe92]. They computed a figure of merit defined as correlations per second per watt, concluding that optical processors still enjoys a significant advantage for large fields of view; for example optical correlators can perform 41.67 to 416.67 corr/s/watt depending on the filter plane modulator, for input scenes of 1024x1024 pixels, whereas digital processors can, in the best case, perform only 4.156 corr/s/watt, that is ten to one hundred times less. However, since their data depend strongly on delicate assumptions and extrapolations of technology trends we prefer to make such a comparison by adopting a different point of view. We will derive the number of floating point operations per second -flops- needed to compute the same number of correlations than those performed by an actual optical correlator: SPOTR. Since the number

of floating point operations per second is an standard measure of computational power we can get a fair idea of its electronic equivalents. Although this approach does not encompass important issues such as the cost of those operations, for example through the power consumption, as the mentioned references, it relies only on non-controversial assumptions.

SPOTR, the acronym for *System for Passive Optical Target Recognition*, is a high-speed multichannel VanderLugt correlator developed by Martin-Marietta and Boulder Nonlinear Systems Inc. for the Transition of Optical Processors into Systems-TOPS project, a program funded by the Advanced Research Projects Agency to produce rugged optical processors [Lin94][Cas94] and represents the state-of-the-art in this technology. It is designed to achieve distortion invariant recognition over a large range of view angles and distances by correlating the input scene with a filter database.

The input scene and the filter are displayed on two ferroelectric spatial light modulators (SLM) of 128x128 pixels and capable of a frame rate in excess of 2000 Hz. The source of coherent illumination is a 690 nm laser diode and the correlations are captured by means of a high speed camera developed by DALSA working at 800 Hz. The prototype consumes less than 126 watts, occupies less than a cubic foot ( $0.028 \text{ m}^3$ ) and weights about 43 pounds (19.50 Kg).

With this architecture, the speed of the optical correlator is limited by the electronic devices used at the input and output planes, namely modulators and cameras. The frame rate allowed by the optical elements, mainly limited by the speed of light, is higher as we will show.

The ferroelectric SLM has a pixel pitch of  $\Delta x=30\mu\text{m}$ . Since it is composed of 128x128 pixels, its size is

$$d=0.03 \times 128=3.84 \text{ mm.} \quad (1.27)$$

The higher spatial frequency that can be displayed on the modulator is  $1/(2\Delta x)$ , which, to be adequately modified at the Fourier plane, must intercept the filtering modulator at  $x'=d/2$ . Owing to the scaling of the optical Fourier transform given by eq. 1.13 we have:



$$\frac{d}{2} = \lambda f' \frac{1}{2\Delta x}, \quad (1.28)$$

whence

$$f' = \frac{d\Delta x}{\lambda} = \frac{3.84 \times 310^{-2}}{6.9 \times 10^{-4}} = 166.96 \text{ mm} \cong 16.7 \text{ cm}, \quad (1.29)$$

and thus the overall length of the correlator is

$$4f' = 4 \times 16.696 = 66.78 \text{ cm}. \quad (1.30)$$

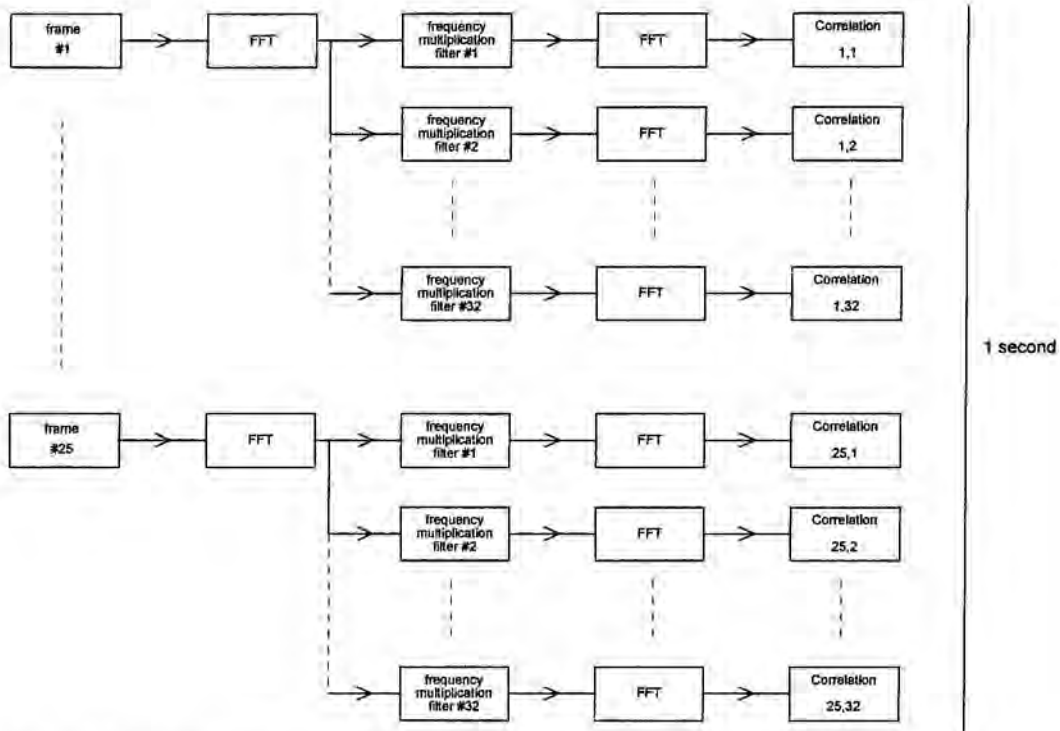
To travel this distance, the light will need:

$$\Delta t = \frac{66.78}{3 \times 10^{10}} = 2.23 \times 10^{-11} \text{ s}, \quad (1.31)$$

which, even if pipelining is not used, equals to the huge frame rate of:

$$v = \frac{1}{\Delta t} = \frac{1}{2.23 \times 10^{-11}} \text{ s}^{-1} = 450 \text{ MHz}. \quad (1.32)$$

and therefore the maximum frame rate is 800 Hz imposed by the camera, the slowest device in the system. If we wish to perform the identification of an object at video rate (25 Hz), this means that we can perform up to 32 correlations per input image ( $32 \times 25 = 800$ ). In practice the SPOTR correlator uses composite filters that are able to include the information of 180 images of the target, which results in  $180 \times 32 = 5760$  views. When the range of viewing angles and distances from which the object may be observed is large this number is not enough; for example, in the initial tests of the correlator was found that a 800 filter database was necessary -a total of 144000 images to characterize the 3-D target- and thus only one identification per second was possible. Since the number of equivalent



**Figure 1.10-** SPOTR correlator digital equivalent operating at video rate.

flops is only slightly different for the two cases we will compute them when the correlator operates performing 32 correlations per input frame at video rate (25 frames per second).

To reproduce the results obtained by the SPOTR correlator by means of digital electronic hardware we need to carry out the Fourier transform of the input scene, multiply it by the precomputed filters on a frequency by frequency basis and obtain the correlations by computing the inverse Fourier transform of those products as sketched in Figure 1.10.

Unfortunately for Optics there exist very efficient algorithms for the digital computation of Fourier transforms known as FFT algorithms -Fast Fourier Transform-, capable of reducing the apparent complexity of  $N^2$  operations to roughly  $N \log_2 N$  for an  $N$ -point sequence. In particular, when the number of points is a power of two, namely  $N=2^k$ , we can use the so-called Radix-2 FFT algorithm, which according to [Nus82] needs for an  $N$ -point complex transform:

$$\begin{aligned}
 M &= \frac{N}{2}(-10 + 3\log_2 N) + 8 \\
 A &= \frac{N}{2}(-10 + 7\log_2 N) + 8
 \end{aligned}
 \tag{1.33}$$

nontrivial real operations, where  $M$  and  $A$  are the number of multiplications and additions respectively. Operations such as multiplications by  $\pm 1$ ,  $\pm j$ , etc, are not taken into account as they imply in the worst case only a change of sign, thus being much less involved than true floating point operations.

The two-dimensional Fourier transform of the  $N \times N$  input scene is carried out by performing the  $N$  one-dimensional row transforms -of  $N$  points each- and subsequently the  $N$  column transforms. Let  $r_{ik}$  be the  $i$  row. Equation 1.33 gives the computational complexity of an  $N$ -point complex one-dimensional Fourier transform. Since the input scene is real we can compute two row transforms with the same number of operations as one complex [Nus82] by forming the auxiliary sequences:

$$\begin{aligned}
 s_{1,k} &= r_{1,k} + jr_{2,k} \\
 s_{p,k} &= r_{2p-1,k} + jr_{2p,k} \\
 s_{N/2,k} &= r_{N-1,k} + jr_{N,k}
 \end{aligned}
 \tag{1.34}$$

that is to say,  $N/2$  complex transforms for which we need:

$$\begin{aligned}
 M_1 &= \frac{N^2}{4}(-10 + 3\log_2 N) + 4N \\
 A_1 &= \frac{N^2}{4}(-10 + 7\log_2 N) + 4N
 \end{aligned}
 \tag{1.35}$$

The individual rows transform can then be separated by using the symmetry properties of the Fourier transforms of pure real and imaginary sequences:

$$\begin{aligned}
\operatorname{Re}\{R_{2p-1,k}\} &= (\operatorname{Re}\{S_{p,k}\} + \operatorname{Re}\{S_{p,-k}\}) / 2 \\
\operatorname{Im}\{R_{2p-1,k}\} &= (\operatorname{Im}\{S_{p,k}\} - \operatorname{Im}\{S_{p,-k}\}) / 2 \\
\operatorname{Re}\{R_{2p,k}\} &= (\operatorname{Im}\{S_{p,k}\} + \operatorname{Im}\{S_{p,-k}\}) / 2 \\
\operatorname{Im}\{R_{2p,k}\} &= (\operatorname{Re}\{S_{p,-k}\} - \operatorname{Re}\{S_{p,k}\}) / 2
\end{aligned} \tag{1.36}$$

at the cost of  $4N \times N / 2 = 2N^2$  additions, where  $\operatorname{Re}()$  and  $\operatorname{Im}()$  symbolize the real and imaginary parts and capital letters represent the Fourier transforms of the corresponding lower-case sequences. Notice that the multiplications by  $1/2$  are not considered because it is an irrelevant -for our purposes- proportionality factor affecting all the transforms. Therefore we have so far:

$$\begin{aligned}
M_2 &= \frac{N^2}{4}(-10 + 3\log_2 N) + 4N \\
A_2 &= \frac{N^2}{4}(-10 + 7\log_2 N) + 4N + 2N^2
\end{aligned} \tag{1.37}$$

The computation of the column transforms is also simplified by the real nature of the input scene. In this case, the real and imaginary parts of the two-dimensional Fourier transform have even and odd symmetry respectively:

$$F_{i,k} = F_{-i,-k}^* = F_{N-i,N-k}^* \quad i, k = 1, \dots, \frac{N}{2} \tag{1.38}$$

and thus only the first  $N/2$  transforms -now complex- must be calculated making a total of:

$$\begin{aligned}
 M_3 &= \frac{N^2}{2}(-10 + 3\log_2 N) + 8N \\
 A_3 &= \frac{N^2}{2}(-6 + 7\log_2 N) + 8N
 \end{aligned}
 \tag{1.39}$$

As the SPOTR correlator uses an FLC device with only two modulation states, +1 and -1, the frequency-to-frequency multiplication between the scene transform and the filters is accomplished with  $N^2$  trivial multiplications, which are consequently discarded. Finally to carry out the Fourier transform of this product, and considering that the filter, and therefore the correlation, may be complex in direct space, we will need  $N$  complex row transforms followed by  $N$  complex column transforms, namely:

$$\begin{aligned}
 M_4 &= N^2(-10 + 3\log_2 N) + 16N \\
 A_4 &= N^2(-10 + 7\log_2 N) + 16N
 \end{aligned}
 \tag{1.40}$$

Finally, to carry out the 32 correlations per frame at a 25 Hz frame rate we need -25 real FFT (input scene) + 32x25 complex FFT (product):

$$\begin{aligned}
 25M_3 + 800M_4 &= \frac{N^2}{2}(-16250 + 4875\log_2 N) + 13000N \\
 25A_3 + 800A_4 &= \frac{N^2}{2}(-16150 + 11375\log_2 N) + 13000N
 \end{aligned}
 \tag{1.41}$$

whence the total number of operations per second is:

$$\text{flops} = \frac{N^2}{2}(-32400 + 16250\log_2 N) + 26000N
 \tag{1.42}$$

Connection Machine-CM2 with 2048 processors (peak value)	20 Gigaflops
IBM SP2 with 12 nodes (peak value)	3,2 Gigaflops
IBM 3090/600 J with 6 processors (? value)	828 Megaflops
<b>SPOTR correlator</b> (theoretical value)	<b>670 Megaflops</b>
CRAY Y-MP 232 with two processors (? value)	666 Megaflops
16x16 mesh of Inmos T800 transputers (peak value)	384 Megaflops
CRAY-1 (peak value)	160 Megaflops
Data Translation DT 2878 Advanced Processor-image processing card (nominal value)	25 Megaflops
My home computer-Intel 486 DX2 at 66 Mhz (computed with LINPACK-Argonne test)	2,2 Megaflops

*Table 1.1- Computational power comparison between several systems.*

and since the bandwidth of the modulators is  $N=128$  we have:

$$\text{flops} = 669747200 \approx 670 \text{ Megaflops} \quad (1.43)$$

which is about the same computational power as a Cray Y-MP 232 supercomputer working in parallel with two processors (666 Megaflops) and roughly 27 times the capacity of off-the-shelf image processing electronic hardware as Table 1.1 summarizes.

In conclusion, because of vision problems are characterized by the necessity of processing a large amount of information in very short times, and since they do not need a high precision in the associated computations - frequently a yes/no result is enough- they are well suited for Optics, an analog and intrinsically parallel technology. In the particular case of correlation, a relatively simple optical setup can compete with parallel supercomputers outperforming them, probably by some orders of magnitude, in cost, size and power consumption.

## Chapter two. Filter design for VanderLugt correlators

### Introduction.

The measure of similarity between patterns of equal energy implicit in the correlation function can be used, as explained in the first chapter, to determine the position of a given object inside a complex scene. The obvious way to do this consists of using as a filter the pattern to be detected, which will naturally produce a maximum correlation at the position in which the image is placed. But this is not only possibility since we can use as a filter a model of the object, perhaps with only its salient features such as the edges and distinctive details.

The information contained by an image is in general very much redundant, a well known fact that is employed in compression algorithms. These algorithms routinely attain large compression ratios [Son94] by exploiting the repeated appearance of strings of data that are frequent in natural images. This effect is even more patent when we don't have to recover the whole image but we are interested in distinguishing a pattern from a different one. In a such a case, and if there is no noise blurring the images, the problem can be trivially solved just by using one of the pixels where the two patterns differ. For a more complicated problem involving noise and more images we will need the use of more information, but in general, the whole target picture will not be necessary. This fact gives us some flexibility in the correlation filters we can use for a reliable detection.

VanderLugt in 1964 proposed the first feasible implementation of a correlation filter by means of the, in those days, recently discovered technique of off-axis holography. Since then and in an almost trial and error process at the beginning, hundreds of different filters have been developed and enhanced until the systematic approach of the optimization procedures emerged as a synthesis.



Filter design for optical correlators is currently a well founded discipline to which this chapter is devoted.

The natural domain in designing filters for VanderLugt correlators is the Fourier space because the correlation product is carried out by means of a frequency filtering process. However the object space is sometimes more useful when the filters are intended for a different architecture such as a joint transform correlator, which works with a filter in direct space or an acoustooptic correlator that performs the correlation directly without using Fourier transforms. An specific design in direct space may overcome some of the difficulties raised by the frequency representation, for example the circularity of the correlations. Inasmuch as we deal only with VanderLugt correlators the Fourier representation is implicitly assumed in what follows unless otherwise specified.

## 2.1 Previous considerations: notation, procedures and quality criteria.

We begin this chapter by making clear the conventions we adopt to express and derive the different results. First of all, the two representations, discrete and continuous, will be used indistinctly passing from one another when mathematical convenience so demands. In general, single-image filters are described in the continuous representation whereas the multiple-image designs, to avoid the Hilbert formalism, will be represented by discrete magnitudes. This will allow us to use the more simple theorems of linear algebra.

To unify the notation we adopt the following conventions:

$x, y$  : represent spatial coordinates in object space.

$u, v$  : represent frequency coordinates in Fourier space.

$h(x,y)$  : represent images in object space and

$H(u,v)$  : in capital letters denotes the Fourier transform of  $h(x,y)$ .

$\theta, \phi(u,v)$  : Greek letters are used to represent angles and phases.

and whenever discrete notation is used:



$x$  : normal letters denote scalar variables.

$\mathbf{x}, x_i$  : bold case letters represent vectors and  $x_i$  its  $i$ -component.

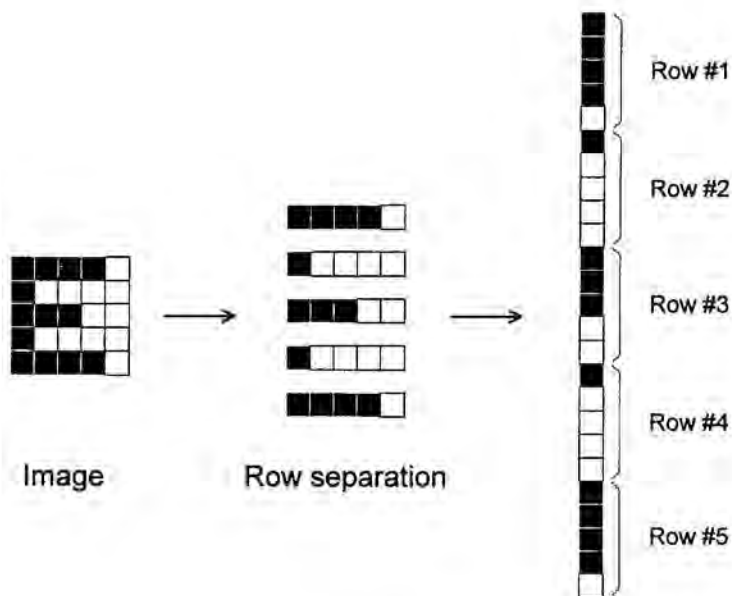
$\tilde{\mathbf{x}}, \tilde{x}_i$  : are vectors and components in Fourier space.

$\mathbf{X}, X_{ij}$  : a capital bold case letter denotes a matrix with components  $X_{ij}$ .

$\tilde{\mathbf{X}}, \tilde{X}_{ij}$  : represent a matrix in Fourier space.

Images in the discrete representation are treated as vectors using lexicographic scanning. Such procedure, as sketched in Figure 2.1, enables us to construct a vector of  $N^2$  components from an image of  $N \times N$  pixels by successively stacking its  $N$  row vectors. With this technique, the correlation can be written as a scalar product allowing the introduction of the powerful concepts of vector and metric spaces. Indeed, the central correlation between the filter  $h(x,y)$  and the image  $r(x,y)$ :

$$c(0,0) = \iint h^*(x,y) r(x,y) dx dy = \iint H^*(u,v) R(u,v) du dv \quad (2.1)$$



Lexicographic scanning

*Figure 2.1- Conversion of a 5x5 image into a vector of 25 components by means of lexicographic scanning.*

can be rewritten for discrete functions as:

$$c_{0,0} = \sum_i h_i^* r_i = \mathbf{h}^+ \mathbf{r} \quad (2.2)$$

where the superscript + means conjugate transpose. Equivalently, expressed in terms of their Fourier transforms we have:

$$c_{0,0} = \frac{1}{N} \sum_i \tilde{h}_i^* \tilde{r}_i = \frac{1}{N} \tilde{\mathbf{h}}^+ \tilde{\mathbf{r}} \quad (2.3)$$

where  $N$  is the image bandwidth.

To illustrate the different results several simulations have been carried out. In all of them, the correlation has been implemented using an FFT routine of the Brenner type found in [Pre88].

Noise is assumed to be additive, Gaussian, stationary -the statistical properties do not vary from point to point- and zeromean. When a white spectrum is supposed it is explicitly mentioned. This model for the noise is the most simple one, so it is worth mentioning that more realistic cases are being currently addressed such as non-Gaussian models [Kas87][Vij89a] or nonoverlapping -with respect to the signal- noise as a first approximation to clutter [Jav93][Ref93].

Finally, mostly following Vijaya Kumar in [Vij90a] we define several quality criteria which enables us to assess the performance of the different filters and consequently to search, compare and select the most appropriate ones. They are:

a) Signal-to-noise ratio (SNR), which is a typical figure-of-merit in signal processing to measure the quality and reliability of a system. Here, the signal is the correlation peak, which for simplicity is supposed to be at the center of the output plane. It is defined as:

$$\text{SNR} = \frac{|E\{c(0,0)\}|^2}{\text{var}\{c(0,0)\}} \quad (2.4)$$

where  $E\{\}$  is the mathematical expectation or more simply the mean and  $var\{\}$  represents the variance of the correlation peak. Large signal-to-noise ratios represent small fluctuations of the correlation signal compared with the total height of the peak, and thus this parameter should be kept as high as possible when thresholding is used to determine their position. Since the noise is additive in our model, the input image  $r(x,y)$  can be written as:

$$r(x,y) = s(x,y) + n(x,y) \quad (2.5)$$

where  $s(x,y)$  is the noncorrupted image and  $n(x,y)$  a noise sample realization. The correlation with filter  $h$  is therefore:

$$\mathbf{h}^+ \mathbf{r} = \mathbf{h}^+ \mathbf{s} + \mathbf{h}^+ \mathbf{n} \quad (2.6)$$

and thus the mean is:

$$E(\mathbf{h}^+ \mathbf{r}) = E(\mathbf{h}^+ \mathbf{s}) + E(\mathbf{h}^+ \mathbf{n}) = \mathbf{h}^+ \mathbf{s} + E(\mathbf{h}^+ \mathbf{n}) \quad (2.7)$$

and since the noise is Gaussian and zeromean we have:

$$E(\mathbf{h}^+ \mathbf{n}) = \mathbf{h}^+ E(\mathbf{n}) = \mathbf{h}^+ \mathbf{0} = 0 \quad (2.8)$$

which after substituting into equation 2.7 gives:

$$E(\mathbf{h}^+ \mathbf{r}) = \mathbf{h}^+ \mathbf{s} \quad (2.9)$$

The variance can also be expressed as a function of the filter components in a simple way:

$$var\{\mathbf{h}^+ \mathbf{r}\} = E\left[\left(\mathbf{h}^+ \mathbf{r} - E(\mathbf{h}^+ \mathbf{r})\right)^2\right] = E\left(\mathbf{h}^+ \mathbf{r}^2\right) - E^2(\mathbf{h}^+ \mathbf{r}) \quad (2.10)$$

and applying again that the noise is Gaussian and zeromean:

$$\begin{aligned} \text{var}\{\mathbf{h}^+ \mathbf{r}\} &= |\mathbf{h}^+ \mathbf{s}|^2 + E\left(|\mathbf{h}^+ \mathbf{n}|^2\right) - |\mathbf{h}^+ \mathbf{s}|^2 = E\left(|\mathbf{h}^+ \mathbf{n}|^2\right) = \\ &= E\left((\mathbf{h}^+ \mathbf{n})(\mathbf{h}^+ \mathbf{n})^*\right) = E[\mathbf{h}^+ \mathbf{n} \mathbf{n}^+ \mathbf{h}] = \mathbf{h}^+ E(\mathbf{n} \mathbf{n}^+) \mathbf{h} = \mathbf{h}^+ \mathbf{S} \mathbf{h} \end{aligned} \quad (2.11)$$

where

$$\mathbf{S} = E(\mathbf{n} \mathbf{n}^+) \quad (2.12)$$

is known as the correlation matrix, the second order matrix or -for zeromean noise- the variance-covariance matrix [Pic93]. For an stationary noise, the Fourier transform,  $\tilde{\mathbf{S}}$ , of the covariance matrix is diagonal, being its entries the components of the power spectrum of that noise.

Finally, the signal-to-noise ratio can be written as:

$$\text{SNR} = \frac{|\mathbf{h}^+ \mathbf{s}|^2}{\mathbf{h}^+ \mathbf{S} \mathbf{h}} = \frac{|\tilde{\mathbf{h}}^+ \tilde{\mathbf{s}}|^2}{N \tilde{\mathbf{h}}^+ \tilde{\mathbf{S}} \tilde{\mathbf{h}}} \quad (2.13)$$

or equivalently in terms of continuous functions:

$$\text{SNR} = \alpha \frac{\left| \iint H^*(u, v) S(u, v) du dv \right|^2}{\iint P_n(u, v) |H(u, v)|^2 du dv} \quad (2.14)$$

where  $P_n(u, v)$  is the power spectrum of the noise and  $\alpha$  a scaling constant.

b) Horner efficiency: This is a quality criterion specifically defined to address the optical properties of correlation filters. In particular, Horner efficiency measures the fraction of the total energy present at the input plane passing through

the filter and forming the correlation so this parameter should be maximized. It is defined as:

$$\eta = \frac{\text{Total energy at the correlation plane}}{\text{Total energy at the input plane}} = \frac{\iint |c(x,y)|^2 dx dy}{\iint |r(x,y)|^2 dx dy} \quad (2.15)$$

which using the Parseval's theorem -see Chapter 1- can be written in Fourier space as:

$$\eta = \frac{\iint |H(u,v)|^2 |R(u,v)|^2 du dv}{\iint |R(u,v)|^2 du dv} \quad (2.16)$$

Notice that the optical systems -holograms, modulators, etc- used to display the filters are passive, and so their transmittance are such that  $|H(u,v)| \leq 1$  and consequently:

$$\eta \leq 1 \quad (2.17)$$

c) Peak-to-correlation energy: A high Horner efficiency is not enough to guarantee a large detection signal when the filter produces broad correlation peaks or if large secondary lobes are present. To measure the fraction of the correlation energy going to form a detection maximum, the peak-to-correlation energy is defined:

$$\text{PCE} = \frac{\text{Energy of the peak}}{\text{Energy of the correlation plane}} = \frac{|c(0,0)|^2}{\iint |c(x,y)|^2 dx dy} \quad (2.18)$$

or in terms of Fourier transforms:

$$\text{PCE} = \alpha \frac{\left| \iint H^*(u, v) R(u, v) \, du \, dv \right|^2}{\iint |H(u, v)|^2 |R(u, v)|^2 \, du \, dv} \quad (2.19)$$

where  $\alpha$  is a normalizing constant. The expression of the peak-to-correlation energy takes a compact form when written in discrete notation which is of great interest in designing multiple-image filters. Such expression is:

$$\text{PCE} = \frac{|\tilde{\mathbf{h}} + \tilde{\mathbf{r}}|^2}{N \sum_i |\tilde{h}_i \tilde{r}_i|^2} = \frac{|\tilde{\mathbf{h}} + \tilde{\mathbf{r}}|^2}{N \sum_i |\tilde{h}_i|^2 |\tilde{r}_i|^2} = \frac{|\tilde{\mathbf{h}} + \tilde{\mathbf{r}}|^2}{N \tilde{\mathbf{h}} + \tilde{\mathbf{D}} \tilde{\mathbf{h}}} \quad (2.20)$$

where  $\tilde{\mathbf{D}}$  is the diagonal matrix:

$$\tilde{\mathbf{D}} = \begin{pmatrix} |\tilde{r}_1|^2 & 0 & \cdot & 0 \\ 0 & |\tilde{r}_2|^2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & |\tilde{r}_N|^2 \end{pmatrix} \quad (2.21)$$

Again this parameter should be maximized to obtain sharp peaks and small sidelobes.



## 2.2 Filter design for a reliable detection. Necessary conditions.

Chapter 1 introduced the correlation as a useful technique for pattern recognition based on intuitive reasonings. Recalling those arguments, we said that the correlation is large when the high values of the two functions being correlated coincide and therefore when they are similar. However one could legitimately ask what would happen when one of several similar objects is to be identified and its position determined. Is it possible to find a filter capable of doing so or correlation is only able to discriminate patterns in trivial cases when the objects are very different?. As we saw, when the filter is a constant -or there is no filter- an optical correlator is merely an imaging system that, of course, gives no distinctive information about the input images so which filters, if any, ensure a reliable identification?. To answer this, we need first an explicit definition of what a reliable filter is and a mathematical condition characterizing them. A reasonable possibility is the following:

When the detection procedure of a given object is implemented by means of the correlation function and the subsequent thresholding of the output plane, the necessary conditions that a filter must fulfill for a reliable identification are:

- a) The correlation between the filter and the target image must be higher than the correlation between that filter and any other image of the same energy.
- b) The maximum of the correlation between the filter and the target image must indicate its position.

With these conditions the appearance of a clear maximum in the output plane would indicate the presence and position(s) of the image of interest in the input scene. Furthermore a threshold can be selected to eliminate the correlations with the nontarget objects. If the output plane shows no sharp peak, the image being considered is not present. It is worth pointing out that the utilization of threshold detectors, although being the easiest, is not the only possibility for interpreting the output plane. The height of the correlation maxima may induce to errors in the identification of an image in an adverse environment with changes in the illumination conditions. For this reason the shape of the peaks has been proposed [Jen87] as a new element of decision. Filter designs such as G-MACE [Cas91] and MSE-SDF [Vij92a] that control the form of the peaks have also been developed. Furthermore the hard decision logic implied by the binarization may be

replaced by soft decision rules such as those used in fuzzy logic or by using neural networks [Raj92]. These different approaches would lead to different definitions of acceptable filters. However, we will not consider them in the following argumentation.

When only the height of the correlation peak is taken into account, the two constraints that we have established ensure the detection. The first condition enables us to distinguish the target pattern from other similar objects and the second one to accurately determine its position inside a large scene. Both can be mathematically expressed in a convenient way through the following

### Proposition

Let  $f(x,y)$  be the pattern to be identified. Let us choose a point of this image that will be referred to as the image center and will be used to define the position of  $f(x,y)$ . In this way, when we say that the image is centered or is placed at the point  $(x_0,y_0)$  we mean that its center is located there. Then, if

$$F(u,v) = |F(u,v)| e^{j\phi_1(u,v)} \quad (2.22)$$

is the Fourier transform of  $f(x,y)$  centered at the origin of coordinates and

$$H(u,v) = |H(u,v)| e^{j\phi_2(u,v)} \quad (2.23)$$

the Fourier transform of the filter  $h(x,y)$  then:

- The sufficient condition -condition b) above- for the filter to give a maximum at the center of the correlation plane is:

$$\phi_1(u,v) = \phi_2(u,v) \quad \forall u,v \quad (2.24)$$

- If the latter requisite is met, the necessary condition -a) above- to obtain a higher correlation with the pattern  $f(x,y)$  than with any other of the same energy is:

$$|H(u,v)| = |F(u,v)| \quad \forall u,v \quad (2.25)$$

Proof.

The intensity of the correlation between the image and the filter can be expressed as

$$\left| [h \star f]_{(x,y)} \right|^2 = \left| \iint H^*(u,v) F(u,v) e^{j(xu+yv)} dudv \right|^2 \quad (2.26)$$

by means of the correlation theorem.

If we express both  $H$  and  $F$  in terms of their magnitudes and phases the correlation can be written as:

$$\left| [h \star f]_{(x,y)} \right|^2 = \left| \iint |H(u,v)| |F(u,v)| e^{j(\phi_2(u,v) - \phi_1(u,v))} e^{j(xu+yv)} dudv \right|^2 \quad (2.27)$$

The maximum value of this expression with respect to the phase, is easily found by considering the complex values as two-dimensional vectors and applying to them the triangle inequality:

$$\begin{aligned} \left| [h \star f]_{(x,y)} \right|^2 &= \alpha \left| \iint |H(u,v)| |F(u,v)| e^{j(\phi_2(u,v) - \phi_1(u,v))} e^{j(xu+yv)} dudv \right|^2 \leq \\ &\leq \alpha \left( \iint |H(u,v)| |F(u,v)| e^{j(\phi_2(u,v) - \phi_1(u,v))} e^{j(xu+yv)} dudv \right)^2 = \\ &= \alpha \left( \iint |H(u,v)| |F(u,v)| dudv \right)^2 = \alpha \left| \iint |H(u,v)| |F(u,v)| dudv \right|^2 \end{aligned} \quad (2.28)$$

where the equality holds if and only if:

$$e^{j(xu+yv+\phi_2(u,v)-\phi_1(u,v))} = e^{j\beta} \quad \forall u,v \quad (2.29)$$

where  $\beta$  is an arbitrary constant.

Since we are looking for the condition ensuring a maximum to appear at the center of the correlation plane, i.e.:

$$x = y = 0 \quad (2.30)$$

it is enough if we demand:

$$\phi_2(u, v) = \phi_1(u, v) + \beta \quad (2.31)$$

The constant  $\beta$  is of little importance because it merely represents a different phase origin between the Fourier transforms of target and filter and can be set to zero with no loss of generality. By doing so, the first part of the proposition is proved.

A few comments about this result are necessary. First, the condition is sufficient but not necessary. It could be possible to obtain the maximum of the correlation plane at the center with a different phase distribution. Notice that we have established the condition to obtain the maximum possible correlation -with a given amplitude- at the center, which obviously implies that the rest of the correlation plane is not higher. However, in many cases, for example with some SDF filters, the center is also the highest with respect to the correlation plane although it is not the maximum achievable value for the correlation with that particular input image. The possibility of establishing a less stringent requisite will be later used in deriving the multiple-image filters.

On the other hand the maximum is not necessarily unique and several points may take the same value as the center. For example when the input image is a periodic pattern, the correlation is also periodic and the maximum appears at regular intervals. However the position of a periodic image is loosely defined since such images are indistinguishable when they are displaced exactly one or several periods. Thus the problem seems to be a consequence of pathological images rather than a deficiency of the recognition procedure.

The second point is proved as follows. The intensity of the correlation between two different images of the same energy can be written as:

$$\begin{aligned}
| [h \star f]_{(x,y)} |^2 &= \alpha \left| \iint |H(u,v)| |F(u,v)| e^{j(\phi_2 - \phi_1 + xu + yv)} dudv \right|^2 \leq \\
&\leq \alpha \left( \iint |H(u,v)|^2 dudv \right) \left( \iint |F(u,v)|^2 dudv \right) = \alpha \left( \iint |H(u,v)|^2 dudv \right)^2 = \\
&= | [h \star h]_{(0,0)} |^2
\end{aligned} \tag{2.32}$$

where the first step has been derived by using the Cauchy-Schwartz inequality and the second one by imposing the equality of the energies. Since the latter integral is the intensity of the central autocorrelation the proposition is proved.

Expressed in words this result permits us to answer the questions raised at the beginning of the section: it is always possible to discriminate a given object from others of similar shape -in fact no matter how similar- and determine its position if the filter is the Fourier transform of the target object. Thus in this sense correlation is a well defined operation and can be used with confidence in pattern recognition problems.

The Fourier transform of the target object, when considered as a correlation filter, is a particular case of the so-called matched filter and has a number of additional interesting properties. In particular for white noise the signal-to-noise ratio is a maximum:

$$\begin{aligned}
\text{SNR}_{H(u,v)} &= \frac{\left| \iint H^*(u,v) S(u,v) dudv \right|^2}{\iint P_n(u,v) |H(u,v)|^2 dudv} \leq \frac{\iint |H(u,v)|^2 dudv \iint |S(u,v)|^2 dudv}{P_n \iint |H(u,v)|^2 dudv} = \\
&= \frac{\iint |S(u,v)|^2 dudv}{P_n} = \frac{\left| \iint |S(u,v)|^2 dudv \right|^2}{P_n \iint |S(u,v)|^2 dudv} = \text{SNR}_{S(u,v)}
\end{aligned} \tag{2.33}$$



where the Cauchy-Schwartz formula has again been used. Furthermore, the noise induced peak location error is a minimum, see [Vij92b] for the details and the filter can be easily and accurately implemented in practice thanks to optical holography following the VanderLugt method [Van64].

Matched filters have also important drawbacks, mainly their overspecialization on noise robustness as we will further analyze. This makes the filter to rank poorly in Horner efficiency and specially in the peak-to-correlation energy criterion. This overspecialization limits the range of possible applications of the correlation based techniques to problems in which noise is the main concern.

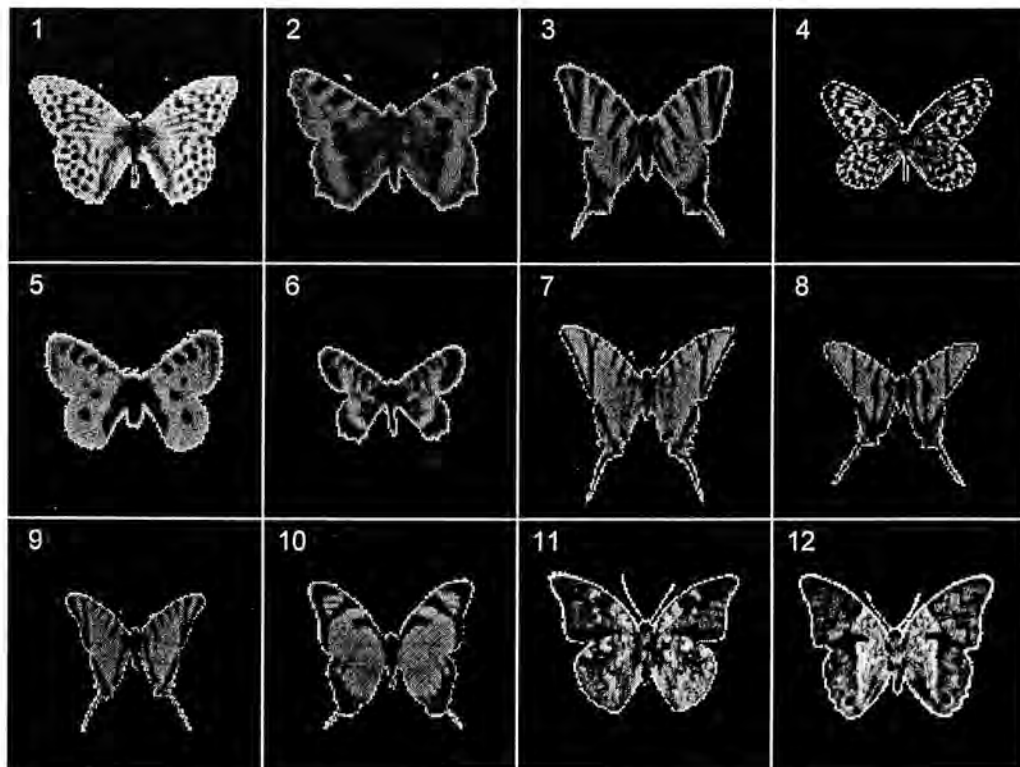
However, one would wish to have different choices to be applied to different situations. For example, a filter giving sharp peaks when noise is not too severe but multiple targets may appear in the field of view or perhaps, a very efficient filter to be used jointly with a low-power laser in a portable correlator. To get this flexibility we must renounce to the maximum correlation condition, passing from having a total certainty that the target correlation is the highest to a reasonable probability. In fact, this condition is too strict. Think, for example in a problem that consists in sorting a set of cutting tools, a practical application solved by an optical correlator as reported in [Caw89] and [Raj92]. The optical processor was developed as part of an ESPRIT program, the high-technology initiative promoted and funded by the European Union, being the final user the German company Fried Krupp GmbH. To reduce the fabrication costs, the company was interested in achieving a dense packing of different tools in a holder where the pieces had to undergo a coating process but avoiding the contact between them. The optical correlator had therefore to classify the pieces and provide information about their position and orientation. Here, the number of images that can appear in the field of view is limited and differ from the ideal shapes within very strict margins -the tolerance of the fabrication process-. In such a case it is unnecessarily demanding to prevent that, say, a set of butterflies give a high correlation with the filters used to identify the tools. Simply, it is unlikely to see a butterfly in such a place.

This matches with the intuitive notion, as commented in the introduction, that the totality of the information of an image is in general not necessary to identify it. For the specific case of using the correlation as the means to carry out

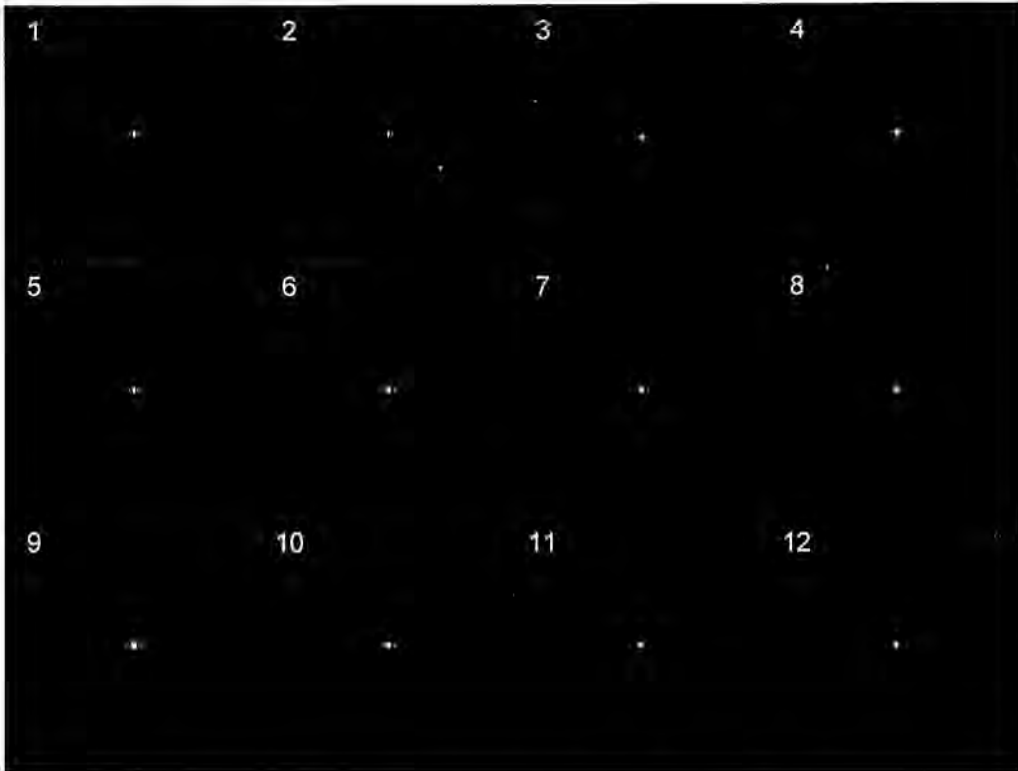


such identification, it may suffice if we use as a filter a model of the target object, that is some version containing only relevant data. This possibility would permit us the utilization of different filters.

The work of Oppenheim and Lim [Opp81] gives the theoretical background to put this qualitative arguments into a more formal basis. In the cited reference, the authors analyze the information content of both, the magnitude and phase of the Fourier transform of natural images, concluding that it is not equally distributed. The magnitude of natural images seems to be very similar thus carrying little specific information about them and consequently it is the phase what really makes the difference. This fact is illustrated in the following example. Figure 2.2 shows a set of twelve different butterflies and Figure 2.2 the magnitude of their respective Fourier transforms; all twelve show a similar behavior, namely a concentration of energy in the lower frequencies giving delta-like distributions. It is worth pointing out that this behavior is the origin of the broad correlation peaks given by matched filters.



*Figure 2.2- Test images.*



*Figure 2.3- Magnitude of the Fourier transforms of butterflies in Figure 2.2.*

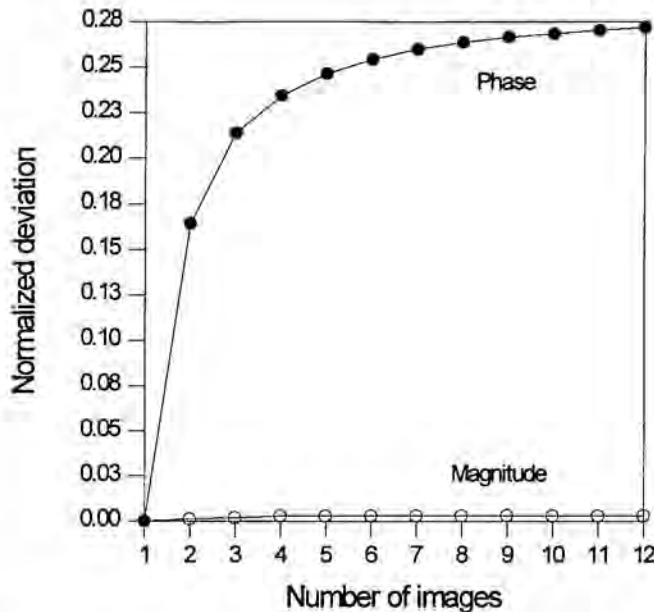
Figure 2.4 gives a numerical comparison of the information content of both, magnitude and phase of a set of images. The graph has been obtained by adding successively a new image of the set in Fig. 2.2 and computing for the two quantities of interest the following deviation:

$$\sigma_s = \frac{\frac{1}{N} \sum_{k=1}^N \left( \frac{1}{N_i^s} \sum_{j=1}^{N_i^s} (x_j[k] - m^s[k])^2 \right)}{\max_{j,k} (x_j[k])} \quad (2.34)$$

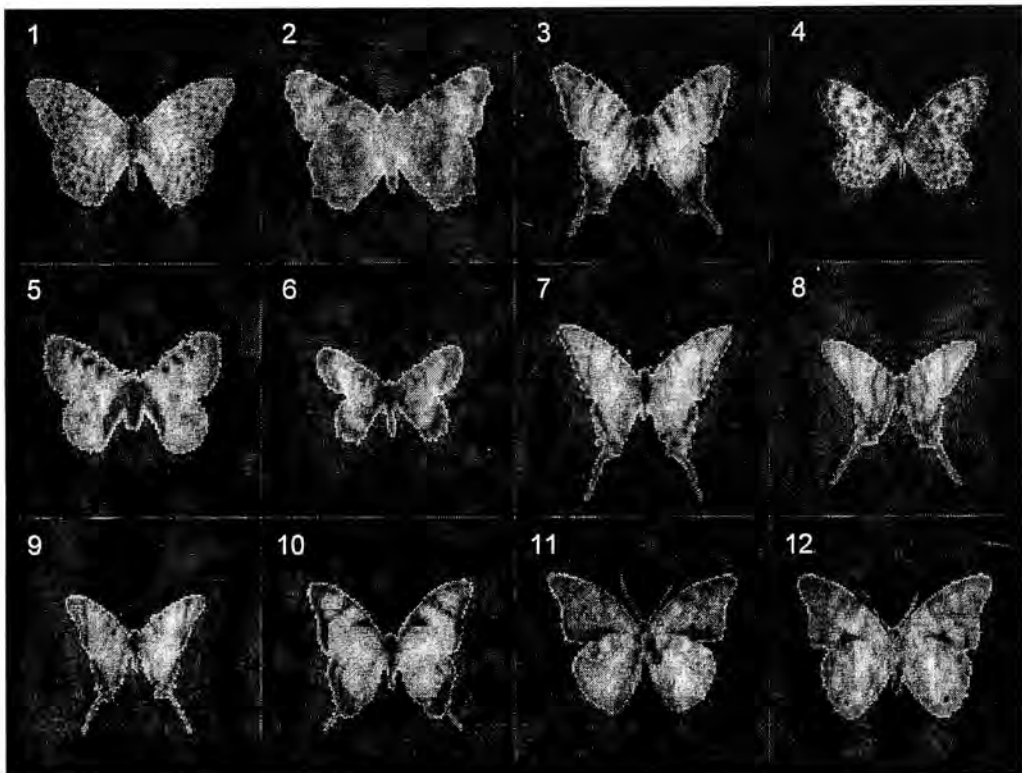
where  $N$  is the bandwidth of the images,  $N_t^S$  the number of images being considered,  $x_j[k]$  the  $k$ th pixel of the image  $j$ , and finally  $m^S[k]$  the mean value of either magnitude or phase of pixel  $k$  for the set of images.

For a given number of patterns, say the first five butterflies, the quantity inside the parenthesis is a measure, for each pixel, of the deviation of the phase and magnitude with respect to the average value of these same quantities for the five butterflies. For example, if for a given pixel the magnitude of all butterflies considered is the same, the parenthesis is zero. This gives then an indication of how varies the magnitude -or the phase- in that pixel for the set of five butterflies. Furthermore, when we add a new image, the new value of the parenthesis reflects the amount of distinctive information introduced. The rest of the expression adds up the variation of all the pixels and normalize it to the maximum value.

The plot in Figure 2.2 shows, although this should be taken with caution since we are reducing the complexity of a set of images to a single numerical value -see [Juv91] for possible exceptions-, that the average variation of the phase is much important than that of the magnitude and consequently that the phase carries most of the specific information of an image.



*Figure 2.4- Normalized variance vs. number of images of the magnitude and phase.*



*Figure 2.5- Reconstructed images.*

A further example, this one closer to our discussion about correlation filters as we will see, is shown in figure 2.5. In this case the twelve butterflies have been reconstructed from Fourier transforms made by conserving their specific phases but by changing all the magnitude distributions to a common one, the average of the set. As can be observed, except for a nonuniform background and a small loss of details, all the images resemble the original ones, and in particular they can be distinguished from similar companions.

The consequence of this asymmetric distribution of the information between the two parts of a Fourier transform, enables us to use with quite precision the approximation that the amplitude distribution of a set of images is the same and does not change from pattern to pattern. In such a case, the two conditions we established to ensure a reliable detection can be met by only maintaining the phase of the target pattern, namely we are allowed to use filters of the type:

$$H(u, v) = |H(u, v)| e^{j\phi_1(u, v)} \quad (2.35)$$

where  $\phi_1(u, v)$  is the phase of the target pattern and  $|H(u, v)|$  an arbitrary positive function. Indeed, let

$$F_1(u, v) = |F(u, v)| e^{j\phi_1(u, v)} \quad (2.36)$$

be the target pattern and

$$F_2(u, v) = |F(u, v)| e^{j\phi_2(u, v)} \quad (2.37)$$

an arbitrary pattern to be discriminated against. Let

$$H(u, v) = |H(u, v)| e^{j\phi_1(u, v)} \quad (2.38)$$

be the filter. We trivially have:

$$\begin{aligned} \left| [h \star f_2]_{(x, y)} \right| &= \left| \iint |H(u, v)| |F(u, v)| e^{j(\phi_2 - \phi_1 + xu + yv)} du dv \right| \leq \\ &\leq \left| \iint |H(u, v)| |F(u, v)| du dv \right| = \left| [h \star f_1]_{(0, 0)} \right| \end{aligned} \quad (2.39)$$

no matter the magnitude of the filter.

Depending on the accuracy of the equal-magnitude assumption we will have a more or less reliable identification. As we saw, for natural images this can be supposed with wide generality, giving us the needed flexibility in the filter design at the cost of a lower confidence in the recognition process.

### 2.3 Design of single-image filters.

The possibility of using arbitrary functions for the filter amplitude was soon realized by Horner and Gianino [Hor84], who proposed a filter to maximize the optical efficiency. This work stimulated the interest in filter design attracting other authors, who developed a long list of new modifications of the matched filter. An incomplete inventory could be as follows:

- Phase-only filter (POF), [Hor84]
- Binary phase-only filter (BPOF), [Hor85a]
- Optimal filter (OF), [Yar86],
- Amplitude compensated matched filter (ACMF), [Mu88]
- Optimal phase-only filter (OPOF), [Vij89b]
- Phase-mostly filter (PMF), [Jud89]
- Quad-phase filter (QPF), [Dic89a]
- Complex ternary matched filter (CTMF), [Dic90]
- Phase-with-constrained-magnitude filter (PCMF), [Kau90]
- Amplitude matched phase-only filter (AMPOF), [Aww90]
- Fractional power filter (FPF), [Vij90a]
- Entropy optimized filter, [Fle90],
- Optimal trade-off filter (OTF), [Réf91a]
- Locally nonlinear matched filter. [Gua93]

and the multiple image designs that will be later commented. Obviously, a detailed discussion of all of these possibilities is out of the scope of this introduction. Moreover, some of these designs introduced only minor modifications and some others are now known to be suboptimal so we will only discuss the principal designs: the matched filter, the phase-only filter, the inverse filter and the generalization of all of them, the optimal trade-off filter.

#### a) Optimization of the signal-to-noise ratio; matched filter.

The signal-to-noise ratio for white noise has been proved to be optimized when the filter is the Fourier transform of the target pattern.



This result can be generalized to colored noise by using again the Cauchy-Schwartz inequality:

$$\text{SNR}_{H(u,v)} = \alpha \frac{\left| \iint H^*(u,v) S(u,v) du dv \right|^2}{\iint P_n(u,v) |H(u,v)|^2 du dv} = \alpha \frac{\left| \iint [H^*(u,v) \sqrt{P_n(u,v)}] \left[ \frac{S(u,v)}{\sqrt{P_n(u,v)}} \right] du dv \right|^2}{\iint P_n(u,v) |H(u,v)|^2 du dv} \leq \quad (2.40)$$

$$\leq \alpha \frac{\iint |H(u,v)|^2 P_n(u,v) du dv \iint \frac{|S(u,v)|^2}{P_n(u,v)} du dv}{\iint |H(u,v)|^2 P_n(u,v) du dv} = \alpha \iint \frac{|S(u,v)|^2}{P_n(u,v)} du dv$$

where the equality holds only when:

$$H(u,v) = \beta \frac{S(u,v)}{P_n(u,v)} \equiv H_{\text{CMF}}(u,v) \quad (2.41)$$

where  $\beta$  is an arbitrary constant.

This filter is known as classical matched filter (CMF) and consists of a whitened version of the Fourier transform of the target image. When the noise is already white, the matched filter reduces to the Fourier transform of the object, except for an unimportant proportionality factor.

The noise robustness of the matched filter has an interesting consequence that it is worth mentioning. When the input scene is affected by noise, the filter receives not the Fourier transform of the target pattern but some distorted version. In such a case, the filter will not be able to cancel out the phase and therefore the maximum value of the correlation plane may be located in an arbitrary position. The deviation between the actual position of a correlation maximum from its nominal location is termed as peak location error and can be proved to be minimized by the matched filter [Vij92b]. Although this seems a logical corollary of the optimality against noise, the proof is far from being obvious and is not given here.

b) Optimization of the optical efficiency: phase-only filter.

The light throughput of an optical correlator can be optimized by using a filter whose transmittance be one at all frequencies, namely:

$$H(u, v) = e^{j\phi(u, v)} \quad (2.42)$$

This filter is called phase-only filter (POF) when  $\phi(u, v)$  is the phase of the Fourier transform of the target. This latter requirement, which is necessary to ensure a reliable detection as we have seen is not an unfavorable restriction but on the contrary an optimal election. In particular the phase-only filter is, among all unit-modulus filters, the optimum in signal-to-noise ratio and peak-to-correlation energy [Vij90b][Dic89b]. The correlation at the origin with a unit-modulus filter of phase  $\phi_H(u, v)$  can be written as:

$$|c(0, 0)| = \left| \iint |S(u, v)| e^{j(\phi_s(u, v) - \phi_H(u, v))} du dv \right| \quad (2.43)$$

which, as previously shown, is a maximum when:

$$\phi_H(u, v) = \phi_S(u, v) \quad (2.44)$$

Since neither the denominator of the signal-to-noise ratio, eq. 2.14, nor that of the peak-to-correlation energy, eq. 2.19, depend on the phase of the filter -they have the same value for all unit-magnitude filters- the statement is proved.

The advantages of phase-only filters over matched filters are numerous. Aside of a higher Horner efficiency that produces a central peak anywhere from 50 to 500 times higher [Hor84][Fla89], the phase-only filter provides an equal or greater peak-to-correlation energy as has been analytically proved by Romero and Dickey in [Rom91] and numerically studied by Kumar et al. in [Vij90b]. For example, the results reported by the latter reference show an improvement of about ten times for the case considered.

Closely related with PCE is the discrimination capability, namely the ease in distinguishing similar patterns, which is consequently higher in the POF than in

the matched. This can be intuitively appreciated if we consider the phase-only filter as a high frequency enhanced version of a matched filter. It is easy to see that the POF is a matched filter that has been divided by its modulus. The modulus of common images concentrates most of the energy at low frequencies -see Fig. 2.3- and thus the denominator attenuates the low frequencies and enhances the high ones. Finally, as the high frequency components of the Fourier transform are related with the edges and the profile of an object, that is with its distinctive elements, the POF becomes more sensitive to differences between images.

A further benefit is a more convenient implementation because only the phase of the incoming light is to be modified. While there are several devices capable of acting as phase-only modulators with more or less accuracy -liquid crystal displays, deformable mirror devices-, the full-complex modulation needed by matched filters presents multiple difficulties -holograms, combination of two modulators in cascade, etc-. In addition, the phase-only filter can be hard-clipped to two values -giving the so-called binary phase-only filters (BPOF) [Hor85]- with a small loss in performance, but with a great saving in storage requirements and with an added simplicity in its implementation. For good surveys of BPOF's and relevant bibliography see [Fla89] and [Réf90b].

On the other hand, the signal-to-noise ratio obtained with phase-only filters is in general small because of the all-pass nature of the filter. Not only the light coming from the images gets through the filter but also the possible noise present at the input scene. A possible remedy [Vij89b] consists in defining a region of support outside which the filter modulus is set to zero. A proper selection of the region of support leads to improvements in the SNR of about 3.5 dB although with a significant loss in peak sharpness.

### c) Optimization of the peak sharpness: inverse filter.

The peak sharpness measured by the peak-to-correlation energy ratio is optimized by the inverse filter (IF), a design probably borrowed from image restoration techniques. Using once again the Cauchy-Schwartz inequality:

$$\begin{aligned}
 \text{PCE}_{H(u,v)} &= \alpha \frac{\left| \iint H^*(u,v) S(u,v) du dv \right|^2}{\iint |H(u,v)|^2 |S(u,v)|^2 du dv} = \alpha \frac{\left| \iint \left[ H^*(u,v) |S(u,v)| \right] \left[ \frac{S(u,v)}{|S(u,v)|} \right] du dv \right|^2}{\iint |H(u,v)|^2 |S(u,v)|^2 du dv} \leq \\
 &\leq \alpha \frac{\iint |H(u,v)|^2 |S(u,v)|^2 du dv \iint du dv}{\iint |H(u,v)|^2 |S(u,v)|^2 du dv} = \alpha \iint du dv = 1
 \end{aligned}
 \tag{2.45}$$

where the equality holds if and only if:

$$H(u,v) = \beta \frac{S(u,v)}{|S(u,v)|^2} \equiv H_{\text{IF}}(u,v)
 \tag{2.46}$$

where  $\beta$  is an arbitrary constant.

The above result indicates that the inverse filter provides a peak-to-correlation energy equal to one, and consequently, that concentrates all the energy passing through the filter in the correlation peak. This can be easily explained in physical terms by looking at the wavefront obtained immediately after the filter:

$$H_{\text{IF}}^*(u,v) \cdot S(u,v) = \beta \frac{S^*(u,v)}{|S(u,v)|^2} \cdot S(u,v) = \beta
 \tag{2.47}$$

namely we obtain a constant or in other words a plane wave. The second lens focuses this plane wave into a single point or equivalently, the second Fourier transform converts the constant function into a delta distribution.

The inverse filter can be considered as a phase-only filter divided by the modulus of the image, and then represents a new step in highlighting the high

frequency components of the matched filter. As a consequence, the discrimination capability of the inverse filter is better than that of the POF although, in this case, this is often a problem rather than a practical advantage; the inverse filter does not only give small correlations with images other than the target but also with slightly distorted versions of the image of interest -the so-called intraclass variations- in such a way that its discrimination properties are frequently presented as a low generalization capability.

The noise sensitivity may also be a difficulty in this design since we are strengthening the frequencies of low energy, and therefore those which are more likely to present a low signal-to-noise ratio. This should be taken with caution because the less the noise distribution differs from the power spectrum of the target, the more the inverse filter approaches the matched filter and thus the better it performs against noise.

Inverse filters present a lower efficiency than the other two designs although this again should be carefully considered before mentioning it as a drawback. For instance in the problem examined in [Réf91a] the inverse filter has a Horner efficiency 9.7 times lower than that of the CMF but the correlation peak is about 56 times more intense than the one given by the matched filter.

Finally, the inverse filter may diverge at some frequencies owing to the mathematical poles of the denominator, a practical problem that can be avoided in different ways as studied in [Car93].

#### d) Putting it all together: the optimal trade-off filter.

The three designs that we have commented so far optimize a different quality criterion. When noise is presumed to be the main problem, one can use a matched filter; if the amount of light at the detector plane is the limiting factor, the phase-only filter would be the adequate choice or finally, if the input scene is composed of multiple targets, the inverse filter with its delta-like peaks could be a good solution. However, all three filters -the phase-only to a lower extent- focus only in one of the quality parameters and, as a consequence, they are overspecialized, that is, they give good performance in one of the ratios of interest but in general, unacceptable low values in the rest.



To overcome the specialization of the filters, we can trade-off, say, peak sharpness for noise robustness or for light efficiency, or for both, to obtain filters that are not optimal in any sense but giving reasonable values for all three criteria. The most effective compromise between SNR, efficiency and PCE is found in the optimal trade-off filters (OT) [Réf91a]. Optimality, according to Réfrégier, means that for given values of two of the criteria, a filter giving a better value for the third one, can not be found with the same degrees of freedom -here the bandwidth of the images-. This definition also implies that it is always possible to find an optimal trade-off that performs better for the three quality parameters than any given filter. They can be obtained by means of the Lagrange minimization method as follows:

Let us call the expression that appears in the denominator of the signal-to-noise ratio in eq. 2.13, the mean-square error (MSE):

$$\text{MSE} = \tilde{\mathbf{h}}^+ \tilde{\mathbf{S}} \tilde{\mathbf{h}} \quad (2.48)$$

and that appearing in the peak-to-correlation energy in eq. 2.20, correlation plane energy (CPE):

$$\text{CPE} = \tilde{\mathbf{h}}^+ \tilde{\mathbf{D}} \tilde{\mathbf{h}} \quad (2.49)$$

These two quantities as well as the Horner efficiency do not depend on the phases of the filter components, and thus the signal-to-noise ratio and the peak-to-correlation energy are optimized with respect to the phases, when the value of the central correlation is the highest. The maximum central correlation is obtained with the phase of the target pattern, a result already obtained in the section concerning phase-only filters. This also matches with our necessity of using the phase of the image to be detected to ensure a reliable identification. Therefore, we need only to derive the magnitudes.

The problem of finding the filter that maximizes the SNR while holding the PCE and the Horner efficiency constant, is equivalent to the problem of minimizing the mean-square error with CPE and the value of the central correlation kept constant. This latter formulation is easier to solve and thus we define the Lagrange function as:



$$E(\alpha, \beta) = \text{MSE} - \alpha \text{CPE} - \beta \tilde{\mathbf{h}} + \tilde{\mathbf{x}} \quad (2.50)$$

where  $\alpha$  and  $\beta$  are the Lagrange multipliers. By introducing the changes:

$$\alpha = \frac{\mu - 1}{\mu} \quad \text{and} \quad \lambda = \frac{1}{2} \mu \beta \quad (2.51)$$

equation 2.50 becomes:

$$E'(\mu, \lambda) = \mu \text{MSE} + (1 - \mu) \text{CPE} - 2 \lambda \tilde{\mathbf{h}} + \tilde{\mathbf{x}} \quad (2.52)$$

where we have the additional constraints:

$$\mu \in [0, 1], \quad \lambda \geq 0, \quad |h_k| \leq 1 \quad (2.53)$$

By making explicit the dependence of the above quantities on the magnitude of the filter components we get:

$$E'(\mu, \lambda) = \mu \sum_{i=1}^N \tilde{S}_{ii} |\tilde{h}_i|^2 + (1 - \mu) \sum_{i=1}^N |\tilde{x}_i|^2 |\tilde{h}_i|^2 - 2 \lambda \sum_{i=1}^N |\tilde{x}_i| |\tilde{h}_i| \quad (2.54)$$

and setting the derivatives of  $E'(\mu, \lambda)$  with respect to  $|\tilde{h}_k|$  to zero we obtain:

$$2 \mu \tilde{S}_{kk} |\tilde{h}_k| + 2(1 - \mu) |\tilde{x}_k|^2 |\tilde{h}_k| - 2 \lambda |\tilde{x}_k| = 0 \quad (2.55)$$

whence

$$|\tilde{h}_k| \left( \mu \tilde{S}_{kk} + (1 - \mu) |\tilde{x}_k|^2 \right) = \lambda |\tilde{x}_k| \quad (2.56)$$

and considering the restrictions of eq. 2.53:

$$\left. \begin{aligned} |\tilde{h}_k| = 1 & \quad \text{iif} \quad \frac{\lambda |\tilde{x}_k|}{\mu \tilde{S}_{kk} + (1-\mu) |\tilde{x}_k|^2} \geq 1 \\ |\tilde{h}_k| = \frac{\lambda |\tilde{x}_k|}{\mu \tilde{S}_{kk} + (1-\mu) |\tilde{x}_k|^2} & \quad \text{otherwise} \end{aligned} \right\} \quad (2.57)$$

which can be compactly written as:

$$|\tilde{h}_k| = \delta_k + (1-\delta_k) \frac{\lambda |\tilde{x}_k|}{\mu \tilde{S}_{kk} + (1-\mu) |\tilde{x}_k|^2} \quad (2.58)$$

where  $\delta_k$  is zero or one depending on whether or not the condition in eq. 2.57 is satisfied. Finally the optimal trade-off filter can be expressed as:

$$\tilde{h}_k = \sigma_\lambda \left[ \frac{\tilde{x}_k}{\mu \tilde{S}_{kk} + (1-\mu) |\tilde{x}_k|^2} \right] \quad (2.59)$$

where

$$\sigma_\lambda(y) = \begin{cases} \lambda y & \text{if } |y| \leq \frac{1}{\lambda} \\ e^{j\psi(y)} & \text{otherwise, where } \psi(y) \text{ is the phase of } y \end{cases} \quad (2.60)$$

The optimal trade-off filter includes the matched, the phase-only and the inverse filters. In particular, by setting:

$$\mu = 1, \quad \lambda < \frac{1}{\text{Max} \left( \frac{|\tilde{x}_k|}{\tilde{S}_{kk}} \right)} \quad (2.61)$$

eq. 2.59 becomes:

$$\tilde{h}_k = \sigma_\lambda \left[ \frac{\tilde{x}_k}{\tilde{S}_{kk}} \right] = \frac{\lambda \tilde{x}_k}{\tilde{S}_{kk}} \quad \text{since } |y| \leq \frac{1}{\lambda} \quad (2.62)$$

that is the matched filter; by setting:

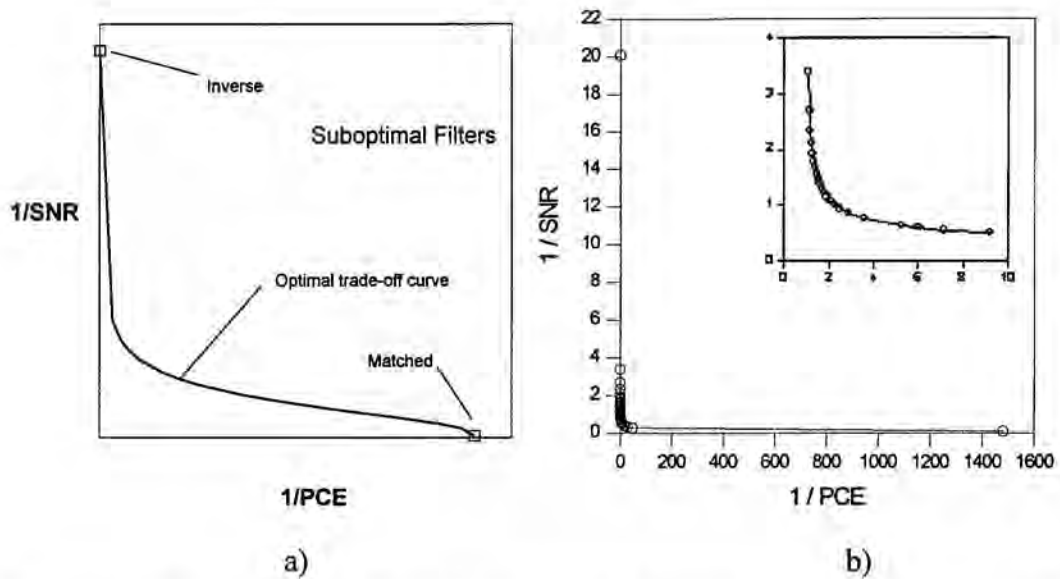
$$\mu = 0, \quad \lambda < \frac{1}{\text{Max} \left( \frac{1}{|\tilde{x}_k|} \right)} \quad (2.63)$$

we obtain:

$$\tilde{h}_k = \sigma_\lambda \left[ \frac{\tilde{x}_k}{|\tilde{x}_k|^2} \right] = \frac{\lambda \tilde{x}_k}{|\tilde{x}_k|^2} \quad \text{since } |y| \leq \frac{1}{\lambda} \quad (2.64)$$

that is the inverse filter, and finally by setting:

$$\lambda = \infty \quad (2.65)$$



**Figure 2.6-** Operating curve of an OT filter. a) Key elements. b) Trade-of between SNR and PCE. The inset curve is a zoom of the elbow area.

we obtain:

$$\tilde{h}_k = \exp j\psi \left( \frac{\tilde{x}_k}{\mu \tilde{S}_{kk} + (1-\mu) |\tilde{x}_k|^2} \right) \quad (2.66)$$

and since the phase of the filter is that of the target, we obtain the phase-only filter.

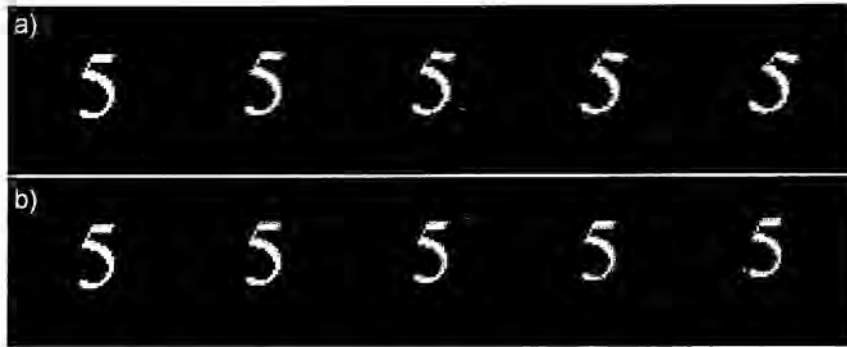
Figure 2.6.a) shows the transition between an inverse filter and a matched filter by means of the optimal trade-off technique -only PCE and SNR are being considered-. The suboptimal filters lie in the region above the curve since for a given value of the SNR they present a lower PCE. The region below the operating curve can not be reached with a filter with  $N$  degrees of freedom. The trade-off approach permits the obtention of filters with good performance in all the quality criteria as can be observed in Figure 2.6.b). The plot represents the operating curve of an optimal trade-off filter obtained to detect the first butterfly of Figure 2.2. The noise is white with a power spectral density of 0.1 of the average value of

that of the target butterfly. This curve is very close to the axes and then it is possible to get an important increasing in the signal-to-noise ratio of an inverse filter with a small loss in the peak-to-correlation energy or vice versa, an important increasing in the peak-to-correlation energy of a matched filter with a small loss in signal-to-noise ratio. As an example, the trade-off filter obtained with  $\mu=0.04$  has both a PCE and a SNR of about 17% of the maximum values. The overspecialization of the matched and the inverse filter is evident in this example: the signal-to-noise ratio of the inverse filter is only 0.5% that of the matched and the peak-to-correlation energy of the matched filter only 0.07% that of the inverse.

In conclusion, the possibility of using an arbitrary function for the modulus enables us to design filters that are optimum with respect to different criteria. We can also combine these different criteria to obtain more flexible filters by putting more emphasis in the more critical parameters but without forgetting the rest. In this way, the filters can be adapted to the application instead of adapting the application to the filter.

## 2.4 Lack of invariance of single-image filters.

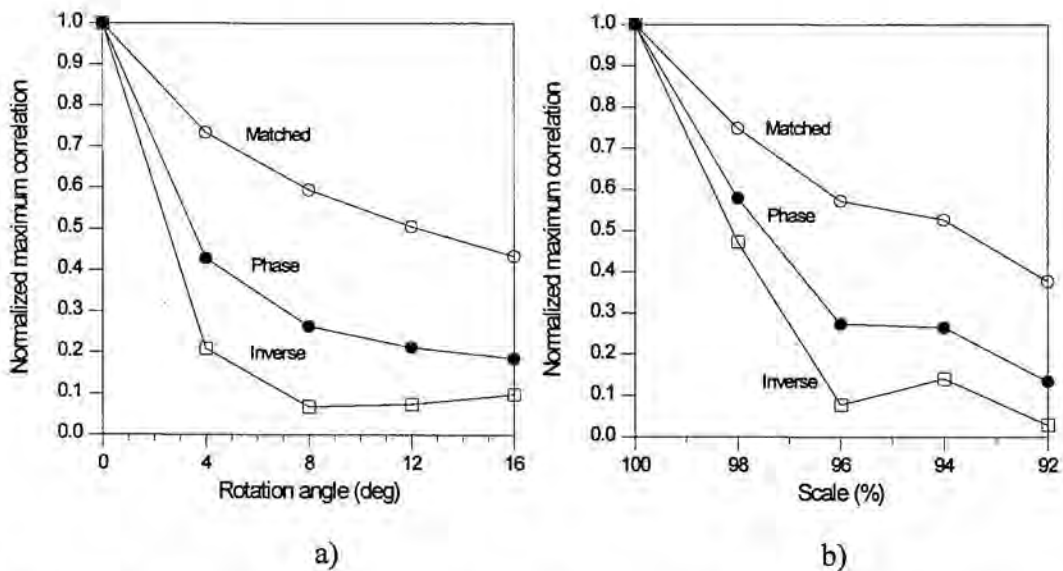
Up to now, we have been assuming that the only degradation that can affect the input images is noise, a sample realization of an stochastic process. We mentioned somewhere that there has been some attempts to modelize the background clutter in a more realistic way than the one provided by the additive-Gaussian-stationary model. Degradations such as defocussing, that can also be present in the scene, are currently being analyzed and some methods to minimize their effects being developed [Cam91][Sal95]. However, the probably more important source of distortions is that affecting the geometry of the images. In practice it is unlikely to observe the object always under the same point of view, at the same distance, and as a consequence, the images may be captured with a variable scale or arbitrarily rotated. A pattern recognition system intended to be used in a practical situation, must, therefore, be able to identify the objects of interest, regardless of their relative positions with respect to the acquisition



**Figure 2.7-** Test images. a) Original and images rotated each 4 degrees. b) Original and images reduced in 2% steps.

module. Such a property is called invariance and unfortunately is absent from the filters considered so far.

Figure 2.7 and 2.8 are an example of the lack of invariance of the matched, phase-only and inverse filters. Figure 2.7.a) shows five images corresponding to  $0^\circ$ ,  $4^\circ$ ,  $8^\circ$ ,  $12^\circ$ , and  $16^\circ$  in-plane rotations. Similarly, Fig. 2.7.b) shows an image at five different scales: 100%, 98%, 96%, 94%, and 92%. Figure 2.8 represents the correlation -the maximum normalized to the central correlation of the original pattern- between images in Fig. 2.7 and the three filters derived from the



**Figure 2.8-** Correlation between single-image filters and a) rotated input b) scaled input.



nondistorted pattern. All three present a marked decay, that in the best case, is about 50% of the original correlation. The excessive discrimination capability of the inverse filter is clearly observed, making it useless to recognize the target even with these small changes in scale and view angle.

There are two approaches that can be followed to overcome the lack of invariance of single-image filters. Both, explicitly or implicitly, involve the multiplexing of several images into the filter as a means to introduce the information about how the distortion affects the original pattern. This implies that the phase of the filter components are not matched to the target image -in fact we have several targets, the different views of the object- and thus, according to the conditions we derive for a reliable detection, the central correlation is not necessarily the maximum of the output plane. These new filters may present large lateral peaks, known as sidelobes, that can induce to false alarms and to an incorrect determination of the position of the image. Several solutions to the problem of the presence of large sidelobes have been proposed and will be discussed in detail in this and in subsequent chapters.

The two possibilities to obtain invariant correlation filters are the ad hoc approach, based on harmonic series, and the general formalism provided by the synthetic discriminant functions (SDF) theory. The first approach is based on the fact that sometimes the target pattern can be expressed as an infinite sum of two-dimensional functions, each one being invariant to the deformation. By using as a filter one of the terms of such a decomposition, we will obtain the same correlation with the original pattern as with any distorted version. Moreover, as the filter conserves certain amount of information about the image it was derived from, one expects to obtain a large value for the central correlation and some discrimination capability. Examples of this technique are the circular harmonic filters (CHF) [Hsu82a][Hsu82b] to handle in-plane rotations and the radial Mellin harmonic filters (RHF) [Cas76] to address scale changes.

The second approach, the SDF theory, is closer to the supervised learning procedures found in neural networks. This technique is quite general and does not need an analytical description of the degradation but only to have abundant examples of distorted images.

## 2.5 Design of invariant filters.

The two approaches to design filters capable of handling large deformations of the input patterns are the topic of this section. The harmonic decomposition method is illustrated by means of an specific example: the circular harmonic expansion.

### 2.5.1 Circular harmonic filters.

The in-plane rotation is one of the cases where the ad hoc approach can be used. The method is based on the expansion of an image into its circular harmonic components and has been adapted to designing correlation filters by Hsu *et al.* [Hsu82a][Hsu82b]. For the sake of simplicity, the following discussion is carried out in direct rather than in Fourier space.

Let us consider the image  $f(x,y)$  expressed in polar coordinates. It will be denoted  $f(r, \theta)$ . In this representation, the image is a periodic function with respect to the polar angle since:

$$f(r, \theta) = f(r, \theta + 2\pi) \quad (2.67)$$

namely a function with period  $2\pi$ , and therefore  $f(r, \theta)$ , for any value of the radius, admits a Fourier expansion, expressed as:

$$f(r, \theta) = \sum_{m=-\infty}^{\infty} f_m(r) e^{jm\theta} \quad (2.68)$$

where  $m$  are integer numbers and the factors  $f_m(r)$  are called circular harmonics - strictly speaking the circular harmonics are the complex exponential terms-.

The coefficients  $f_m(r)$  can be found in the following way:

$$f(r, \theta) e^{-jn\theta} = \sum_{m=-\infty}^{\infty} f_m(r) e^{j(m-n)\theta} \quad (2.69)$$

and integrating both sides with respect to the polar angle we get:

$$\int_0^{2\pi} f(r, \theta) e^{-jn\theta} d\theta = \sum_{m=-\infty}^{\infty} f_m(r) \int_0^{2\pi} e^{j(m-n)\theta} d\theta = \sum_{m=-\infty}^{\infty} f_m(r) 2\pi \delta_{m-n} = 2\pi f_n(r) \quad (2.70)$$

whence:

$$f_n(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) e^{-jn\theta} d\theta \quad (2.71)$$

If we use as a filter one of the infinite terms of the sum in equation 2.68, namely:

$$h(r, \theta) = f_n(r) e^{jn\theta} \quad (2.72)$$

the modulus of the central correlation with the target image will be the same regardless of the rotation suffered by the object. The correlation in polar coordinates between the image rotated a certain angle  $\alpha$ ,  $f(r, \theta - \alpha)$ , and a circular harmonic filter can be expressed as:

$$[f^\alpha * h]_{(0,0)} = \int_0^{\infty} \int_0^{2\pi} h^*(r, \theta) f(r, \theta - \alpha) r dr d\theta \quad (2.73)$$

which, by using eqs. 2.68 and 2.72, becomes:

$$[f^\alpha \star h]_{(0,0)} = \int_0^\infty \int_0^{2\pi} f_n^*(r) e^{-jn\theta} \left( \sum_{m=-\infty}^{\infty} f_m(r) e^{jm(\theta-\alpha)} \right) r dr d\theta \quad (2.74)$$

that can be rearranged to obtain:

$$\begin{aligned} [f^\alpha \star h]_{(0,0)} &= \int_0^\infty f_n^*(r) \left( \sum_{m=-\infty}^{\infty} f_m(r) e^{-jm\alpha} \int_0^{2\pi} e^{j(m-n)\theta} d\theta \right) r dr = \\ &= e^{-jn\alpha} \int_0^\infty f_n^*(r) \left( \sum_{m=-\infty}^{\infty} f_m(r) 2\pi \delta_{m-n} \right) r dr = 2\pi e^{-jn\alpha} \int_0^\infty f_n^*(r) f_n(r) r dr = \\ &= e^{-jn\alpha} [f^0 \star h]_{(0,0)} \end{aligned} \quad (2.75)$$

and as we detect energies, namely the squared modulus of the later expressions, we finally have:

$$\left| [f^\alpha \star h]_{(0,0)} \right|^2 = \left| e^{-jn\alpha} [f^0 \star h]_{(0,0)} \right|^2 = \left| [f^0 \star h]_{(0,0)} \right|^2 \quad (2.76)$$

which is the desired invariance for the central correlation.

The analysis of the zero-order circular harmonic could be enlightening to understand how these filters provide rotation invariance. The zero-order filter is merely, see eqs. 2.71 and 2.72:

$$h(r, \theta) = f_0(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) d\theta \quad (2.77)$$

This function does not depend on the polar angle -it has the same value whatever the angle  $\theta$ - and thus presents circular symmetry. As a consequence, all images rotated with respect to this symmetry center will see the same filter and will give the same correlation.

It is worth mentioning that a circular harmonic filter can be interpreted as being formed by a linear combination of all possible rotated images. For the particular case of the zero-order harmonic the coefficients of the linear combination are all equal and thus it becomes the average of those images. This can be easily shown. The average of all possible rotated images can be written as:

$$\bar{f}(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta + \alpha) d\alpha \quad (2.78)$$

By splitting the interval of integration into two parts we obtain:

$$\bar{f}(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi-\theta} f(r, \theta + \alpha) d\alpha + \frac{1}{2\pi} \int_{2\pi-\theta}^{2\pi} f(r, \theta + \alpha) d\alpha \quad (2.79)$$

that can be written as:

$$\bar{f}(r, \theta) = \frac{1}{2\pi} \int_{\theta}^{2\pi} f(r, \beta) d\beta + \frac{1}{2\pi} \int_0^{\theta} f(r, \beta) d\beta \quad (2.80)$$

if the following changes are performed:

$$\begin{aligned} 1^{\text{st}} \text{ integral} &\rightarrow \beta = \theta + \alpha \\ 2^{\text{nd}} \text{ integral} &\rightarrow \beta = \theta + \alpha - 2\pi \end{aligned} \quad (2.81)$$

and so finally we have:

$$\bar{f}(r, \theta) = \bar{f}(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \beta) d\beta \equiv f_0(r) \quad (2.82)$$

namely the zero-order harmonic. This result will be later used and justifies the statement that the harmonic based filters implicitly involve the multiplexing of samples of distorted images to achieve their invariance properties.

### 2.5.2 The general approach: synthetic discriminant functions (SDF's).

The former method needs an analytical description of the degradation, which is not always available and thus can only be applied to a few specific cases. In general, the relationship between two distorted images is unknown although, in contrast, we can obtain as many examples of such degraded images as we wish. The synthetic discriminant functions theory allows the design of invariant filters based on the information provided by these sample patterns and therefore, with no previous knowledge about the degradation affecting them. The set of patterns used to compute the filter is called the training set and, if it is representative enough, the resulting solution captures the underlying connection between the images, in such a way that we obtain the desired correlation even with patterns not used in the learning process. *Representative enough* is, as in other learning-by-examples procedures not well defined, being the size of the training set required to obtain generalization capabilities dependent on the type of filter. These designs does not guarantee as in the harmonic expansion approach, that the maximum correlation will be found pointing to the position occupied by the target. This means that we must provide solutions to the problem of the appearance of sidelobes.



The filters we are looking for can be characterized in mathematical terms in the following way:

Let  $\mathbf{h}$  be the filter whose expression is to be found and let  $\mathbf{x}_i$  be the  $i$ th training image. Let  $c_i$  the value we wish to obtain as central correlation with the pattern  $\mathbf{x}_i$ , namely:

$$\mathbf{h}^+ \mathbf{x}_i = c_i, \quad i = 1, \dots, K \quad (2.83)$$

where  $K$  is the number of training patterns. Notice that the correlation can be expressed in the above form in both, object and Fourier space, eqs. 2.2 and 2.3, so the following discussion is equally valid whatever the representation we use. This equation can be written in a more compact form by defining an  $N \times K$  matrix,  $\mathbf{X}$ , whose  $i$ th column be the  $i$ th training vector  $\mathbf{x}_i$ . In this way equation 2.83 becomes:

$$\mathbf{h}^+ \mathbf{X} = \mathbf{c}^t \quad (2.84)$$

where  $\mathbf{c}$  is the column vector of components  $c_i$ . This is a system of linear equations for which the superposition principle holds:

$$\mathbf{h}_G = \mathbf{h}_P + \mathbf{h}_H \quad (2.85)$$

namely, that the general solution of 2.84,  $\mathbf{h}_G$ , can be written as a particular solution,  $\mathbf{h}_P$ , plus the general solution,  $\mathbf{h}_H$ , to the homogeneous system:

$$\mathbf{h}_H^+ \mathbf{X} = 0 \quad (2.86)$$

In fact, if  $\mathbf{h}$  is a solution of the system it can be decomposed as:

$$\mathbf{h} = \mathbf{h}_P + (\mathbf{h} - \mathbf{h}_P) \quad (2.87)$$

where

$$(\mathbf{h} - \mathbf{h}_p)^+ \mathbf{X} = \mathbf{h}^+ \mathbf{X} - \mathbf{h}_p^+ \mathbf{X} = \mathbf{c} - \mathbf{c} = \mathbf{0} \quad (2.88)$$

The particular solution is found by means of an *ansatz*: a tentative solution that in our case is formed by a linear combination of the training images:

$$\mathbf{h}_p = \sum_{i=1}^K a_i \mathbf{x}_i = \mathbf{X} \mathbf{a} \quad (2.89)$$

where  $a_i$  are the coefficients of the combination and  $\mathbf{a}$  the column vector whose components are  $a_i$ . By substituting eq. 2.89 into eq. 2.84 we get:

$$\mathbf{a}^+ \mathbf{X}^+ \mathbf{X} = \mathbf{c}^t \quad (2.90)$$

The  $K \times K$  matrix  $(\mathbf{X}^+ \mathbf{X})$  is regular when the training patterns are linearly independent, something that can be supposed with wide generality and a requirement of most pattern recognition techniques. The solution for the values  $a_i$  can then be found as:

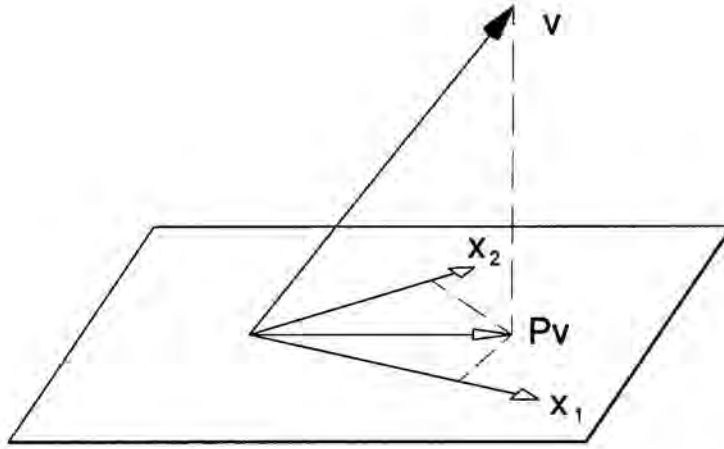
$$\mathbf{a}^+ = \mathbf{c}^t (\mathbf{X}^+ \mathbf{X})^{-1} \Leftrightarrow \mathbf{a} = (\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* \quad (2.91)$$

and finally the filter  $\mathbf{h}_p$  is:

$$\mathbf{h}_p = \mathbf{X} \mathbf{a} = \mathbf{X} (\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* \quad (2.92)$$

A few comments about this expression are in order. First, it is instructive to check that it is actually a solution of equation 2.84:

$$\mathbf{h}_p^+ \mathbf{X} = \mathbf{c}^t (\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \mathbf{X} = \mathbf{c}^t \quad (2.93)$$



**Figure 2.9-** Orthogonal projection of the vector  $v$  onto the subspace generated by the training patterns.

On the other hand by substituting this later equality into the filter expression, equation 2.92, we obtain:

$$\mathbf{h}_P = \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \mathbf{h}_P \equiv \mathbf{P} \mathbf{h}_P \quad (2.94)$$

The  $N \times N$  matrix  $\mathbf{P}$  is an orthogonal projector, that is, an operator that projects any vector onto the hyperplane spanned by the  $K$  training images in an orthogonal fashion as shown in Figure 2.9. Its main properties are:

a)  $\mathbf{P}$  is a Hermitian matrix:

$$\begin{aligned} \mathbf{P}^+ &= \left( \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right)^+ = \mathbf{X}^{++} \left( \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \right)^+ = \mathbf{X} \left( \left( (\mathbf{X}^+ \mathbf{X})^{-1} \right)^+ \mathbf{X}^+ \right) = \\ &= \mathbf{X} \left( \left( (\mathbf{X}^+ \mathbf{X})^+ \right)^{-1} \mathbf{X}^+ \right) = \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ = \mathbf{P} \end{aligned} \quad (2.95)$$

b)  $\mathbf{P}$  is idempotent, i.e.  $\mathbf{P}^2 = \mathbf{P}$  (and thus  $\mathbf{P}^n = \dots = \mathbf{P}^3 = \mathbf{P}^2 = \mathbf{P}$ ):

$$\begin{aligned}
 \mathbf{P}^2 &= \left( \mathbf{X}(\mathbf{X}^+\mathbf{X})^{-1}\mathbf{X}^+ \right) \left( \mathbf{X}(\mathbf{X}^+\mathbf{X})^{-1}\mathbf{X}^+ \right) = \mathbf{X}(\mathbf{X}^+\mathbf{X})^{-1} \underbrace{\mathbf{X}^+\mathbf{X}(\mathbf{X}^+\mathbf{X})^{-1}}_{\text{Id}} \mathbf{X}^+ = \\
 &= \mathbf{X}(\mathbf{X}^+\mathbf{X})^{-1}\mathbf{X}^+ = \mathbf{P}
 \end{aligned}
 \tag{2.96}$$

which means that once the vector has been projected and belongs to the hyperplane, additional applications of the operator  $\mathbf{P}$  do not result in any further modification. The projection of a vector already belonging to the subspace is the vector itself, being equation 2.94 an example.

The effect of the projection matrix over a vector,  $\mathbf{v}$ , can be easily visualized when the image vectors  $\mathbf{x}_i$  are orthogonal. In this case we have:

$$\mathbf{x}_i^+ \mathbf{x}_j = \|\mathbf{x}_i\|^2 \delta_{ij} \Leftrightarrow (\mathbf{X}^+\mathbf{X})^{-1} = \begin{pmatrix} \frac{1}{\|\mathbf{x}_1\|^2} & & \\ & \ddots & \\ & & \frac{1}{\|\mathbf{x}_k\|^2} \end{pmatrix}
 \tag{2.97}$$

and multiplying it by the matrix  $\mathbf{X}$  we get:

$$\mathbf{X}(\mathbf{X}^+\mathbf{X})^{-1} = \left( \left( \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|^2} \right) \cdot \left( \frac{\mathbf{x}_k}{\|\mathbf{x}_k\|^2} \right) \right)
 \tag{2.98}$$

On the other hand, the product  $\mathbf{X}^+ \mathbf{v}$  is a  $K$ -dimensional column vector whose components are the scalar products:

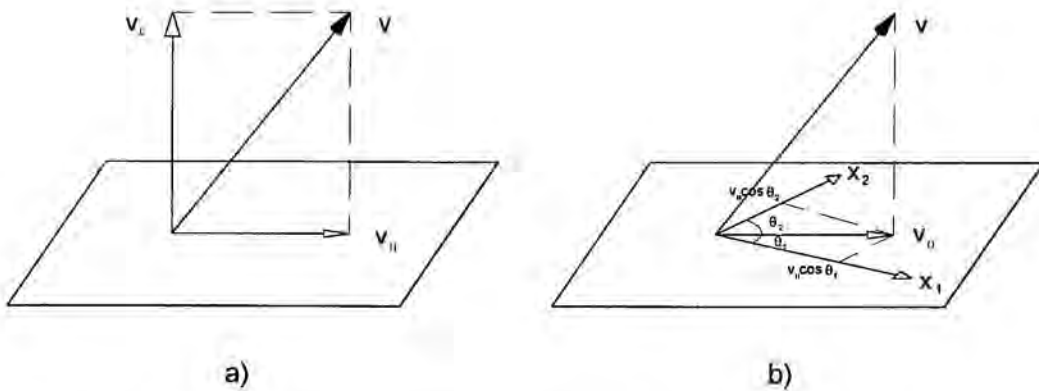
$$\mathbf{X}^+ \mathbf{v} = \begin{pmatrix} \mathbf{x}_1^+ \mathbf{v} \\ \vdots \\ \mathbf{x}_k^+ \mathbf{v} \end{pmatrix} \quad (2.99)$$

so by combining equations 2.98 and 2.99, we can express the projection of the vector  $\mathbf{v}$  as:

$$\mathbf{Pv} = \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} \frac{\mathbf{x}_1^+ \mathbf{v}}{\|\mathbf{x}_1\|} + \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} \frac{\mathbf{x}_2^+ \mathbf{v}}{\|\mathbf{x}_2\|} + \dots + \frac{\mathbf{x}_k}{\|\mathbf{x}_k\|} \frac{\mathbf{x}_k^+ \mathbf{v}}{\|\mathbf{x}_k\|} \quad (2.100)$$

and since the vector  $\mathbf{v}$  can be written as the sum of a parallel -with respect to the set of images- and a perpendicular component, see Figure 2.10.a), the scalar products in the above equation become:

$$\mathbf{x}_i^+ \mathbf{v} = \mathbf{x}_i^+ (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) = \mathbf{x}_i^+ \mathbf{v}_{\parallel} + 0 = \|\mathbf{x}_i\| \|\mathbf{v}_{\parallel}\| \cos \theta_i \quad (2.101)$$



**Figure 2.10-** a) Decomposition of a vector b) Perpendicular projection.

where  $\mathbf{v}_{\parallel}$  is the projected vector and  $\theta_i$  the angles formed by  $\mathbf{v}_{\parallel}$  and the training images as shown in Figure 2.10.b). Finally, we have:

$$\mathbf{P}\mathbf{v} = \mathbf{e}_1 \|\mathbf{v}_{\parallel}\| \cos\theta_1 + \mathbf{e}_2 \|\mathbf{v}_{\parallel}\| \cos\theta_2 + \dots + \mathbf{e}_k \|\mathbf{v}_{\parallel}\| \cos\theta_k = \mathbf{v}_{\parallel} \quad (2.102)$$

that clearly shows how the multiplication by the matrix  $\mathbf{P}$  gives the component of the vector lying on the hyperplane spanned by the images of the training set.

This formalism is very appropriate to characterize orthogonal vectors and therefore allows us to find the homogeneous solution with ease. Let  $\mathbf{v}$  be an arbitrary vector. According to the above discussion it can be written as:

$$\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} = \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} + \mathbf{v}_{\perp} \quad (2.103)$$

whence

$$\mathbf{v}_{\perp} = \mathbf{v} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} = \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} = [\mathbf{I} - \mathbf{P}] \mathbf{v} \equiv \mathbf{P}' \mathbf{v} \quad (2.104)$$

where  $\mathbf{I}$  is the  $N \times N$  identity matrix. The matrix  $\mathbf{P}'$  has exactly the same properties as  $\mathbf{P}$ , namely it is Hermitian and idempotent, because it also represents an orthogonal projection, now over the subspace generated by the orthogonal vectors to the training images. As a consequence, the kernel of the matrix  $\mathbf{X}$  can be readily expressed as:

$$\text{Ker } \mathbf{X} \equiv \left\langle \mathbf{v} \mid \mathbf{X}^+ \mathbf{v} = \mathbf{v}^+ \mathbf{X} = \mathbf{0} \right\rangle = \left\{ \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \mid \forall \mathbf{v} \in \mathbf{C}^N \right\} \quad (2.105)$$

and thus the homogeneous solution is:

$$\mathbf{h}_H = \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \quad (2.106)$$



where  $\mathbf{v}$  is an arbitrary  $N$ -dimensional complex vector. Indeed, we have:

$$\mathbf{h}_H^+ \mathbf{X} = \mathbf{v}^+ \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{X} = \mathbf{v}^+ \mathbf{X} - \mathbf{v}^+ \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \mathbf{X} = \mathbf{v}^+ \mathbf{X} - \mathbf{v}^+ \mathbf{X} = \mathbf{0} \quad (2.107)$$

and thus finally, the general solution of 2.84 is:

$$\mathbf{h}_G = \mathbf{h}_P + \mathbf{h}_H = \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* + \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \quad (2.108)$$

This design procedure enables us not only to design invariant filters, but also filters with controlled response -the values for  $c_i$  can be chosen at one's wishes- to be used in classification systems or to include information about images that we need not to confuse with the targets. Thus, the SDF formalism makes correlation a flexible tool for pattern recognition and image classification. Furthermore, equation 2.108 implies that we have available an infinite number of filters fulfilling the conditions we imposed on the correlation values -an  $(N-K)$ -dimensional hyperplane for real or a  $(2N-K)$ -dimensional one if the filters are allowed to be complex in object space- from which we can select the most suitable solution for a given problem. In particular, we could search for optimal filters with respect to the quality criteria exposed in section 2.1. Optimal filters are usually derived using Lagrange optimization procedures but we will continue with the geometrical point of view we have been using, which offers more intuitive insights.

The first interesting result in this sense is that the linear combination filter,  $\mathbf{h}_P$ , is the minimum-norm solution to equation 2.84:

$$\begin{aligned} \mathbf{h}^+ \mathbf{h} = \|\mathbf{h}\|^2 &= \left\| \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* + \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \right\|^2 = \left\| \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* \right\|^2 + \\ &+ \left\| \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \right\|^2 \geq \left\| \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* \right\|^2 = \|\mathbf{h}_P\|^2 = \mathbf{h}_P^+ \mathbf{h}_P \end{aligned} \quad (2.109)$$

whatever the filter  $\mathbf{h}$ .

The filter  $h_p$ , called projection SDF or composite filter is important because of historical reasons. It was proposed by Hester and Casasent [Hes80] in 1980 being the first filter of the SDF type. Since it is a weighted sum of several images, it can be implemented with conventional optical holography through a multiexposure VanderLugt filter. However, with the advent and development of digital holography, the optical feasibility is no longer a constraint and the linear combination assumption is an unnecessary restriction. Let us find more interesting designs.

All the correlation quality criteria we have established in section 2.1, involve purely quadratic terms. For these simple expressions it is possible to derive the optimal filters exploiting a result already obtained. To optimize the different figures of merit -Horner efficiency is an exception- and because of the central correlations are now fixed, we need only to minimize their respective denominators. In mathematical terms we look for the filter that makes

$$E = \mathbf{h}^+ \mathbf{M} \mathbf{h} \quad (2.110)$$

minimum while fulfilling

$$\mathbf{h}^+ \mathbf{X} = \mathbf{c}^t \quad (2.111)$$

and where  $\mathbf{M}$  is a real symmetrical matrix. Equation 2.110 may also be written as:

$$E = \mathbf{h}^+ \mathbf{M} \mathbf{h} = \mathbf{h}^+ \mathbf{M}^{1/2} \mathbf{M}^{1/2} \mathbf{h} = (\mathbf{M}^{1/2} \mathbf{h})^+ (\mathbf{M}^{1/2} \mathbf{h}) \equiv \mathbf{h}'^+ \mathbf{h}' \quad (2.112)$$

and similarly equation 2.111:

$$\mathbf{h}^+ \mathbf{X} = \mathbf{c}^t \Leftrightarrow \mathbf{h}^+ \mathbf{M}^{1/2} \mathbf{M}^{-1/2} \mathbf{X} = \mathbf{c}^t \Leftrightarrow (\mathbf{M}^{1/2} \mathbf{h})^+ (\mathbf{M}^{-1/2} \mathbf{X}) = \mathbf{c}^t \Leftrightarrow \mathbf{h}'^+ \mathbf{X}' = \mathbf{c}^t \quad (2.113)$$

Thus we want a filter so that:

$$\begin{aligned} \mathbf{h}'^+ \mathbf{X}' &= \mathbf{c}^t \\ \mathbf{h}'^+ \mathbf{h}' &\text{ minimum} \end{aligned} \quad (2.114)$$

The solution of the problem in terms of  $\mathbf{h}'$  and  $\mathbf{X}'$  is known to be the composite filter, namely:

$$\mathbf{h}' = \mathbf{X}' (\mathbf{X}'^+ \mathbf{X}')^{-1} \mathbf{c}^* \quad (2.115)$$

and if we express the above equation as a function of the original quantities, we finally have:

$$\begin{aligned} \mathbf{M}^{1/2} \mathbf{h} = \mathbf{h}' &= \mathbf{X}' (\mathbf{X}'^+ \mathbf{X}')^{-1} \mathbf{c}^* = \mathbf{M}^{-1/2} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1/2} \mathbf{M}^{-1/2} \mathbf{X})^{-1} \mathbf{c}^* = \\ &= \mathbf{M}^{-1/2} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* \Leftrightarrow \mathbf{h} = \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* \end{aligned} \quad (2.116)$$

However, a different approach to solve the same problem provides a richer interpretation and a description both unified and elegant. By rewriting the equation about the central correlations as:

$$\mathbf{h}^+ \mathbf{X} = \mathbf{c}^t \Leftrightarrow \mathbf{h}^+ \mathbf{M} (\mathbf{M}^{-1} \mathbf{X}) = \mathbf{c}^t \Leftrightarrow \mathbf{h}^+ \mathbf{M} \mathbf{X}' = \mathbf{c}^t \quad (2.117)$$

where the new definitions are obvious, our problem turns out to be:

$$\begin{aligned} \mathbf{h}^+ \mathbf{M} \mathbf{X}' &= \mathbf{c}^t \\ \mathbf{h}^+ \mathbf{M} \mathbf{h} &\text{ minimum} \end{aligned} \quad (2.118)$$

Notice that equation 2.84 can be considered as a particular case of 2.117 for  $M=I$ . Moreover, in our case the matrices  $M$  are not only real and symmetrical but also positive definite, that is:

$$\mathbf{v}^+ \mathbf{M} \mathbf{v} \geq 0 \quad \forall \mathbf{v} \quad (2.119)$$

and then they can be considered as a metric. It should be taken into account in what follows that a new metric implies a new definition of the orthogonality between vectors:

$$\mathbf{u} \perp \mathbf{v} \Leftrightarrow \mathbf{u}^+ \mathbf{M} \mathbf{v} = 0 \quad (2.120)$$

and also a new definition of the norm of a vector:

$$\|\mathbf{u}\|_{\mathbf{M}}^2 \equiv \mathbf{u}^+ \mathbf{M} \mathbf{u} \quad (2.121)$$

With these elements, the problem can be reinterpreted as the search of the minimum norm vector whose projections onto the training images are known, being the norm and the projections defined through a generalized metric  $M$ . The general solution to:

$$\mathbf{h}^+ \mathbf{M} \mathbf{X}' = \mathbf{c}^t \quad (2.122)$$

is completely analogous to that in eq. 2.84 and is obtained through similar deductions. As before, the particular solution is formed by a linear combination of the training images:

$$\mathbf{h}'_{\mathbf{p}} = \mathbf{X}' \mathbf{a} \quad (2.123)$$

and substituting into eq. 2.117:

$$\mathbf{a}^+ \mathbf{X}'^+ \mathbf{M} \mathbf{X}' = \mathbf{c}^t \Leftrightarrow \mathbf{a}^+ = \mathbf{c}^t \left( \mathbf{X}'^+ \mathbf{M} \mathbf{X}' \right)^{-1} \Leftrightarrow \mathbf{a} = \left( \mathbf{X}'^+ \mathbf{M} \mathbf{X}' \right)^{-1} \mathbf{c}^* \quad (2.124)$$

and therefore the filter  $\mathbf{h}'_{\mathbf{P}}$  is:

$$\mathbf{h}'_{\mathbf{P}} = \mathbf{X}'\mathbf{a} = \mathbf{X}'(\mathbf{X}' + \mathbf{M}\mathbf{X}')^{-1} \mathbf{c}^* = \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{M}\mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{c}^* = \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{c}^* \quad (2.125)$$

By looking at this later expression we define the orthogonal projector - according to the new metric- onto the hyperplane generated by the the modified images  $\mathbf{X}'$  as:

$$\mathbf{P} = \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \quad (2.126)$$

and the orthogonal projector onto the orthogonal subspace:

$$\mathbf{P}' = [\mathbf{I} - \mathbf{P}] = \left[ \mathbf{I} - \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \right] \quad (2.127)$$

with which the general solution can be written as:

$$\mathbf{h}'_{\mathbf{G}} = \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{c}^* + \left[ \mathbf{I} - \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} = \mathbf{h}'_{\mathbf{P}} + \mathbf{h}'_{\mathbf{H}} \quad (2.128)$$

It should be noted that the two components of  $\mathbf{h}'_{\mathbf{G}}$  are perpendicular only according to the metric defined by the matrix  $\mathbf{M}$ :

$$\begin{aligned} \mathbf{h}'_{\mathbf{P}}{}^+ \mathbf{M} \mathbf{h}'_{\mathbf{H}} &= \mathbf{c}^+ (\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \mathbf{M}^{-1} \mathbf{M} \left[ \mathbf{I} - \mathbf{M}^{-1}\mathbf{X}(\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} = \\ &= \mathbf{c}^+ (\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} - \mathbf{c}^+ (\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X} (\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} = \\ &= \mathbf{c}^+ (\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} - \mathbf{c}^+ (\mathbf{X} + \mathbf{M}^{-1}\mathbf{X})^{-1} \mathbf{X}^+ \mathbf{v} = 0 \end{aligned} \quad (2.129)$$

As before, the minimum norm filter is the particular solution given by equation 2.125:

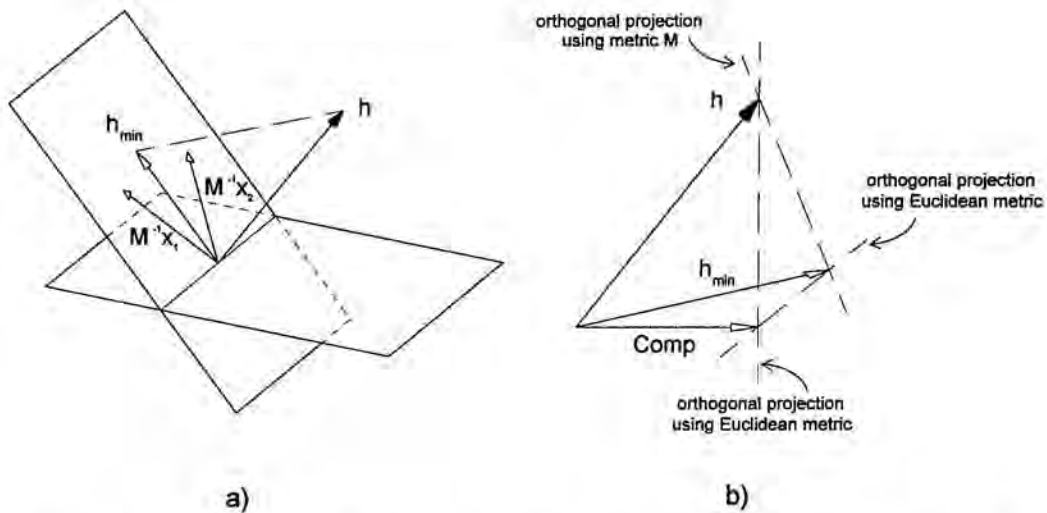
$$\begin{aligned}
 \mathbf{h}^+ \mathbf{M} \mathbf{h} = \|\mathbf{h}\|_{\mathbf{M}}^2 &= \left\| \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* + \left[ \mathbf{I} - \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \right\|_{\mathbf{M}}^2 = \\
 &= \left\| \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* \right\|_{\mathbf{M}}^2 + \left\| \left[ \mathbf{I} - \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \right\|_{\mathbf{M}}^2 \geq \\
 &\geq \left\| \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* \right\|_{\mathbf{M}}^2 = \|\mathbf{h}_P\|_{\mathbf{M}}^2 = \mathbf{h}_P^+ \mathbf{M} \mathbf{h}_P
 \end{aligned} \tag{2.130}$$

where the following equivalence has been used:

$$\begin{aligned}
 \left\| \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* + \left[ \mathbf{I} - \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \right\|_{\mathbf{M}}^2 &= \|\mathbf{h}_P + \mathbf{h}_H\|_{\mathbf{M}}^2 = \\
 (\mathbf{h}_P + \mathbf{h}_H)^+ \mathbf{M} (\mathbf{h}_P + \mathbf{h}_H) &= \mathbf{h}_P^+ \mathbf{M} \mathbf{h}_P + \mathbf{h}_P^+ \mathbf{M} \mathbf{h}_H + \mathbf{h}_H^+ \mathbf{M} \mathbf{h}_P + \mathbf{h}_H^+ \mathbf{M} \mathbf{h}_H = \\
 = \|\mathbf{h}_P\|_{\mathbf{M}}^2 + 0 + 0 + \|\mathbf{h}_H\|_{\mathbf{M}}^2 &= \left\| \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{c}^* \right\|_{\mathbf{M}}^2 + \left\| \left[ \mathbf{I} - \mathbf{M}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{X}^+ \right] \mathbf{v} \right\|_{\mathbf{M}}^2
 \end{aligned} \tag{2.131}$$

The geometrical interpretation of this result is as follows. We have found the filters that would give as central correlations with a set of patterns the values we want. We see that we have an infinite number of solutions. These filters can be expressed as a sum of two orthogonal terms, one being a linear combination of the training patterns and the other one a vector orthogonal to all of them. But in turn this decomposition is not unique, and can also be expressed as the sum of a linear combination of the training images modified in certain way and an orthogonal vector to all of them, where the orthogonality is now defined through a generalized metric. In all cases the projection of the filter onto the hyperplane generated by the images, that is the linear combination component is optimal with respect to a given quadratic criterion as sketched in Figure 2.11.a). Figure 2.11.b)





**Figure 2.11-** a) Orthogonal projection of a vector, according to a metric  $M$ , onto the hyperplane generated by the modified images b) Two possible decompositions for an SDF filter.

is an example of two possible decompositions for an SDF filter. The filter  $h$  can be expressed as the sum of the composite filter, called **Comp** in the figure, plus an orthogonal vector according to the Euclidean metric, i.e.  $M=I$ . On the other hand, we can also express this filter as the sum of a vector that minimizes  $h^+Mh$ , called  $h_{min}$ , plus an orthogonal vector following the metric  $M$ . Furthermore, the filter  $h_{min}$  can also be expressed in terms of the composite filter or the composite in terms of  $h_{min}$ .

The first analysis is a particular case of the later when the metric used is the identity, being the composite the associated optimal filter. Other cases of interest are:

- a) The metric is the average image energy,  $M = \tilde{D}$  with:

$$\tilde{D} = \frac{1}{K} \sum_{i=1}^K \alpha_i \tilde{D}^i \quad (2.132)$$

where

$$\tilde{\mathbf{D}}^i = \text{diag} \left( \left| \tilde{x}_1^i \right|^2, \dots, \left| \tilde{x}_N^i \right|^2 \right) \quad (2.133)$$

is a diagonal matrix that represents the energy of the  $i$ th image. The quantity minimized is therefore:

$$E = \tilde{\mathbf{h}}^+ \tilde{\mathbf{D}} \tilde{\mathbf{h}} = \frac{1}{K} \sum_{i=1}^K \alpha_i \tilde{\mathbf{h}}^+ \tilde{\mathbf{D}}^i \tilde{\mathbf{h}} \quad (2.134)$$

namely a weighted average of the correlation energy with the set of training images. The resulting optimal design is called minimum average correlation energy filter (MACE) and was proposed by Mahalanobis *et al.* in [Mah87]. The filter provides an indirect method of controlling the whole correlation plane rather than only the central point. By minimizing the average energy of the correlation plane the sidelobes are forced to be small and consequently, this filter gives sharp peaks and reduces the number of false alarms. The MACE filter can be considered the generalization to multiple images of the inverse filter. When there is only one image in the training set both MACE and inverse filter are the same:

$$\tilde{\mathbf{h}} = \tilde{\mathbf{D}}^{-1} \tilde{\mathbf{x}} \left( \tilde{\mathbf{x}}^+ \tilde{\mathbf{D}}^{-1} \tilde{\mathbf{x}} \right)^{-1} \mathbf{c}^* \quad (2.135)$$

The matrix inside the parenthesis is in this case,  $1 \times 1$ , that is a scalar, whose value is merely  $1/N$ . On the other hand, we have:

$$\tilde{\mathbf{D}}^{-1} \tilde{\mathbf{x}} = \begin{pmatrix} \frac{1}{\left| \tilde{x}_1 \right|^2} & & & & \\ & \dots & & & \\ & & \frac{1}{\left| \tilde{x}_N \right|^2} & & \\ & & & \dots & \\ & & & & \frac{1}{\left| \tilde{x}_N \right|^2} \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \cdot \\ \cdot \\ \tilde{x}_{N-1} \\ \tilde{x}_N \end{pmatrix} = \frac{\tilde{x}_i}{\left| \tilde{x}_i \right|^2} \quad (2.136)$$

and therefore

$$\tilde{h}_i = \beta \frac{\tilde{x}_i}{|\tilde{x}_i|^2} \quad (2.137)$$

where  $\beta$  is a certain constant.

The major drawbacks of the MACE are a notable noise sensitivity and a low generalization capability, giving small peaks with almost every image not included in the training set.

b) The metric is either the noise covariance matrix or the noise power spectral density, i.e.  $\mathbf{M} = \mathbf{S}$ , eq. 2.12, or  $\mathbf{M} = \tilde{\mathbf{S}}$ , eq. 2.13. In such a case the quantity to be minimized is the noise induced peak variance in either object or Fourier space:

$$\begin{aligned} E &= \mathbf{h}^+ \mathbf{S} \mathbf{h} \quad \text{or} \\ E &= \tilde{\mathbf{h}}^+ \tilde{\mathbf{S}} \tilde{\mathbf{h}} \end{aligned} \quad (2.138)$$

The resulting optimal filter is called minimum variance synthetic discriminant function (MVSDF) and was proposed by Vijaya Kumar [Vij86] in an attempt to obtain stable peaks in the presence of input noise. For white noise this filter reduces to the composite filter since the covariance -or equivalently the power spectral density- matrix becomes:

$$\mathbf{S} = \beta \mathbf{I} \quad (2.139)$$

whence

$$\mathbf{h} = \mathbf{S}^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{S}^{-1} \mathbf{X})^{-1} \mathbf{c}^* = \frac{1}{\beta} \mathbf{I} \mathbf{X} \left( \frac{1}{\beta} \mathbf{X}^+ \mathbf{I} \mathbf{X} \right)^{-1} \mathbf{c}^* = \frac{\beta}{\beta} \mathbf{X} (\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* = \mathbf{X} (\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* \quad (2.140)$$

Analogously to the MACE filter it is possible to show that the MVSDF is the generalization to multiple images of the classical matched filter. Therefore, the main difficulties of this design are the same as those of the CMF, namely a low discrimination capability, which makes necessary the inclusion of nontarget patterns in the training set, and broad correlation peaks. In addition, we find the specific problems of the SDF technique: large lateral peaks and sometimes mathematical problems in inverting the noise matrix. Notice that  $\mathbf{S}$  is an  $N \times N$  matrix, which for example for 128x128 pixel images gives 16384x16384 elements. In contrast to the matrix associated to MACE filters, this matrix is diagonal or easily invertible, say Toeplitz [Vij89a], only in certain cases.

c) The metric is a combination of the two cases above,  $\mathbf{M} = \mu \tilde{\mathbf{S}} + (1 - \mu) \tilde{\mathbf{D}}$  with  $\mu \in [0, 1]$ . The minimized criterion is a trade-off between the noise resistance and the peak sharpness, where the importance of these two conflicting goals is balanced by means of the parameter  $\mu$ :

$$E = \mu \tilde{\mathbf{h}}^+ \tilde{\mathbf{S}} \tilde{\mathbf{h}} + (1 - \mu) \tilde{\mathbf{h}}^+ \tilde{\mathbf{D}} \tilde{\mathbf{h}} \quad (2.141)$$

The optimal filter is called optimal trade-off synthetic discriminant function (OTSDF) [Réf90a] and for the single-image case coincides with the well known Wiener filter [Réf91a][Réf91b]. For  $\mu=0$ , the OTSDF is just the MACE filter and conversely, for  $\mu=1$ , the filter becomes the MVSDF. For intermediate values the optimal trade-off design gives, as its single-image counterpart, filters with intermediate characteristics between the two extremes in such a way that no filter can be found giving a better trade-off with the same degrees of freedom.

It is worth mentioning that in contrast with the single-image case, the optical efficiency can not be easily introduced in the SDF approach, because it involves nonlinear constraints for the filter components. This leads to a rather complicated mathematical problem for which an analytical solution has not been found and only a few proposals for designing phase-only SDF filters by means of iterative algorithms has been published. The problem of computing phase-only synthetic discriminant functions is one of the research topics of this dissertation and will be extensively discussed in Chapter 6.

## 2.6 Equivalence of circular harmonic filters and invariant SDFs.

The synthetic discriminant functions theory is a powerful mathematical tool to design filters with controlled responses. As we have seen this enables us to utilize the correlation to perform image classification or invariant pattern recognition. Moreover, the theory is general enough to encompass, as particular cases, many designs obtained on an independent basis, as for example the matched or the inverse filter. This seems to be also valid for the invariant filters developed by means of the decomposition into harmonic series as we show in this section. Although only the equivalence with the circular harmonics has been explicitly proved -to our knowledge anyway- it is our believe that this is the rule rather than the exception and constitutes a very interesting theoretical result that should be further studied. The equivalence of both approaches for the case of in-plane rotations has been proved by Réfrégier in [Réf90b], but since his discussion had a different goal -more devoted to optimal trade-off comparisons- we give here a different and more specific proof.

Let us consider a composite filter,  $\mathbf{h}$ , designed to achieve invariance with respect to in-plane rotations. The training set is formed by  $P$  rotated images spanning an angle of  $2\pi$ , that is they are tilted each  $2\pi/P$  radians. According to equations 2.89 and 2.91 the filter can be written as:

$$\mathbf{h} = \sum a_i \mathbf{x}_i = \mathbf{X}\mathbf{a} \quad (2.142)$$

where

$$\mathbf{a} = (\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* \equiv \mathbf{A}^{-1} \mathbf{c}^* \quad (2.143)$$

The matrix  $\mathbf{A}$  in the latter equation has as elements the correlations between the different images in the training set. The central correlation between two rotated images depends only on the relative angle between them and not on their absolute tilts. As a consequence,  $\mathbf{A}$  is a special type of matrix called Toeplitz, characterized by:

$$a_{ij} = a_{|i-j|} \quad (2.144)$$

and for which the following relation is valid [Réf90b]:

$$A_{lm}^{-1} = \frac{1}{P} \sum_{k=1}^P \tilde{A}_k^{-1} e^{j(2\pi/P)k(l-m)} \quad (2.145)$$

Substituting eq. 2.145 into 2.143 we get:

$$a_l = \sum_m A_{lm}^{-1} c_m^* = \frac{1}{P} \sum_m \sum_k \tilde{A}_k^{-1} e^{j(2\pi/P)k(l-m)} c_m^* \quad (2.146)$$

Now, if we force the SDF to give the same correlations that a circular harmonics filter would give:

$$c_m = e^{-j(2\pi/k)\alpha m} \quad (2.147)$$

where  $\alpha$  is the order of the harmonic, equation 2.146 becomes:

$$\begin{aligned} a_l &= \frac{1}{P} \sum_m \sum_k \tilde{A}_k^{-1} e^{j(2\pi/P)k(l-m)} e^{j(2\pi/P)\alpha m} = \\ &= \frac{1}{P} \sum_k \tilde{A}_k^{-1} e^{j(2\pi/P)kl} \sum_m e^{j(2\pi/P)(\alpha-k)m} \end{aligned} \quad (2.148)$$

which considering that:

$$\frac{1}{P} \sum_m e^{j(2\pi/P)(\alpha-k)m} = \delta_{\alpha-k} \quad (2.149)$$

leads to:

$$a_l = \sum_k \tilde{A}_k^{-1} e^{j(2\pi/P)kl} \delta_{\alpha-k} = \tilde{A}_\alpha^{-1} e^{j(2\pi/P)\alpha l} \equiv C e^{j(2\pi/P)\alpha l} \quad (2.150)$$

Substitution of eq. 2.150 into 2.142 gives:

$$\mathbf{h} = \sum_l \mathbf{x}_l a_l = \sum_l \mathbf{x}_l e^{j(2\pi/P)\alpha l} \quad (2.151)$$

where the unimportant constant factor has been omitted for convenience. If we remove the lexicographic scanning and retrieve the two-dimensional notation we have:

$$h_{i,j} = \sum_l x_{i,j}^l e^{j(2\pi/P)\alpha l} \quad (2.152)$$

where now the superscript indicates the image and the subscripts the pixel count. The above expression, in polar coordinates, can be written as:

$$h_{r,\theta} = \sum_l x_{r,\theta}^l e^{j(2\pi/P)\alpha l} \quad (2.153)$$

Assuming that the images are ordered by growing angles, the following identity can be established:

$$x_{r,\theta}^l = x_{r,(\theta+1)}^{l+1} \Rightarrow x_{r,\theta}^l = x_{r,(\theta+k)}^{l+k} \quad (2.154)$$

and therefore:



$$h_{r,\theta} = \sum_l x_{r,\theta}^l e^{j(2\pi/P)\alpha l} = \sum_k x_{r,(\theta-k)}^0 e^{j(2\pi/P)\alpha k} \quad (2.155)$$

which can be finally rewritten as:

$$h_{r,\theta} = \sum_{\omega} x_{r,\omega}^0 e^{j(2\pi/P)\alpha(\theta-\omega)} = \left[ \sum_k x_{r,\omega}^0 e^{j(2\pi/P)\alpha\omega} \right] e^{j(2\pi/P)\alpha\theta} \quad (2.156)$$

This expression is the discrete equivalent to eqs. 2.71 and 2.72, that is the circular harmonic filter of degree  $\alpha$ , a result that suggest that invariant filters are unique although can be computed in different ways.

## **Chapter three. Limitations of single-filter correlators**

### **Introduction.**

As commented on in the introduction of the first chapter, pattern recognition is of vital importance in several fields that have a great commercial interest. This has attracted people from different areas of mathematics, physics, computer science, biology, psychology, etc, and therefore it is not surprising to find a wide variety of methods and recipes to attack this difficult task. A given problem may be solved by means of statistical pattern recognition, neural nets, syntactic pattern recognition, graph matching, or correlation techniques to mention only a few -see for example [Son94] for more information-. In addition, some of these methods seems to be deeply related. For example, correlation may be viewed as a neural network [Ghe89], it can be considered as a discriminant function -see for example [Cas84]-, or mathematical morphology operations can be implemented through correlation [Cas92]. Thus it could be difficult to determine in general whether a problem is more suitable to be solved by means of correlation or neural networks or discriminant functions or mathematical morphology or cellular automata, or etc, and good comparison methods are then necessary. This chapter is devoted to adapt the formalism of the decision boundaries -one possibility to do so- to the specific case of correlation. In this way, the pros and cons of the correlation based techniques could be more easily understood, enabling the comparison with other classification procedures or between different optical architectures. This kind of analysis has been vaguely used before in the optical pattern recognition literature and we intend to carry out a more formal analysis and to take profit of the results. However such results are rather preliminary and should be considered as a starting point for future research.

### 3.1 Correlation as a discriminant function.

Classification procedures are, without doubt, about complexity. Any classification procedure, no matter whether we are classifying animals or stellar spectra, is intended to simplify the description of a complex phenomenon. If we can order a large set of elements, showing some of them similar features, into a smaller set of groups, the study of such objects becomes much simpler. For example we can analyze in detail one conspicuous representative of each class and extend the results to the rest of the group. That is the reason why classification techniques have a long tradition in the history of science.

The same argument apply to the classification of images. By grouping them into a small number of prespecified categories we may solve certain problems in a simpler form. This is usually accomplished by computing some kind of similarity measure between the input images and a set of reference patterns of each class.

The simplest problem is the two-category classification and its importance relies on the frequency with which simple decisions of this kind must be taken in pattern recognition applications. In addition, it forms the basis of more complex classification systems. For instance, the classification of an object as belonging to one of four possible categories (**A**, **B**, **C**, **D**) can be decomposed into two two-class problems in the following way: first, decide whether the objects belongs to (**A**  $\cup$  **B**), and second whether it belongs to (**A**  $\cup$  **C**), where the symbol **U** indicates the union of the two sets. Each object is completely determined: the only object which belongs to both is **A**; **B** only belongs to the first, **C** only to the second, and **D** to neither of them. In what follows we will deal only with the two class problem.

One of the most used methods in optical pattern recognition is that based on the correlation. The correlation between the input scene and a filter that codifies the information of the class is the measure of similarity utilized in these processes. The intensity of the correlation plane is binarized by means of a threshold function. Then, the input scene is assigned to the class if the correlation intensity is higher than the threshold. Otherwise the image is rejected. Because of a sampled image can be considered as a point in a  $P$ -dimensional space -thanks to the lexicographic scanning trick-, such a dichotomic decision divides the

hyperspace into two regions: the points that give a zero output and those that give an output equal to one.

The hypersurface that separates the two regions is called the decision boundary. A given classification problem can be solved if there is a decision boundary that separates the images according to the desired assignment of classes. For a procedure without arbitrarily shaped boundaries it is possible to find some problems that can not be correctly solved. The more complex the boundary functions are the higher the probability that a given problem can be solved by the procedure. For this reason it is of great interest to know what boundaries optical correlators produce and the number of dichotomies classifiable with these devices.

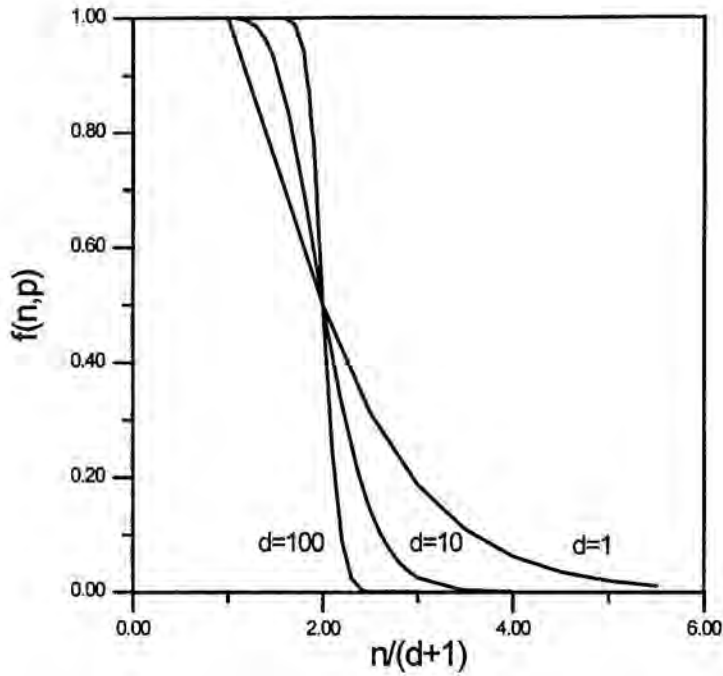
The simplest boundary is an hyperplane: the first order surface. It is produced by the well known linear classifier -Linear Discriminant Function (LDF)-. Let the vectors  $\mathbf{x}$  and  $\mathbf{h}$  represent the input scene and the vector that defines the LDF. The classification process is as follows: if and only if the inner product between both vectors surpasses a prefixed value  $\theta$ , the image is accepted as a member of the class. Therefore, the decision boundary can be written as:

$$\mathbf{h}^t \mathbf{x} - \theta = h_1 x_1 + h_2 x_2 + \dots + h_p x_p - \theta = 0 \quad (3.1)$$

where  $t$  means transpose. Let  $\mathbf{x}_i$ ,  $i=1, \dots, K$  be the set of images we want to accept, and let  $\mathbf{x}_i$ ,  $i=K+1, \dots, N$  those we want to reject. The problem is linearly classifiable if we can find an hyperplane  $\mathbf{h}$  and a threshold  $\theta$  so that

$$\begin{aligned} \mathbf{h}^t \mathbf{x}_i - \theta &> 0 & i = 1, \dots, K \\ \mathbf{h}^t \mathbf{x}_i - \theta &< 0 & i = K+1, \dots, N \end{aligned} \quad (3.2)$$

Obviously, this is not always possible. The probability for an arbitrary dichotomy to be linearly separable depends on the number of patterns  $-N-$  to be classified, and on the number of components of such images  $-P-$ . There are  $2^N$  possible problems of classification: the  $2^N$  different labelings of  $N$  objects with a binary digit. A certain fraction  $f(N, P)$  are linearly separable [Cov65][Dud73]



*Figure 3.1- The fraction of linearly separable dichotomies as a function of  $N/(D+1)$ .*

$$f(N,P) \leq \begin{cases} 1 & N \leq P+1 \\ \frac{2}{2^N} \sum_{i=0}^P \binom{N-1}{i} & N > P+1 \end{cases} \quad (3.3)$$

where the identity holds when the points are in general position, namely when no subset of  $P+1$  points fall in a  $(P-1)$ -dimensional space. Figure 3.1 represents the maximum given by equation 3.3 for several values of  $P$  and  $N$ . All ensembles of  $N < P+1$  are linearly separable. At  $N=2(P+1)$  half of the problems are linearly separable. When  $N$  is several times  $P$  the solution is unlikely.

If instead of a plane, the decision boundary is a polynomial surface, the number of classifiable dichotomies depends, in general, on the degree of the polynomial. The fraction is given by the same formula -eq. 3.3- but substituting

the dimension  $P$  by the number of adjustable parameters of the surface, i.e. the number of degrees of freedom [Cov65], which are given by the binomial

coefficient  $\binom{P+R}{R}$ , where  $R$  is the degree of the polynomial.

### 3.2 Decision boundaries in an optical correlator.

In this subsection we will represent the correlation filters as linear discriminant functions (LDF). From this analogy it follows some limits on the discrimination capability of correlation filters. Let the vectors  $\mathbf{x}$  and  $\mathbf{h}$  represent the input scene and the filter respectively in the spatial domain. For simplicity we will suppose that the filter is normalized, namely

$$\mathbf{h}^t \mathbf{h} = 1 \quad (3.4)$$

The central correlation is given by:

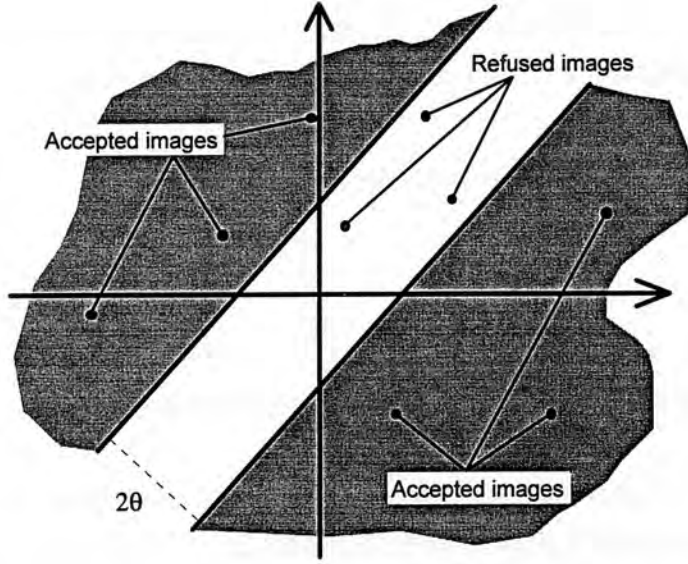
$$\mathbf{h}^t \mathbf{x} = c \quad (3.5)$$

This latter expression is a Linear Discriminant Function -see eq. 3.1-, but now the classification process is slightly different because of the square law detectors usually utilized: we detect the intensity of the correlation. The object is recognized if the correlation intensity  $c^2$  is higher than a threshold  $\theta^2$ . The new decision boundaries can be written as:

$$(\mathbf{h}^t \mathbf{x})^2 - \theta^2 = 0 \quad (3.6)$$

whence





*Figure 3.2- Decision boundaries produced by an optical correlator.*

$$(\mathbf{h}^t \mathbf{x} - \theta) (\mathbf{h}^t \mathbf{x} + \theta) = 0 \Rightarrow \begin{cases} (\mathbf{h}^t \mathbf{x} - \theta) = 0 \\ \text{or} \\ (\mathbf{h}^t \mathbf{x} + \theta) = 0 \end{cases} \quad (3.7)$$

From this expression we can deduce that the decision boundary is a pair of hyperplanes separated a distance  $2\theta$  apart, see Figure 3.2. The regions that are associated to the true class are those shaded in the Figure 3.2, namely when either both expressions  $(\mathbf{h}^t \mathbf{x} - \theta)$ , and  $(\mathbf{h}^t \mathbf{x} + \theta)$  are positives or both negatives. When they have opposite signs the total result is negative and the image will be rejected.

Since we have a second order surface with  $P+1$  adjustable parameters -the coefficients of the filter and the threshold-, the maximum number of separable dichotomies is limited by  $f(N,P)$  -eq. 3.3-. In figure 3.3 two separable problems by an optical correlator and the decision boundaries are shown. Two non-separable problems are also presented.

The possibility of arbitrarily classifying as many images as the dimension of the space deduced from eq. 3.3. seems to be a very good result. For example



for 128x128 pixels images this means that we can introduce in a single filter the information of 16384 patterns. Such number is probably enough for the majority of applications. However, until now we have only considered the central correlations. One of the advantages of the correlation methods is that they are position invariant. The correlations with the shifted images are also evaluated and the position of the peak indicates the position of the target. To profit this information it is necessary to control the correlation in all points  $-P-$  to avoid false alarms due to the sidelobes. The correlation in a given point is the same as the central correlation with the image shifted to that point. This implies that we pass from the problem of detecting one object, to the problem of detecting the image and rejecting all the  $P-1$  possible displacements. So, for an object we have  $N=P$  patterns to classify. According to Equation 3.3 this can be always solved with a single-filter correlation. The solution is the inverse filter that gives 1 at the origin and 0 elsewhere.

A geometrical interpretation may be enlightening. The displaced images lies in a manifold of as much as  $P-1$  dimensions. As the decision space is  $P$ -dimensional, there is always an orthogonal direction  $h$  to all of them. Thus, if the centered image has a non-zero projection onto  $h$ , the correlation will be a delta function by using  $h$  as the filter. But the problem appears when more than one object  $-N_o-$  has to be classified. The number of vectors in the feature space is  $N=P.N_o$ , and the probability for this problem to be separable decreases drastically. To obtain rotation invariance it is necessary to control the correlation with  $N_r$  rotated versions of the target, so for  $N_o$  objects the number of vectors will be  $N= P.N_o.N_r$ . That is the reason why high sidelobes are frequent in this case.

What appeared to be a very good method when only the central

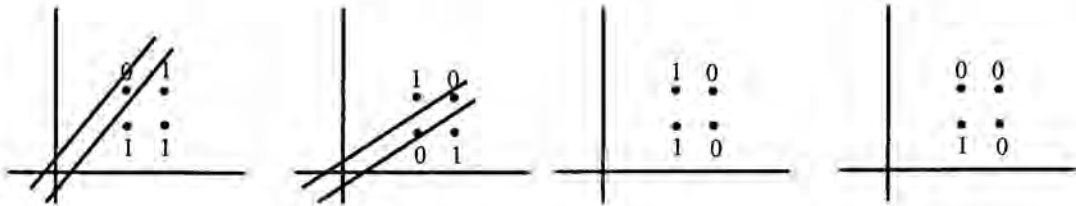


Figure 3.3- Two separable and two non-separable problems in a correlator

correlations were considered, seems to be a very bad one when invariances are required: no more than one image can be arbitrarily classified with total reliability when the position invariance is taken into account. Additional invariances produce the appearance of sidelobes with almost total certainty. However a more detailed analysis must be carried out because the experimental results are not so pessimistic. For instance the MACE filter permits to include several images producing delta like correlations. Also the generalization of the MACE filter [Rav91] seems to work properly for rotation invariant correlation. Possible explanations for this behavior are: on one hand, the projections of the false class images onto the filter must not be equal to zero, but lower than the threshold value  $\theta$ . On the other hand, and perhaps more important, we have to consider the possibility for the images to belong to a manifold of lower dimension than that expected. This means that the number of quasi-orthogonal directions may be higher and this fact permits to include more images. This question may be elucidated by means of a principal component analysis which enables to find the directions that accumulates the majority of the variance of the data.

Equation 3.3 gives an expression for an arbitrary classification problem, and for points distributed without restrictions, except the general position condition. The problems that can be solved by a given decision boundary are not always the same. Depending on the relative locations of the images, a given subset of the  $2^N$  possible dichotomies are separable. For a distinct arrangement a different set of problems can be solved. This implies that the probability for any dichotomy to be separable is limited by  $f(N,P)$  and the maximum is attained when the images are distributed in general position. If this condition is not fulfilled because the set of points lie in a lower dimensional subspace, the probability for an arbitrary dichotomy to be separated is smaller. However for some problems the probability may be higher. The problem of avoiding sidelobes is always the same dichotomy, namely, we want to reject the shifted images and to accept the centered ones. Therefore it is possible that by rejecting a fraction of the shifted images we accomplish to reject all of them being the problem equivalent to one with less images involved and thus being more probable.

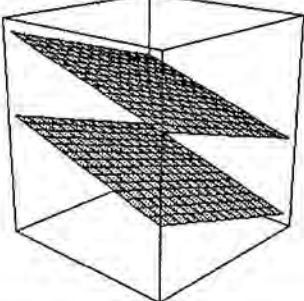
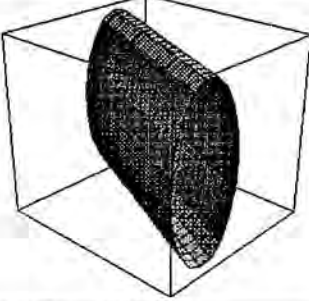
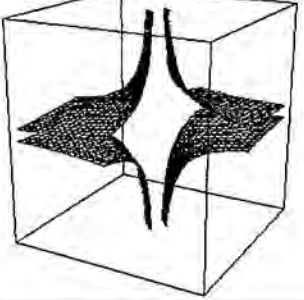
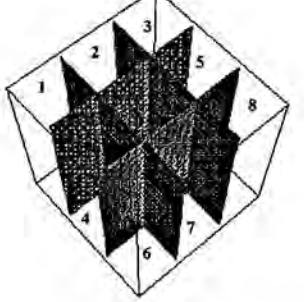
Decision boundaries	Combinations between Channels
	$\left(\sum \text{Channels}\right)^2 \Rightarrow$ $\left(h_x X + h_y Y + h_z Z\right)^2 > \theta$
	$\sum (\text{Channels})^2 \Rightarrow$ $\left(h_x X + h_y Y\right)^2 + \left(h_z Z\right)^2 > \theta$
	$\prod (\text{Channels})^2 \Rightarrow$ $\left(h_x X + h_y Y\right)^2 \cdot \left(h_z Z\right)^2 > \theta$
	<p style="text-align: center;"><b>Logic Operations</b></p> $\text{AND} \begin{cases} \left(h_x X + h_y Y\right)^2 > \theta_1 \\ \left(h_z Z\right)^2 > \theta_2 \end{cases}$ $\text{OR} \begin{cases} \left(h_x X + h_y Y\right)^2 > \theta_1 \\ \left(h_z Z\right)^2 > \theta_2 \end{cases}$

Figure 3.2- Different combination between two channels and the corresponding decision boundaries.

Even in this latter case the relative simple decision boundaries produced by an optical correlator seem to limit a great deal the possibilities of single filter correlations for handling the large number of images required for real-world problems. The solution may be the use of more complex discriminant functions by using multichannel filtering architectures as illustrated in Figure 3.4.

The Figure shows -for the two-filter case- how the information provided by the different correlations can be combined in a later postprocessing stage to obtain decision boundaries with more complex shapes. Now we have an added flexibility, with which new classification problems can be solved or perhaps, the old problems can be solved with more ease. The rest of this dissertation is an attempt to profit by this general idea, developing specific methods and algorithms for multichannel correlators. It is worth pointing out that the price of having a more powerful framework, is of course a more complex procedure. This leads to several practical problems that are also analyzed in subsequent chapters.

## **Chapter four. Sidelobe removal by a multichannel procedure: Real case**

### **Introduction.**

The design of filters for the VanderLugt optical correlator has undergone great progress during the past years as surveyed in chapter two. Inasmuch as filters with optimal properties can be obtained, the obstacles encountered in correlation should be considered as a product of the inherent limitations treated in the preceding chapter rather than as a fault in the design techniques. Let us recall some of the problems. The matched filter gives the maximum signal-to-noise ratio in the correlation plane but is unable to discriminate between similar objects. The discrimination capabilities can be enhanced as well as the light efficiency by using a phase-only filter whose main drawback is its small noise resistance. The inverse filter presents a low efficiency, a small generalization capability and an extreme noise sensitivity.

Furthermore, the response of these filters depends on the scale, orientation and in general any deformation in the input pattern. A possible solution to this problem is to expand the reference objects into a set of orthogonal functions that are invariant to one of these deformations. For instance with a circular harmonic expansion (CHC), rotation invariance can be achieved. A different possibility is the use of SDF filters. With them, the information of the objects to be recognized and the objects to be discriminated against can be introduced simultaneously.

The main difficulty of the last two approaches is the appearance of sidelobes. This phenomenon was explained in chapter three as a problem of classification -involving the centered and the displaced images- that can not be solved by means of the simple decision regions of single filter correlations. Here we present a method to eliminate them through a multifilter procedure, thus



illustrating how those limitations can be defeated by using multichannel correlations. This chapter is devoted to the case of real-valued correlation planes leaving the general case for the next chapter. The reason for this is twofold. On the one hand, the discussion of the real case is simpler and may help in understanding the general method. On the other hand, filters giving real correlations are the most frequently used -for example SDF filter with real constraints- and so they deserve an specific treatment.

#### 4.1 Method.

If we perform the correlation of an input image with two differently designed filters restricted to produce the same central correlation, we can postprocess the output distributions in order to improve the results. For example, if we binarize them by applying a given threshold and then multiply the output planes pixel by pixel -which is equivalent to process them with the logical operation AND-, we can eliminate the sidelobes that are more than the threshold value that are not common to both planes, and we can maintain the central peak.

The existence of false alarms in such a procedure depends on the appearance of common sidelobes. The method we propose ensures, provided certain conditions are met, that the sidelobes above the threshold are spatially disjoint in both correlations.

The idea of an image being processed by means of several digital or optical processes and of the final image being obtained by the pointwise multiplication of each single output has been applied in the past in various problems. The procedure is similar to operations commonly used in mathematical morphology. Casasent *et al.* [Cas92], using the correlation with two different filters, detected simple geometrical shapes immersed in high clutter and noise. The procedure involved the binarization of the correlations and the pixel-by-pixel multiplication of both results. The method was an optical implementation of the morphological hit-miss transform. More recently, Crowe *et al.* [Cro93] proposed the utilization of a similar method to reduce the sidelobes appearing in imaging

systems owing to the finite size of the pupil, thus improving the spatial resolution.

We will restrict our study to the case in which we obtain real-valued correlations. The general case of obtaining complex distributions -such as those produced by circular harmonic filters- is treated in the next chapter.

The method can be applied to a wide variety of filters -which is referred to as the base filter in what follows- and consists of addition and subtraction of a new filter -called the correcting filter- designed so that the following conditions are fulfilled:

(I) The filter is orthogonal to every image in the training set -the filter and the images are treated as vectors with the usual lexicographic ordering. This requirement ensures that the central correlations produced by the base filter remain unchanged.

(II) The correlation between the correcting filter and the images produces a constant plane of a predefined value. This condition is accomplished only in an approximate form by means of a Lagrange minimization process.

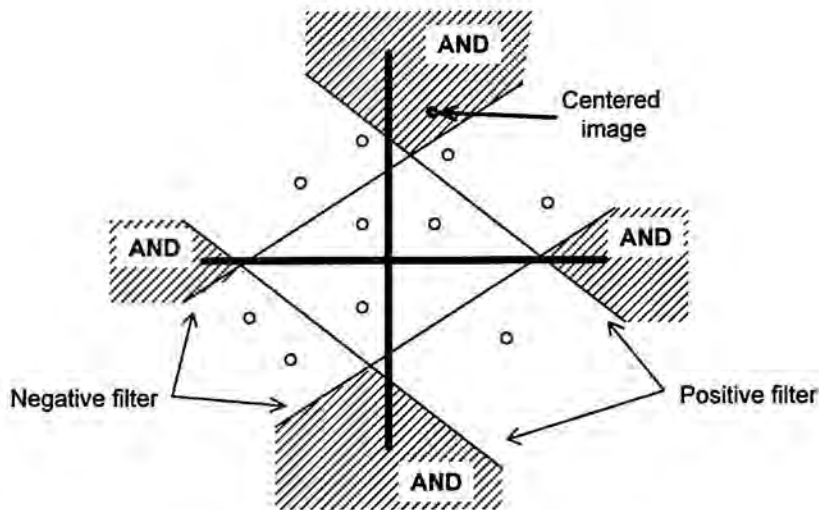
The output distributions obtained with the new filters have two terms:

$$(\mathbf{h}_b + \mathbf{h}_c) \star \mathbf{x}_i = \mathbf{h}_b \star \mathbf{x}_i + \mathbf{h}_c \star \mathbf{x}_i \quad (4.1)$$

$$(\mathbf{h}_b - \mathbf{h}_c) \star \mathbf{x}_i = \mathbf{h}_b \star \mathbf{x}_i - \mathbf{h}_c \star \mathbf{x}_i \quad (4.2)$$

where  $\mathbf{h}_b$  is the base filter,  $\mathbf{h}_c$  is the correcting filter,  $\mathbf{x}_i$  is one of the images and  $\star$  means correlation. If the training images and the filter are real the expression  $\mathbf{h}_b \star \mathbf{x}_i$  is real and may take positive and negative values. The term  $\mathbf{h}_c \star \mathbf{x}_i$  in equation 4.1, which is constant over the whole plane, increases the positive sidelobes and decreases the negative ones. Conversely, in equation 4.2,  $\mathbf{h}_c \star \mathbf{x}_i$  increases the negative sidelobes while decreasing the positive ones. Suitable choice of the value of the constant plane and the threshold ensures that no sidelobe is common to both binarized correlations. By applying the logical operation AND the sidelobes are eliminated.





**Figure 4.1-** Decision regions.

The correcting filter is a particular case of the minimum squared error synthetic discriminant function (MSE-SDF) introduced by Vijaya Kumar [Vij92a] and its mathematical expression is derived in the appendix of the next chapter.

The intuitive idea on which the procedure is based is sketched in Figure 4.1. The combination of the two filters reduces the region where the sidelobes will give false alarms. The centered image is represented by a black point while several displaced versions are represented as white points. All the sidelobes will be removed in the case represented in the figure.

One of the advantages of this method in front of the filters designed to reduce sidelobes such as the MACE filter, is that those designs try to be as orthogonal as possible with the displaced images and, as a consequence, the projection of the centered image is also very small. This means that the efficiency of these filters is usually very poor. In the proposed method the sidelobes are suppressed by the combined action of two filtering processes, thus not being necessary to impose the orthogonality of each single filter.

## 4.2 The correcting filter.

We derive in this section the mathematical expression for the correcting filter.

Let  $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_K$  denote the Fourier transforms of the  $K$  training images of  $N$  components. Let  $\tilde{\mathbf{h}}_c$  be the Fourier transform of the correcting filter.

Condition (I) in section 4.1 can be written as follows:

$$\sum_{j=1}^N \tilde{\mathbf{h}}_c^j \left( \tilde{\mathbf{x}}_i^j \right)^* = 0 \quad i = 1, \dots, K \quad (4.3)$$

where the superscript  $j$  indicates the pixel.

Condition (II) can only be achieved approximately by minimizing the following error function:

$$E = \frac{1}{K} \sum_{i=1}^K \sum_{j=1}^N \left| \tilde{\mathbf{d}}^j - \tilde{\mathbf{h}}_c^j \left( \tilde{\mathbf{x}}_i^j \right)^* \right|^2 \quad (4.4)$$

where  $\tilde{\mathbf{d}}$  represents the Fourier transform of the desired shape for the correlation between the images and the filter -a plane in our case-. Expression 4.4 is therefore a measure of the mean error between the correlations obtained and those desired. This filter is a particular case of the minimum squared error synthetic discriminant function (MSE-SDF) design introduced by Vijaya Kumar *et al* [Vij92a].

Equation 4.3 can be written compactly as

$$\tilde{\mathbf{S}}^+ \tilde{\mathbf{h}}_c = \mathbf{0} \quad (4.5)$$

where  $\tilde{\mathbf{S}}$  is an  $N \times K$  matrix formed by the Fourier transforms of the images arranged in columns:

$$\tilde{\mathbf{S}} = \begin{pmatrix} \tilde{x}_1^1 & \tilde{x}_2^1 & \dots & \tilde{x}_K^1 \\ \tilde{x}_1^2 & \tilde{x}_2^2 & \dots & \tilde{x}_K^2 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_1^N & \tilde{x}_2^N & \dots & \tilde{x}_K^N \end{pmatrix} \quad (4.6)$$

and the superscript + means the conjugate transpose of the matrix. Finally,  $\mathbf{0}$  represents the  $K$  dimensional vector with all its components zero.

Equation 4.4 can be expressed with the same formalism by defining the  $N \times N$  diagonal matrix:

$$\tilde{\mathbf{P}}_i = \begin{pmatrix} \tilde{x}_i^1 & 0 & \dots & 0 \\ 0 & \tilde{x}_i^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{x}_i^N \end{pmatrix} \quad (4.7)$$

With such definition the error can be written as:

$$E = \frac{1}{K} \sum_{i=1}^k [(\tilde{\mathbf{d}} - \tilde{\mathbf{P}}_i^* \tilde{\mathbf{h}}_c)^+ (\tilde{\mathbf{d}} - \tilde{\mathbf{P}}_i \tilde{\mathbf{h}}_c)] = \tilde{\mathbf{d}}^+ \tilde{\mathbf{d}} - \tilde{\mathbf{h}}_c^+ \tilde{\mathbf{r}} - \tilde{\mathbf{r}}^+ \tilde{\mathbf{h}}_c + \tilde{\mathbf{h}}_c^+ \tilde{\mathbf{P}} \tilde{\mathbf{h}}_c \quad (4.8)$$

where

$$\tilde{\mathbf{r}} = \frac{1}{K} \sum_{i=1}^K (\tilde{\mathbf{P}}_i \tilde{\mathbf{d}}) \quad (4.9)$$

and

$$\tilde{\mathbf{P}} = \frac{1}{K} \sum_{i=1}^K (\tilde{\mathbf{P}}_i^* \tilde{\mathbf{P}}_i) \quad (4.10)$$

Conditions (I) and (II) can be accomplished simultaneously by minimizing the following expression:

$$L[\tilde{\mathbf{h}}_c] = (\tilde{\mathbf{d}}^+ \tilde{\mathbf{d}} - \tilde{\mathbf{h}}_c^+ \tilde{\mathbf{r}} - \tilde{\mathbf{r}}^+ \tilde{\mathbf{h}}_c + \tilde{\mathbf{h}}_c^+ \tilde{\mathbf{P}} \tilde{\mathbf{h}}_c) - 2\boldsymbol{\lambda}^+ (\tilde{\mathbf{S}}^+ \tilde{\mathbf{h}}_c) \quad (4.11)$$

by means of a Lagrange optimization process. In expression 4.11,  $\boldsymbol{\lambda}$  denotes a  $K$ -dimensional complex vector containing the Lagrange multipliers.

By calculating the gradients with respect to the filter components and the Lagrange multipliers and setting them to zero we obtain the following expression for the correcting filter. The mathematical details can be found in [Vij92a].

$$\tilde{\mathbf{h}}_c = \left[ \mathbf{I} - \tilde{\mathbf{P}}^{-1} \tilde{\mathbf{S}} (\tilde{\mathbf{S}}^+ \tilde{\mathbf{P}}^{-1} \tilde{\mathbf{S}})^{-1} \tilde{\mathbf{S}}^+ \right] \tilde{\mathbf{P}}^{-1} \tilde{\mathbf{r}} \quad (4.12)$$

### 4.3 Necessary conditions.

Let us suppose we have designed a generalized SDF for solving a two-class problem. The prespecified values for the correlation with classes **A** and **B** are called  $p_0$  and  $p_1$ , respectively, and with no loss of generality we suppose that  $p_0 > p_1$ . The criterion for classification of an image as a member of one of the two classes is the following: if the correlation intensity at the center exceeds a given threshold, the image is classified as belonging to class **A**; otherwise, the image is assigned to class **B**.

As we have commented on in section 4.1, the method for elimination of sidelobes involves the correlations with two filters that have an opposite effect.

The first, which we call positive filter, reduces the negative sidelobes and enhances the positives ones. The second filter, called negative, reduces the positive peaks and enhances the negative ones.

Our goal is to determine the proper settings for the threshold and the constant plane resulting from the correlations between the images and the correcting filter in such a way that the only point in the output intensity distributions that passes the threshold in both cases is the central peak.

The equations that ensure the above statement can be written as follows:

$$\theta (|x| + c)^2 < p_0^2 \quad (4.13)$$

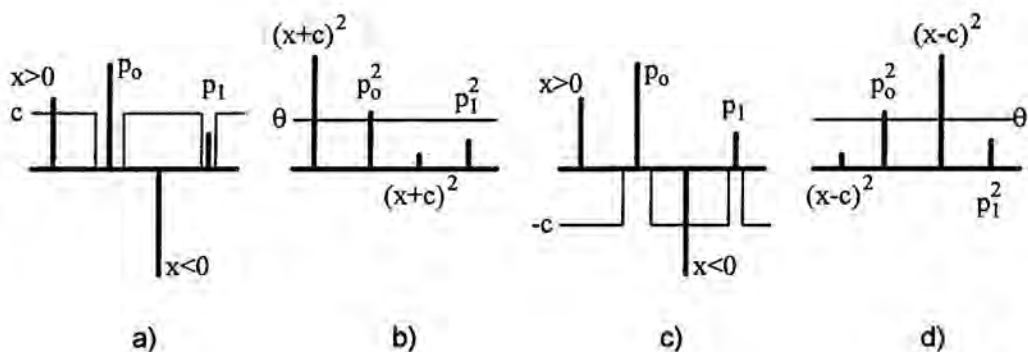
$$(|x| - c)^2 < \theta p_0^2 \quad (4.14)$$

$$c^2 < \theta p_0^2 \quad (4.15)$$

$$p_1^2 < \theta p_0^2 \quad (4.16)$$

where  $\theta$  is a factor between zero and one that represents the threshold,  $c$  is the value of the constant plane, and  $x$  is the height of the maximum sidelobe to be suppressed.

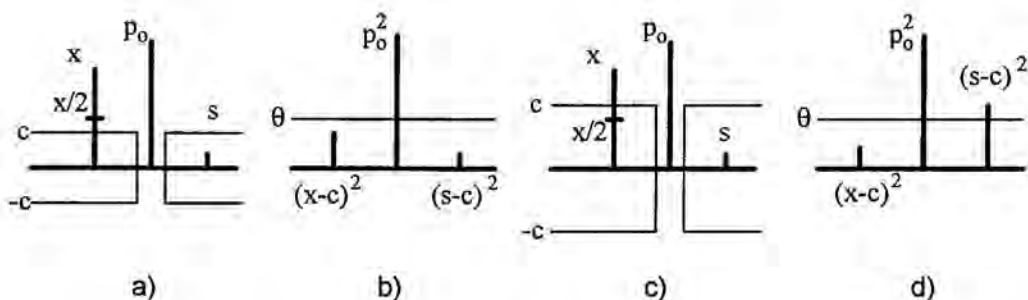
The necessity of conditions 4.13-4.16 is discussed in the following considerations. Maximum sidelobe  $x$  increased by constant  $c$  may become higher than value  $p_0$  of the correlation for class A. This situation is illustrated in Figure 4.2. In Fig. 4.2.a) the positive sidelobe  $-x>0-$  is increased by positive constant  $c$ . The resulting intensity can be seen in Fig. 4.2.b), in which it appears higher than the intensity of the central correlation  $p_0^2$ . For negative sidelobes we have an equivalent situation. In Figure 4.2.c) the negative sidelobe  $-x<0-$  is increased -in absolute value- by negative constant  $-c$ , and the resulting intensity -Fig. 4.2.d)- is also higher than the central correlation for class A. Inequality 4.13 is then necessary to guarantee that  $p_0$  passes the threshold. The expression  $(|x| + c)^2$  is the intensity of the increased peak no matter whether the maximum sidelobe is



**Figure 4.2-** a) Amplitude of the correlation plane produced by the base and the correcting filters when both are added (positive filter). The thin line represents the constant amplitude correlation value given by the correcting filter. The full lines represent the correlation peaks and the sidelobes obtained with the base filter. b) Intensity distribution for the positive filter. c) and d) Same as a) and b) for the case of the negative filter.

positive or negative. In other words, this factor takes into account the case  $(x - c)^2$  for negative peaks and  $(x + c)^2$  for positive ones.

Inequality 4.14 represents the condition for the decreased maximum sidelobe to be eliminated in the binarization, i.e., when the absolute value of the sidelobe is decreased, the resulting intensity has to be lower than the limit marked by the threshold value -see Figs. 4.2.a) and b) for  $x < 0$  and Figs. 4.2.c) and d) for  $x > 0$ -. The condition is written assuming that the threshold is applied



**Figure 4.3-** Illustration of the necessity of inequality 4.5: a) Case  $c < x/2$ . The full lines represent the amplitude of the correlation with the base filter, which produces a high (peak  $x$ ) and a small (peak  $s$ ) sidelobe. b) Intensity distribution corresponding to a). If  $c < x/2$ , sidelobes lower than  $x$  are suppressed by eq. 4.4. c) The same as a) when  $c > x/2$ . d) Intensity distribution corresponding to c). When  $c > x/2$ , small sidelobes increased by the constant may surpass the threshold value.

on  $p_0$ , but the maximum value in the output intensity distribution might be different. For instance, if we had two maximum sidelobes, both with the same absolute value but with different sign, and if the increased term  $(|x| + c)^2$  were higher than  $p_0^2$ , the threshold value would be  $\theta(|x| + c)^2$  because when one decreases, the other one increases -the case represented in Fig. 4.2-. By setting a more restrictive condition, such as equation 4.14, we can cover all the cases.

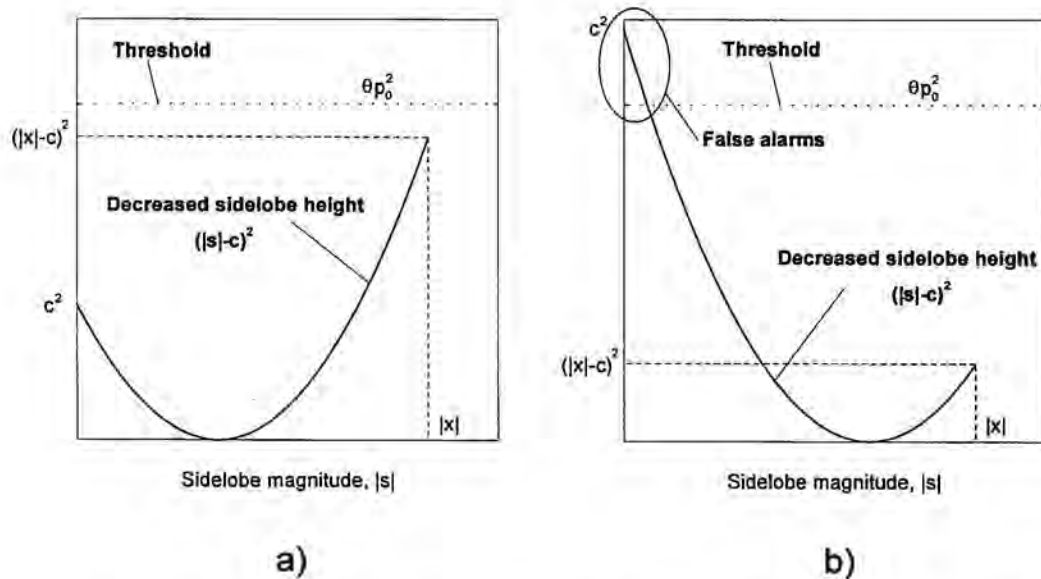
In addition, we need to ensure that the sidelobes lower than  $x$  disappear. We have to treat two cases separately:

a) If

$$c < |x|/2 \quad (4.17)$$

we have

$$(|x| - c)^2 > (|x| - |x|/2)^2 > c^2 \quad (4.18)$$



**Figure 4.4-** a) When  $c < |x|/2$ , elimination of maximum sidelobe  $x$  ensures the elimination of smaller sidelobes. b) When  $c > |x|/2$  a new condition is necessary to prevent false alarms.



In this situation equation 4.14 guarantees the elimination of every sidelobe ranging from zero to  $x$  -see Figs. 4.3.a), b) and Figure 4.4.a)-.

b) If  $c > |x|/2$  the former condition is not assured, and small sidelobes may surpass the threshold. This situation is depicted in Figs. 4.3.c),d) and 4.4.b).

Inequality 4.15 is then necessary to take into account case (b). Finally, inequality 4.16 is required for a correct classification of class **B**. In the typical case in which  $p_0=1$  and  $p_1=0$ , inequalities 4.13-4.16 become:

$$|x| - \theta^{1/2} < c < \theta^{-1/2} - |x| \quad (4.19)$$

$$c < \theta^{1/2} \quad (4.20)$$

As can be observed in the above expressions, the value of  $c$  is not completely determined by the conditions, but there is a range of permissible values that is necessary for the reliability of the method. This necessity is caused by, as mentioned above, the correlation planes with the correcting filter not being exact planes but approximate versions obtained by means of a minimization process.

Since the leftmost part of equation 4.19 is an increasing function with respect to  $|x|$  and the rightmost part is a decreasing function, the maximum sidelobe that fulfills both inequalities is obtained when:

$$|x| - \theta^{1/2} = \theta^{-1/2} - |x| \quad (4.21)$$

whence

$$|x| = \frac{1}{2}(\theta + 1)\theta^{-1/2} \quad (4.22)$$

On the other hand, by considering inequality 4.20 and the leftmost part of inequality 4.19, or:

$$c < \theta^{1/2}$$

$$c > |x| - \theta^{1/2}$$

we obtain the maximum  $|x|$  when

$$|x| - \theta^{1/2} = \theta^{1/2} \quad (4.23)$$

whence  $|x|$  is

$$|x| = 2\theta^{1/2} \quad (4.24)$$

and finally, the maximum sidelobe can be written as:

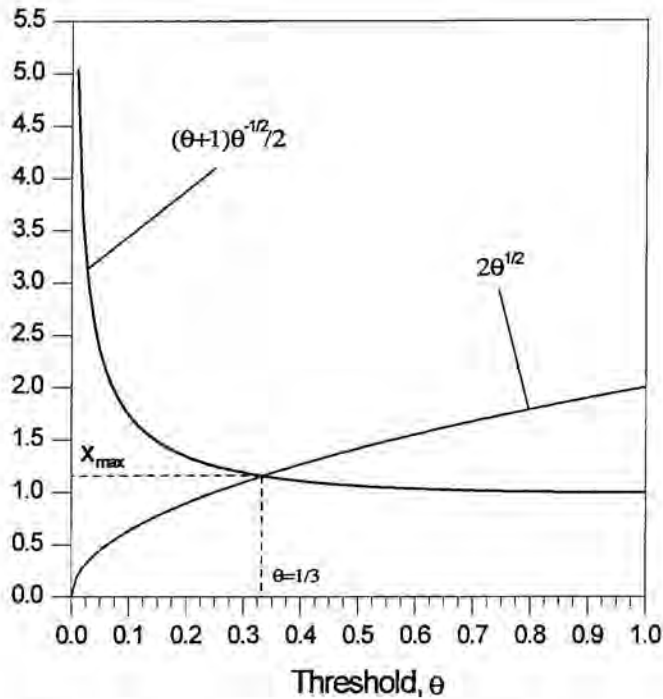
$$|x|_{\max} = \min \left[ 2\theta^{1/2}, \frac{1}{2}(\theta+1)\theta^{-1/2} \right] \quad (4.25)$$

The two expressions that determine the maximum sidelobe have been plotted in Figure 4.5. As can be observed, while the first is an increasing function of  $\theta$ , the second is a decreasing one. Therefore the maximum sidelobe that can be suppressed by the method is found at the point where the curves intersect each other, or in mathematical terms:

$$2\theta^{1/2} = \frac{1}{2}(\theta+1)\theta^{-1/2} \Rightarrow \theta = \frac{1}{3} \quad (4.26)$$

and substituting into eq. 4.25 we get  $|x|_{\max} \approx 1.15$  (i.e. 115% of the central correlation) in amplitude or  $|x|^2_{\max} \approx 1.32$  (132% of the central correlation) in intensity.

In practice it is not possible to reach this limit because the range of permissible values for  $c$  in inequality 4.19 is reduced to a single point. However, in most practical situations the method enables the elimination of sidelobes



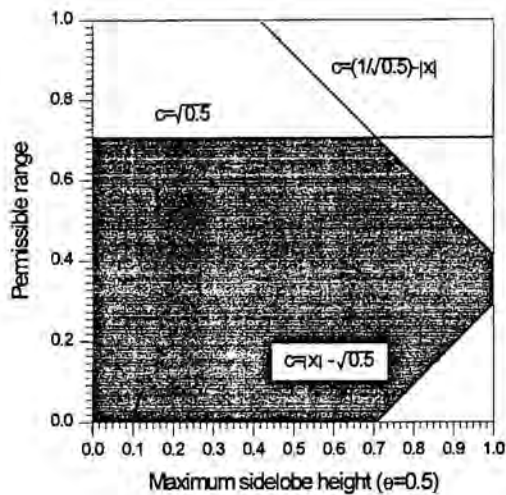
**Figure 4.5-** The plot shows the two conditions that determine the maximum sidelobe that can be suppressed.

higher than the central peak, which cannot be corrected by binarization of the output intensity produced by the base filter with a single threshold.

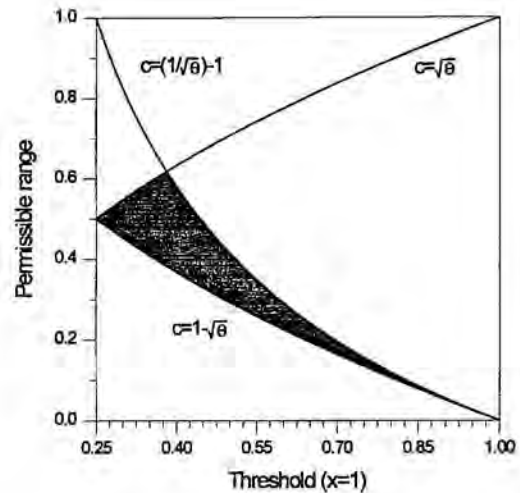
Figure 4.6 represents the permissible variation of constant  $c$  -shaded area- when the threshold value is fixed ( $p_0=1$ ,  $p_1=0$ ,  $\theta=0.5$ ). The graph shows that for small variations in  $c$ , large sidelobes can be eliminated, but for wide variations the height of the sidelobe must be small. By varying the threshold and by fixing the desired height of the sidelobe to be suppressed, we obtain Figure 4.7 ( $x=1$ ).

As can be observed, there is a value of  $\theta$  for which the permissible range of  $c$  is maximum because of the monotonic behaviour of the restrictions in inequalities 4.19 and 4.20. This optimum threshold can be calculated by use of the following equation:

$$\theta_{op}^{1/2} = \theta_{op}^{-1/2} - |x| \quad (4.27)$$



**Figure 4.6-** Permissible range for constant  $c$  as a function of the maximum sidelobe to be eliminated.



**Figure 4.7-** Permissible range for constant  $c$  as a function of the threshold.

Consequently, we can determine all the parameters envisaged by the method by selecting the maximum height of the sidelobe to suppress, by calculating the optimum threshold by means of eq. 4.27, and finally by choosing the constant  $c$  as the midpoint value of the range given by inequality 4.19.

The validity of the method is determined by the extent to which the actual correlations between the images and the correcting filter are constant planes with the expected accuracy. The possibility of the sidelobes being eliminated in the range from 0 to a prespecified value  $|x|$ , depends on whether the minimization procedure is capable of producing a correcting filter so that the correlations obtained with the images in the training set satisfy inequality 4.19.

Therefore the performance of the method is determined by the deviations in the correlation distribution from the expected plane; in other words, the maximum sidelobe that can be suppressed depends on the range of variation of the points in the correlation plane. On the other hand, the height of the sidelobes depends on the similarity between images in different classes; i.e., the more similar the images with different conditions are, the larger the expected values of the sidelobes become.

The method gives more accurate correlations planes, i.e., the correcting power is higher, when the images in the training set are more similar and

therefore when higher sidelobes appear. To demonstrate this property, let us assume we have two  $N$ -dimensional images whose Fourier transforms are  $X_1(w)$  and  $X_2(w)$ .

The expression for the error in equation 4.4 can then be written as:

$$E = \frac{1}{2} \left( \sum_{w=1}^N |D(w) - H_c(w) X_1^*(w)|^2 + \sum_{w=1}^N |D(w) - H_c(w) X_2^*(w)|^2 \right) \quad (4.28)$$

or in vectorial form

$$E = \frac{1}{2} [(\mathbf{d} - \mathbf{P}_1^* \mathbf{h}_c)^+ (\mathbf{d} - \mathbf{P}_1^* \mathbf{h}_c) + (\mathbf{d} - \mathbf{P}_2^* \mathbf{h}_c)^+ (\mathbf{d} - \mathbf{P}_2^* \mathbf{h}_c)] \quad (4.29)$$

Let us suppose that  $\mathbf{P}_1$  and  $\mathbf{P}_2$ , which are diagonal matrices, can be inverted -that is, there is no frequency for which the Fourier transform of the images is zero- and let us define filters  $\mathbf{h}_1$  and  $\mathbf{h}_2$  so that:

$$\mathbf{h}_1 = (\mathbf{P}_1^*)^{-1} \mathbf{d} \quad (4.30)$$

$$\mathbf{h}_2 = (\mathbf{P}_2^*)^{-1} \mathbf{d} \quad (4.31)$$

$\mathbf{h}_1$  and  $\mathbf{h}_2$  represent the filters that give exactly the desired shape when they are correlated with images  $X_1$  and  $X_2$  respectively. The assumption that  $\mathbf{P}_1$  and  $\mathbf{P}_2$  are invertible is not so restrictive, and similar requirements are needed in other filter designs. In particular, in MACE filters, the matrix that represents the average energy of the images in the training set must also be full rank.

Because  $\mathbf{h}_c$  is the optimal filter, by substituting  $\mathbf{h}_c$  for  $\mathbf{h}_1$  in equation 4.29 we have:

$$\begin{aligned}
E &\leq \frac{1}{2} \left[ (\mathbf{d} - \mathbf{P}_1^* \mathbf{h}_1)^+ (\mathbf{d} - \mathbf{P}_1^* \mathbf{h}_1) + (\mathbf{d} - \mathbf{P}_2^* \mathbf{h}_1)^+ (\mathbf{d} - \mathbf{P}_2^* \mathbf{h}_1) \right] = \\
&= \frac{1}{2} \left[ (\mathbf{d} - \mathbf{P}_2^* \mathbf{h}_1)^+ (\mathbf{d} - \mathbf{P}_2^* \mathbf{h}_1) \right] = \frac{1}{2} \left[ (\mathbf{P}_2^* \mathbf{h}_2 - \mathbf{P}_2^* \mathbf{h}_1)^+ (\mathbf{P}_2^* \mathbf{h}_2 - \mathbf{P}_2^* \mathbf{h}_1) \right] = \\
&= \frac{1}{2} \left\{ [\mathbf{P}_2^* (\mathbf{h}_2 - \mathbf{h}_1)]^+ [\mathbf{P}_2^* (\mathbf{h}_2 - \mathbf{h}_1)] \right\}
\end{aligned} \tag{4.32}$$

and from eqs 4.30 and 4.31:

$$\mathbf{P}_2^* \mathbf{h}_2 - \mathbf{P}_1^* \mathbf{h}_1 = \mathbf{0} \tag{4.33}$$

If we express image  $X_2$  as a function of  $X_1$  we can write:

$$\mathbf{P}_2 = \mathbf{P}_1 + \Delta \tag{4.34}$$

and therefore

$$(\mathbf{P}_1^* + \Delta^*) \mathbf{h}_2 - \mathbf{P}_1^* \mathbf{h}_1 = \mathbf{0} \tag{4.35}$$

$$\mathbf{P}_1^* (\mathbf{h}_2 - \mathbf{h}_1) = -\Delta^* \mathbf{h}_2 = -\Delta^* (\mathbf{P}_1^* + \Delta^*)^{-1} \mathbf{d} \tag{4.36}$$

As  $\mathbf{P}_1$  is a full-rank matrix, if  $\Delta$  tends to zero, then  $(\mathbf{h}_2 - \mathbf{h}_1)$  tends to zero, and in consequence the error in inequality 4.32, which depends on this difference, becomes increasingly small:

$$\Delta \rightarrow \mathbf{0}, \mathbf{P}_2 \rightarrow \mathbf{P}_1 \Leftrightarrow (\mathbf{h}_2 - \mathbf{h}_1) \rightarrow \mathbf{0} \Leftrightarrow E \rightarrow 0 \quad \text{Q.E.D.} \tag{4.37}$$

We carried out an experimental verification of this property. The details are given in section 4.4.

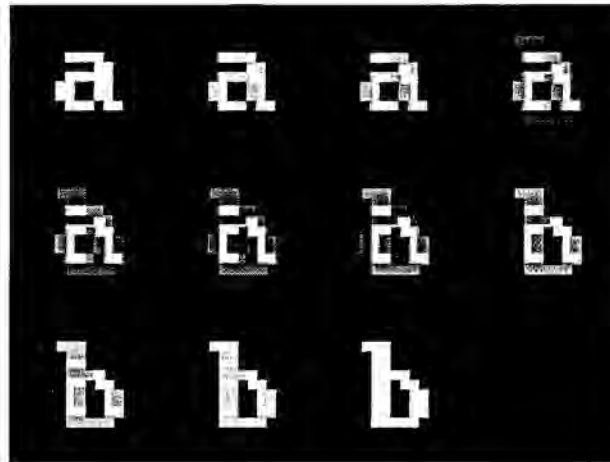
#### 4.4 Computer experiments.

In this section we present the results of several experiments carried out by means of a computer simulation in order to test the suitability of the method in practical situations. We performed a study of the dependence between the value of the expected sidelobes and the correcting capabilities of our method by using the images depicted in Figure 4.8. A set of ten correcting filters was designed, each of them calculated by use of a pair of images from the sequence  $ab(0)$ - $ab(i)$ ; namely, filter 1 was calculated with  $ab(0)$  and  $ab(1)$ , filter 2 with  $ab(0)$  and  $ab(2)$ , and so on. The measure of the similarity between image  $ab(i)$  and  $ab(0)$  was calculated with the following expression:

$$S[ab(0), ab(i)] = \frac{\left| [ab(0) \star ab(i)]_{(0,0)} \right|^2}{\left| [ab(0) \star ab(0)]_{(0,0)} \right|^2} \quad (4.38)$$

where the symbol  $\star$  means correlation.

The error function in eq. 4.4 as well as the deviation from a perfect plane



*Figure 4.8-* Sequence of images used in the simulation. Letter  $a$  is  $ab(0)$ , letter  $b$  is  $ab(10)$ , and the intermediate patterns are  $ab(i)$ , with  $i=1, \dots, 9$ .



( $c=0.35$ ) for each single image was computed for every correcting filter, and the results are represented in Fig. 4.9. The graph shows the dependence between these deviations from the expected shape and the similarity measure given by eq. 4.38. For very similar images such as  $ab(0)$  and  $ab(1)$  - $S(ab(0),ab(1)) = 0.96$ - the error is small - $E=0.25$ - and conversely, when the similarity between images is small - $S(ab(0),ab(1)) = 0.57$  for  $ab(0)$  and  $ab(10)$ - the error is high - $E=10.3$ -. The result shows the expected behavior; i.e., when large sidelobes are more likely to appear, owing to the similarity between images with different constraints, the procedure we propose is more powerful because of a smaller variation with respect to the desired plane.

The increasing correcting power enables the elimination of sidelobes, even if they are higher than the central peak as illustrated in Figure 4.10. In Figure 4.10, two images with a similarity  $S=0.90$  were used to build a MACE filter by imposition of images (a.1) and (a.2) to give values of 1 and 0,

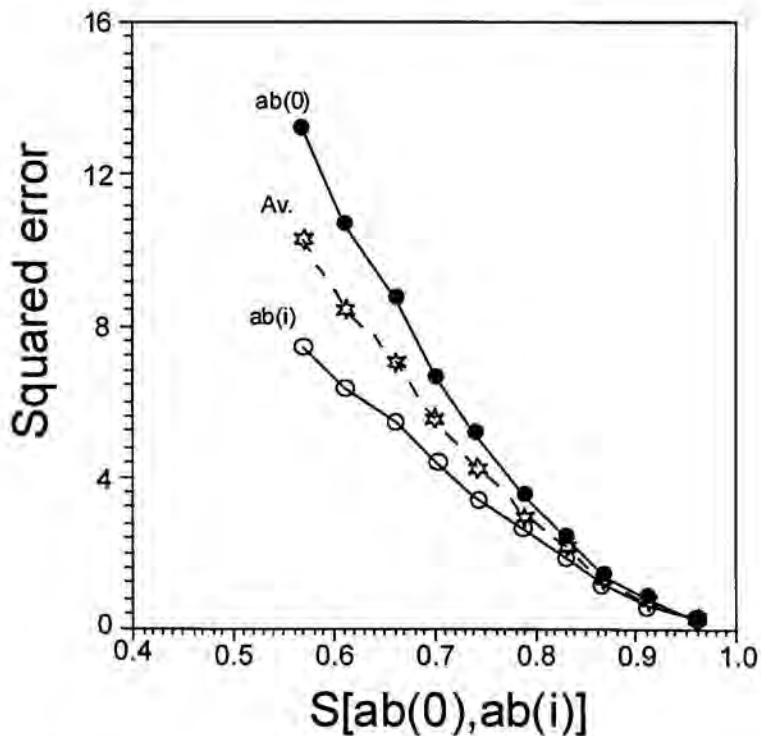
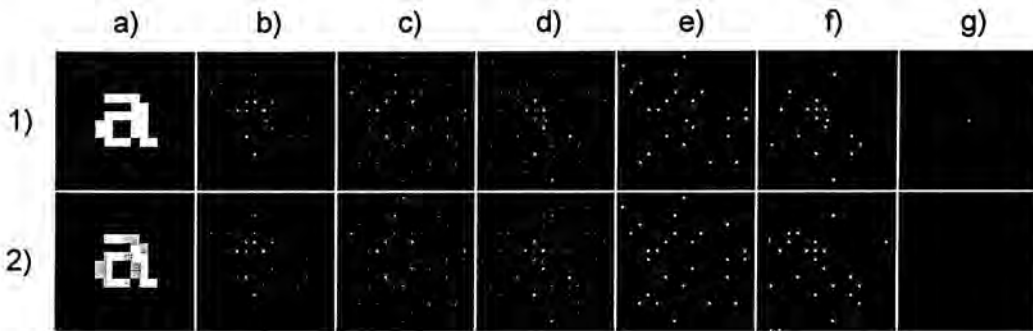


Figure 4.9- Squared error as a function of the similarity between images.



**Figure 4.10-** a.1, a.2, Images used to design the MACE filter. The imposed values were 1 for image a.1 and 0 for image a.2. b.1, b.2, Intensity of the correlation between the MACE filter and images a.1 and a.2, respectively. c.1, c.2, Intensity of the correlation between the positive filter and images a.1 and a.2. d.1, d.2, Intensity of the correlation between the negative filter and images a.1 and a.2. e.1 and e.2, Same as images c.1 and c.2 binarized with  $\theta=0.36$ . f.1, f.2, Same as images d.1 and d.2 binarized with  $\theta=0.36$ . g.1, Result of pixel-by-pixel multiplication of images e.1 and f.1. g.2, Result of pixel-by-pixel multiplication of images e.2 and f.2.

respectively. The MACE filter is an antisidelobe design, but in this situation it gives several lateral peaks, the largest of which has a value of 126% of the central correlation -in intensity- as shown in images (b.1) and (b.2).

In order to eliminate the sidelobes, we prepared the correcting filter using the following set of parameters:

$$\begin{aligned} |x| &= (1.26)^{1/2} = 1.12 \\ \theta_{op} &= 0.36 \\ c &= 0.53 \end{aligned}$$

The results of the correlations between the positive and negative filters are shown in images (c.1)-(d.2) of Figure 4.10, and the binarized results are shown in images (e.1)-(f.2). After pixel-by-pixel multiplication of both binarized planes, images (g.1) and (g.2) were obtained. As can be observed, all the sidelobes are suppressed and a perfect detection of the central correlations is possible.

The method is then capable of producing a significant increase in the discriminant abilities of the SDF filters, including those such as the MACE design, which are specifically designed to avoid the appearance of sidelobes. However, there is no noise resistance included in the minimum squared error SDF filter, so the procedure is highly sensitive to noisy inputs.

The possible solution to this problem is the same as that used to introduce noise resistance in the MACE design: the optimal trade-off technique. A further study on the suitability of this solution should therefore be carried out.

#### 4.5 Optical results.

The method proposed for elimination of sidelobes was tested by use of a convergent correlator [Van92]. This setup has an advantage in that it permits easy matching between the scales of both the input image and the filter.

The filters were built by means of computer-generated holograms codified by Burkhardt's method [Bur70], displayed on a laser printer, and photoreduced. The holograms were sandwiched to avoid uncontrolled phases owing to thickness variations in the photographic film. A low-power He-Ne laser provided the coherent illumination. Finally, a CCD camera and a frame grabber were used to capture the resulting correlation distributions.

The images that were used in the design of the filters are shown in Figure 4.11. The values imposed for the correlation at the origin were 1, 1, 0 for S, C, and E respectively. The correlations between the three letters and a composite



*Figure 4.11- Images used to design the filters. The imposed values for the central correlations were 1 for S, 1 for C, and 0 for E.*

filter are shown in the first column of Figure 4.12, in which large sidelobes can be observed. Figure 4.13 presents a three-dimensional plot of these correlations.

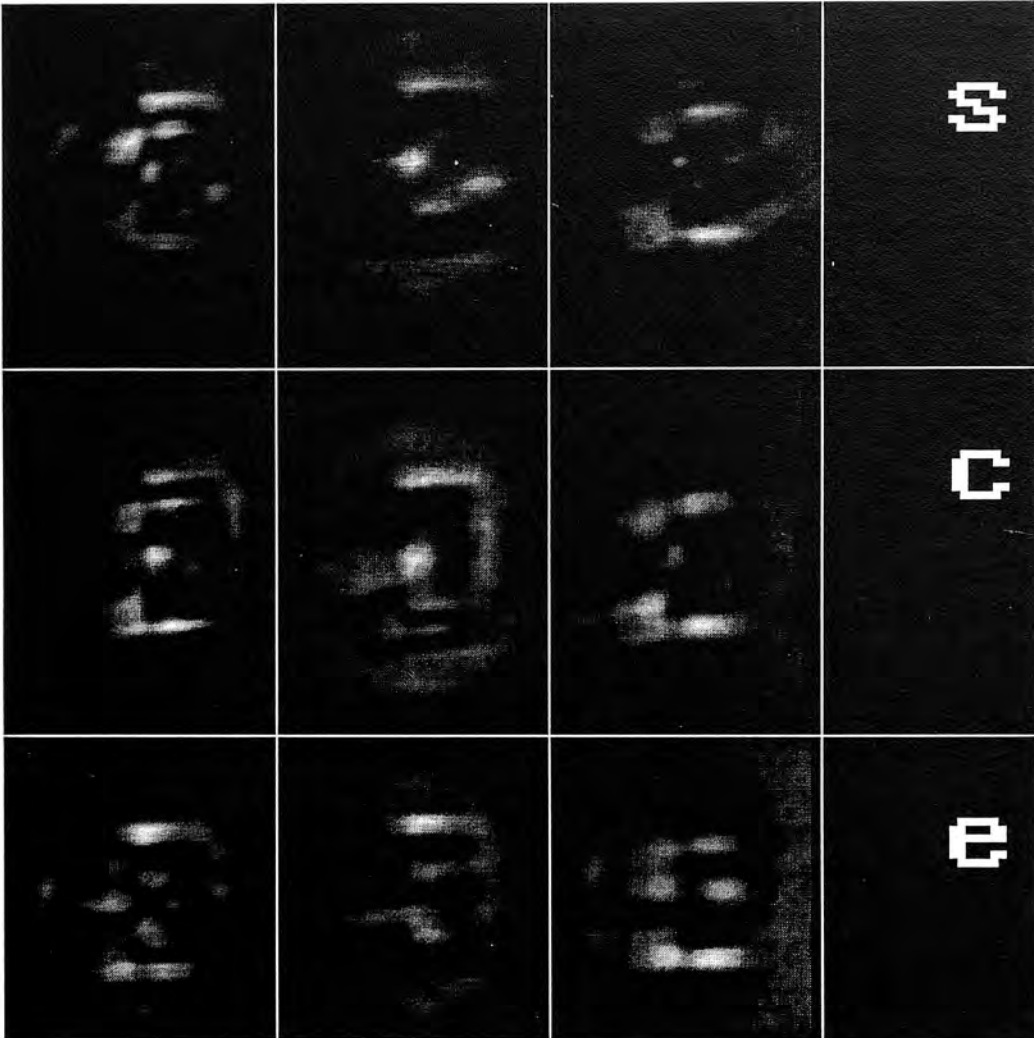
The correlations obtained with the positive and the negative filters are depicted in the second and third columns of Figure 4.12 and in a three-dimensional view in Figures 4.14 and 4.15, respectively. As can be seen, the effects of the two filters are the opposite; the sidelobes reduced by the first one are enhanced by the other and vice versa, as expected. By binarizing the results obtained with the positive and negative filter with a threshold of 0.40, we obtain the two first columns of Figure 4.16. Finally, in the third column of Figure 4.16, the results of pixel-by-pixel multiplication of binarized images are represented.

The results observed in the optical implementation were satisfactory and showed good agreement with previous computer simulations. Therefore the method seems to be suitable for application in a practical situation.

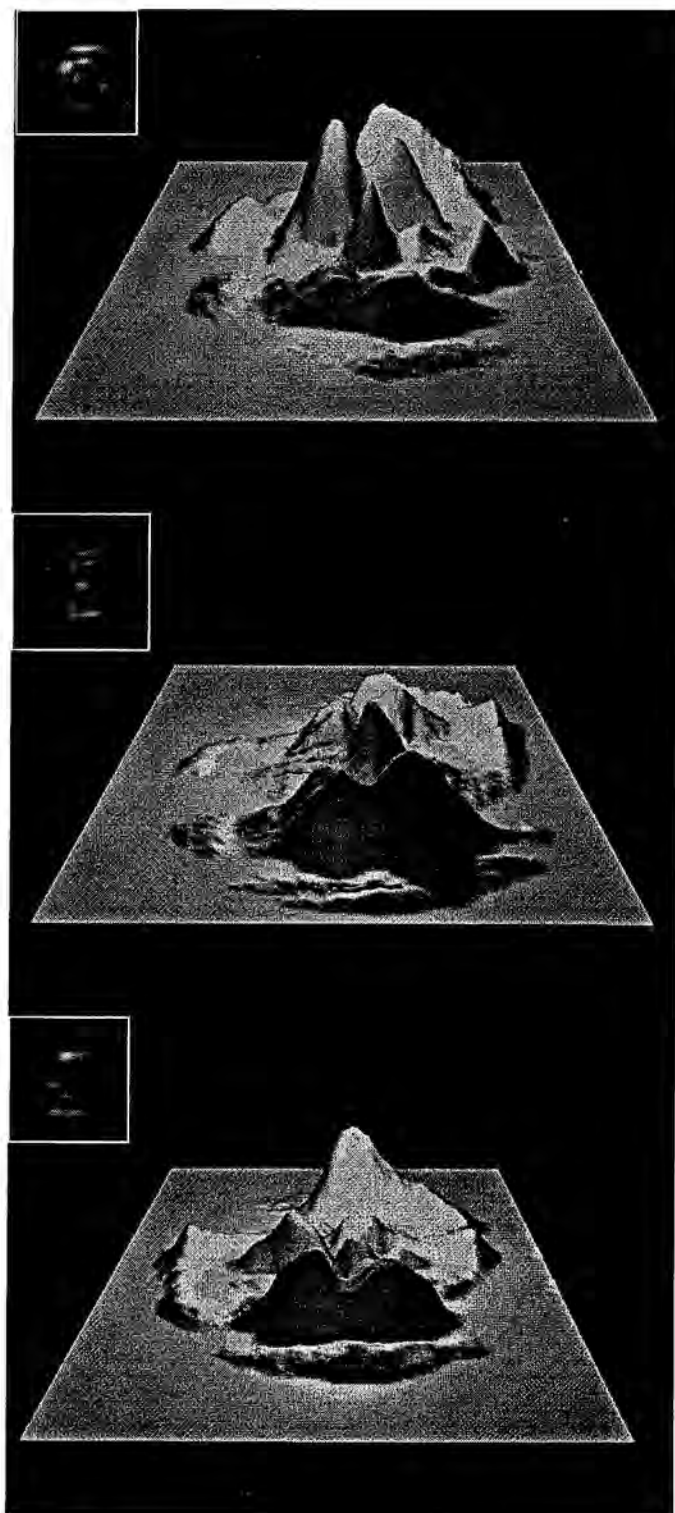
The conclusions of this chapter can be summarized as follows. The existence of lateral peaks is one of the most important problems in optical pattern recognition by means of correlation. In this study we present a method that eliminates every sidelobe within a given range, provided certain conditions are fulfilled. The method has the following properties:

- It can be applied to a wide variety of filters.
- The method ensures the elimination of the sidelobes if certain conditions are satisfied.
- Sidelobes higher than the central correlation can be suppressed.
- The method is more powerful when higher sidelobes are expected.

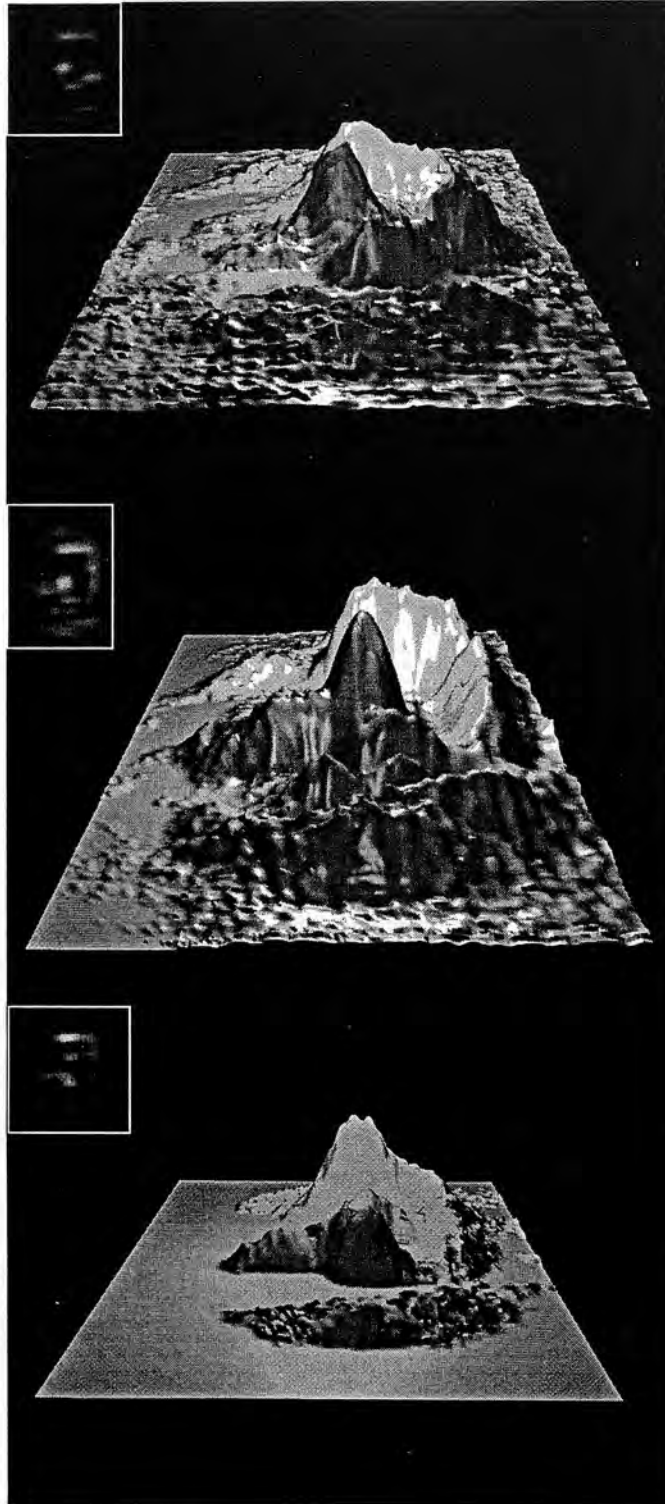
The procedure was tested with simulation and optical implementation and gave satisfactory results in both cases. The method presents a high sensitiveness to noisy inputs although the use of trade-off techniques may presumably overcome that difficulty.



*Figure 4.12- First column shows the results obtained with a composite filter designed to recognize letters *s* and *c* and to reject letter *e*. As can be seen the appearance of sidelobes makes impossible the correct classifications of the letters. Second and third columns show the correlation with the base plus the correcting filters and with the base minus the correcting filters, respectively. Note the opposite effect of the two filters. The high sidelobes on the second column correspond to low sidelobes on the third, and vice versa.*

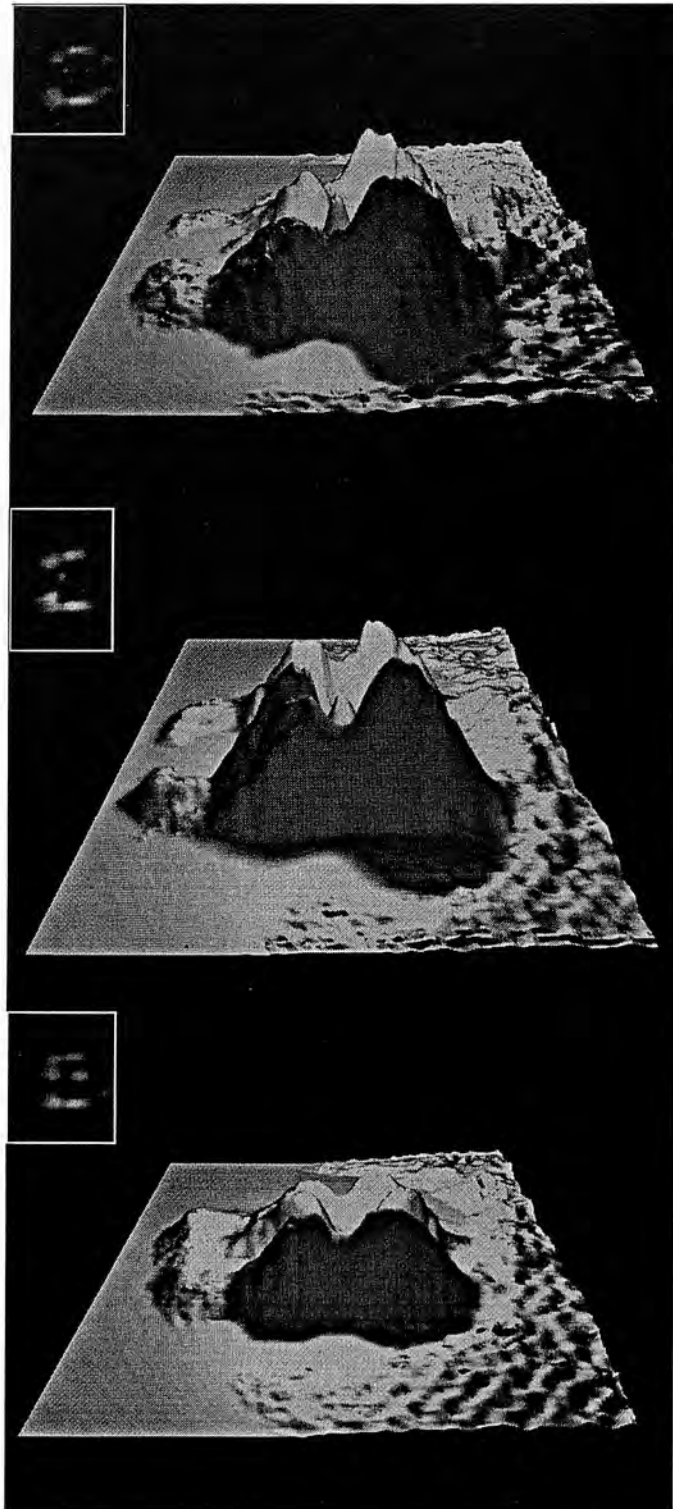


*Figure 4.13- 3-D plots of the correlations with the composite filter.*

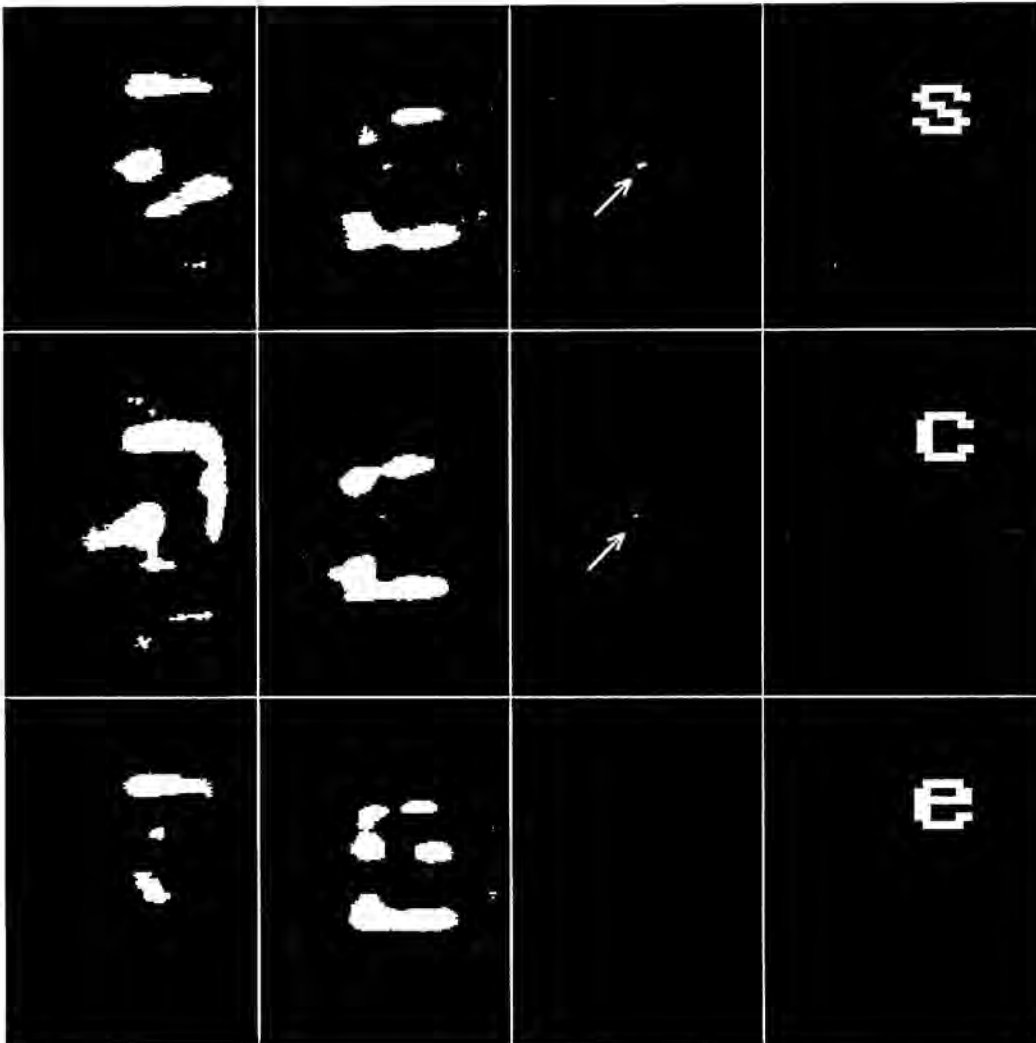


*Figure 4.14- 3-D plots of the correlations with the positive filter*





*Figure 4.15- 3-d plots of the correlations with the negative filter*



*Figure 4.16- First and second columns are the binarized versions of the columns two and three of figure 4.12. The threshold used was  $\theta=0.4$  of the maximum value. As can be seen the sidelobes in both columns are spatially disjoint. Column three is the pixelwise multiplication of columns one and two. A perfect detection is possible since all the sidelobes are removed.*

## **Chapter five. Sidelobe removal by a multichannel procedure: Complex case**

### **Introduction.**

Chapter 4 introduced a novel procedure to eliminate sidelobes by means of a two-filter correlation. The procedure shows clear advantages over the single-filter solutions such as the minimum average correlation energy (MACE) filter, and can be applied to them as well, to increase their discrimination capabilities when necessary. However the method was derived under the simplifying assumption of real-valued correlations. As already commented this is a very common situation, appearing when both filter and scene are real-valued in object space, but is not the most general case. There are several filters that do not fulfill the former condition, that is, their impulse response can be complex-valued, such as the circular harmonic filters or SDF filters made by imposing complex correlations with the training images. This latter case may be interesting to further maximize the quality criteria by taking profit of the additional degrees of freedom resulting from the removal of the symmetry constraints. For example Vijaya Kumar in [Vij88] and Réfrégier in [Réf90c] explore this possibility.

In this chapter the sidelobe elimination method is extended to complex correlation planes. The process is completely analogous to that of chapter four: it involves several correlations, the binarization of the results and a pointwise logic operation between the output planes. The main difference is that now we need more filters to obtain similar results. For the sake of clarity the method is presented from the beginning sometimes duplicating reasonings already developed in the preceding chapter. We believe that this may help to understand better the rather complicated arguments without appearing repetitive.

## 5.1 Method.

The procedure can be intuitively described as a method to share between several filters the task of eliminating the lateral peaks. The traditional approach to prevent the appearance of false alarms due to large sidelobes consists in designing a filter that gives high values for the central correlations with the target images while minimizing the average energy of the correlation plane. This leaves little energy for the rest of the correlation plane and thus makes unlikely the presence of large sidelobes. However, in some cases such minimization is not enough, for example when the filter is forced to give different correlations with similar objects. Moreover, the minimization of the energy leaves little energy not only for the sidelobes but also for the central correlations, that is the optical efficiency of this kind of filters is in general low. This proposal permits to distribute the load between a set of filters in such a way that the first filter has only to eliminate a subset of the total amount of sidelobes, the second one a different subset and so on. With an appropriate design of the filter database we can ensure the suppression of the whole set of lateral peaks.

The multichannel method ensures that after the filtering, binarization and multiplication stages all the sidelobes are eliminated if a set of conditions are satisfied. The higher the number of filters used the easier the conditions are met. This procedure can be applied to a wide variety of filters -which will be referred to as the base filters in what follows- and consists in adding a new filter -called the correcting filter-, weighted with a complex value of unit magnitude, with the following characteristics:

(I) The correcting filter is orthogonal to every image in the training set; the filter and the images are treated as vectors using the usual lexicographic scanning. This requirement ensures that the central correlations produced by the base filter remain unchanged.

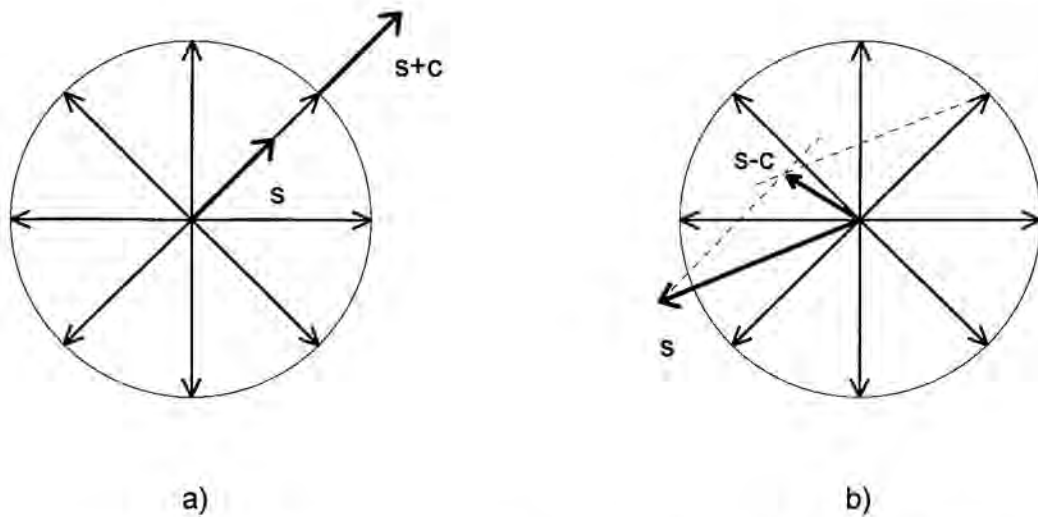
(II) The correlation between the correcting filter and the images produces a constant plane with a predefined real value that we will denote  $c$ . This condition is accomplished only in an approximate form by means of a Lagrange minimization process.

The correcting filter, is the same as that utilized in chapter 4. The output distribution obtained with the  $k$ th filter has the following two terms:

$$(h_b + e^{j\theta_k} h_c) \star x_i = h_b \star x_i + e^{j\theta_k} h_c \star x_i \quad (5.1)$$

where  $h_b$  is the base filter,  $h_c$  is the correcting filter,  $x_i$  is one of the images in the training set and the symbol  $\star$  means correlation. Letter  $j$  represents the imaginary unit and  $\theta_k$  indicates the phase of the  $k$ th complex value applied to the correcting filter.

The expression  $h_b \star x_i$  is a bidimensional distribution of complex values. Since  $h_c \star x_i$  is constant and real over the correlation plane -except at the center where it is null-, the expression  $e^{j\theta_k} h_c \star x_i$  is a constant complex plane of value  $ce^{j\theta_k}$ . The effect of the new filter on the correlation given by the base filter is therefore to increase the sidelobes pointing to directions close to that represented by the angle  $\theta_k$  and decrease the sidelobes pointing to directions close to  $\theta_k + \pi$  - Fig. 5.1-. By equally distributing the angles  $\theta_k$  on the unit circle -i.e.  $\theta_k = 2k\pi/n$  where  $n$  is the number of filters- we ensure that there will always be a filter whose correlation points close to the opposite direction to that of every sidelobe.



**Figure 5.1-** Effect of the correcting correlations on the height of the sidelobe. a) The sidelobe is increased by a correcting correlation pointing in the same direction. b) The sidelobe is decreased by a correlation pointing in a nearly opposite direction.

That filter will decrease the magnitude of the sidelobe. Suitable choice of the value of the constant plane  $c$ , the number of filters  $n$ , and the threshold ensures that no sidelobe is common to all binarized correlations. These considerations will be discussed in detail in section 5.2.

As before our analysis is devoted to the two-class problem. If the prespecified values for the magnitude of the correlation with classes **A** and **B** are  $p_0$  and  $p_1$  respectively, -supposing  $p_0 > p_1$  with no loss of generality-, the decision is taken by comparison with a threshold: if the correlation value at the center is higher than that of the threshold, the input pattern is associated with class **A** and vice versa. The necessary conditions that the multifilter procedure has to fulfill to prevent either a false alarm or a detection miss in a classification problem are as follows:

a) The classification procedure imposes a limit on the threshold value since the correct assignment of each class implies:

$$p_1^2 < \theta p_0^2 \quad (5.2)$$

where  $\theta$  is a value between zero and one that represents the threshold. The maximum value at the output plane is supposed to be the peak for the true class  $p_0^2$  although it may be different, for instance when a high sidelobe appears. The more restrictive condition represented by eq. 5.2 is needed to cover all cases.

b) The magnitude of the sidelobes are sometimes increased owing to the constant planes produced by the correcting filters. In those situations we must ensure that these increased sidelobes do not hide the peak corresponding to the true class, namely:

$$\theta(x+c)^2 < p_0^2 \quad (5.3)$$

where  $c$  is the magnitude of the complex plane given by the correcting filters and  $x$  the magnitude of the maximum sidelobe to be eliminated. The equation has



been written in the most unfavorable case, which appears when there is a sidelobe pointing in the same direction as the correcting plane -Fig. 5.1.a)-.

c) To avoid false alarms, we need that for every sidelobe exists at least one correcting filter being able to decrease its magnitude below the threshold value. As in point b) the equation is written in the worst case which now appears when the sidelobe points between two correcting correlations -Fig. 5.1.b)-. For such peaks we need:

$$(x - c \cos \frac{\pi}{n})^2 + c^2 \sin^2 \frac{\pi}{n} < \theta p_0^2 \quad (5.4)$$

where  $n$  is the number of correcting filters.

d) Finally, we must ensure that every sidelobe ranging from zero to  $x$  disappear after the processing with the threshold function. This fact is not always covered by eq. 5.4 which only accounts for the maximum sidelobe  $x$ . In particular when

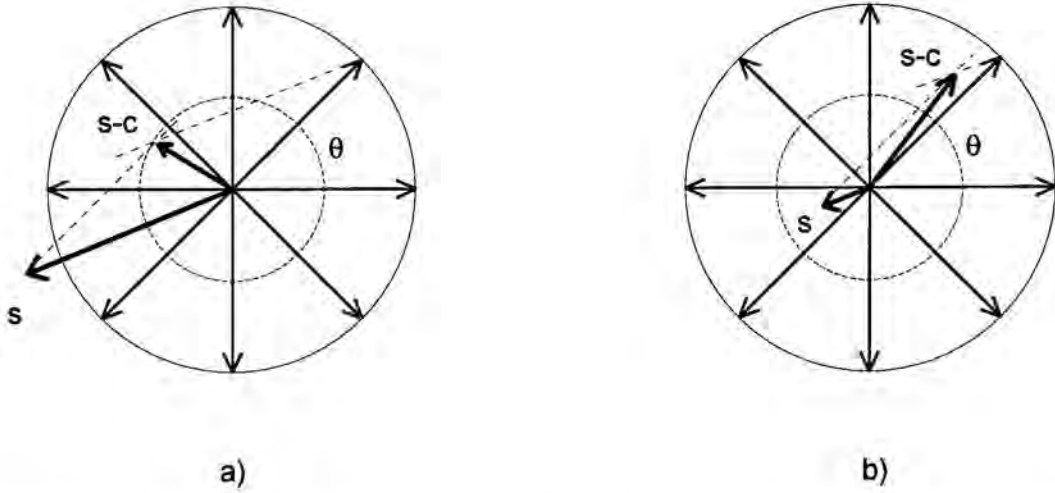
$$(x - c \cos \frac{\pi}{n})^2 + c^2 \sin^2 \frac{\pi}{n} < c^2 \quad (5.5)$$

small sidelobes, for example one with zero value, may surpass the threshold -Fig. 5.2-. After a little of algebra, the above equation can be written as:

$$c > \frac{x}{2 \cos \frac{\pi}{n}} \quad (5.6)$$

To avoid the appearance of these small sidelobes in the final output we need the additional constraint:





**Figure 5.2-** Sometimes the elimination of a high sidelobe (a) does not ensure the elimination of a small one (b).

$$c^2 < \theta p_0^2 \quad (5.7)$$

The method is intended to be applied to the general case in which the correlation plane is fully complex. When the sidelobes can only be positive or negative (their phases only 0 or  $\pi$ ), that is the real case, to increase or decrease them we only need two filters, one giving a constant plane of value  $Ce^{j0}=C$  and a second one giving  $Ce^{j\pi}=-C$ . However when we deal with complex correlations, the sidelobes have arbitrary phases, they can point in any direction, and so we need to use multiple filters. The equivalent situation to that of chapter four is the use of an infinite number of filters. In such a case, for every possible direction that a sidelobe may take, we have a correcting plane pointing in the opposite one. For example, the condition in eq. 5.6, becomes for  $n=\infty$ ,  $c>x/2$ , which was that obtained for the real case, although as we will show, good results can be obtained with a reduced filter database.

## 5.2 Optimal settings for the parameters.

The elimination of all sidelobes through the multiple correlation scheme proposed is guaranteed when the above conditions are fulfilled. For the potential user, the important parameters involved in these conditions are the highest sidelobe to be suppressed and the number of filters that one is willing to use. The other two variables, the value for the constant  $c$  and the threshold  $\theta$  are only internal parameters of the procedure, whose values have no special interest for the user and can be selected with total freedom. According to this, we derive in this section, the values for  $c$  and  $\theta$  that make compatible the conditions with values for  $x$  and  $n$  arbitrarily chosen. Obviously this is not always possible, for example when we try to eliminate a high sidelobe with a small number of filters, and so the minimum number of filters to eliminate a given sidelobe is also deduced here. To do so, we first make explicit the constraints that equations 5.2, 5.3, 5.4 and 5.7 impose over the constant  $c$  and which have to be used in subsequent steps.

### 5.2.1 Constraints over the value of constant $c$ .

For simplicity we will treat the usual case where  $p_0=1$  and  $p_1=0$ . In such a situation equation 5.2 automatically holds unless the threshold is zero -which is meaningless- and thus we have no longer this constraint.

a) The first restriction comes from equation 5.3, which in the general case is:

$$c_1 < -x \pm p_0 \theta^{-1/2} \quad (5.8)$$

and assuming  $p_0=1$  becomes:

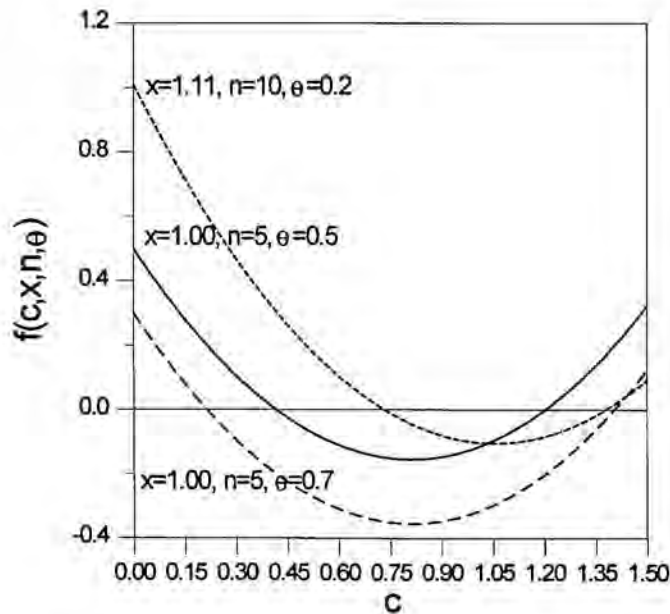
$$c_1 < -x \pm \theta^{-1/2} \Rightarrow c_1 < -x + \theta^{-1/2} \quad (5.9)$$

It should be noted that  $\theta$  is a dimensionless parameter, although in the following equations may apparently take different dimensions due to the simplification we made. In the latter equation, the minus sign has been discarded because  $c$  is always a positive quantity. The subscript indicates the different permitted values for  $c$  due to the different constraints.

b) The second restriction is posed by equation 5.4, which leads to:

$$f(c, x, n, \theta) = c^2 - 2cx \cos \frac{\pi}{n} + (x^2 - \theta) < 0 \quad (5.10)$$

In figure 5.3,  $f(c, x, n, \theta)$  has been plotted for several values of  $x$ ,  $n$  and  $\theta$ . It shows that the range of allowed values for  $c$ , namely those for which  $f(c, x, n, \theta) < 0$ , is the region comprised between the solutions of the second order equation  $f(c, x, n, \theta) = 0$ . By solving this equation we get:



*Figure 5.3- Range of permitted values for  $c$  arising from the necessity of avoiding false alarms.*

$$c_2^+ = x \cos \frac{\pi}{n} + [\theta - x^2 \sin^2 \frac{\pi}{n}]^{1/2} \quad (5.11)$$

and

$$c_2^- = x \cos \frac{\pi}{n} - [\theta - x^2 \sin^2 \frac{\pi}{n}]^{1/2} \quad (5.12)$$

and therefore the permitted range for  $c$  due to condition 5.4 is:

$$\max(0, c_2^-) < c_2 < c_2^+ \quad (5.13)$$

where the lower limit for  $c$  is zero when the minus solution is negative. Since we can write:

$$\theta - x^2 \sin^2 \frac{\pi}{n} = x^2 \cos^2 \frac{\pi}{n} + (\theta - x^2) \quad (5.14)$$

the value of  $c_2^-$  is positive when  $(\theta - x^2) < 0$ , a result that will be later used. This result can be interpreted as follows: if the sidelobe to be suppressed is higher than the threshold, i.e. if  $x^2 > \theta$ , the constant can not be zero and thus we have the lower limit given by eq. 5.12. Otherwise the sidelobe can be eliminated by the binarization alone, that is with a zero constant.

c) Finally, the constraint associated with eq. 5.7 can be expressed as:

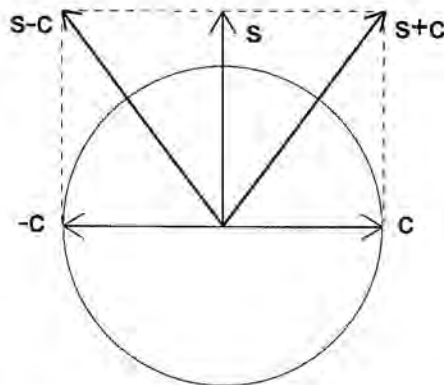
$$c_3 < \theta^{1/2} \quad (5.15)$$

### 5.2.2. Maximum eliminable sidelobe with a given number of filters.

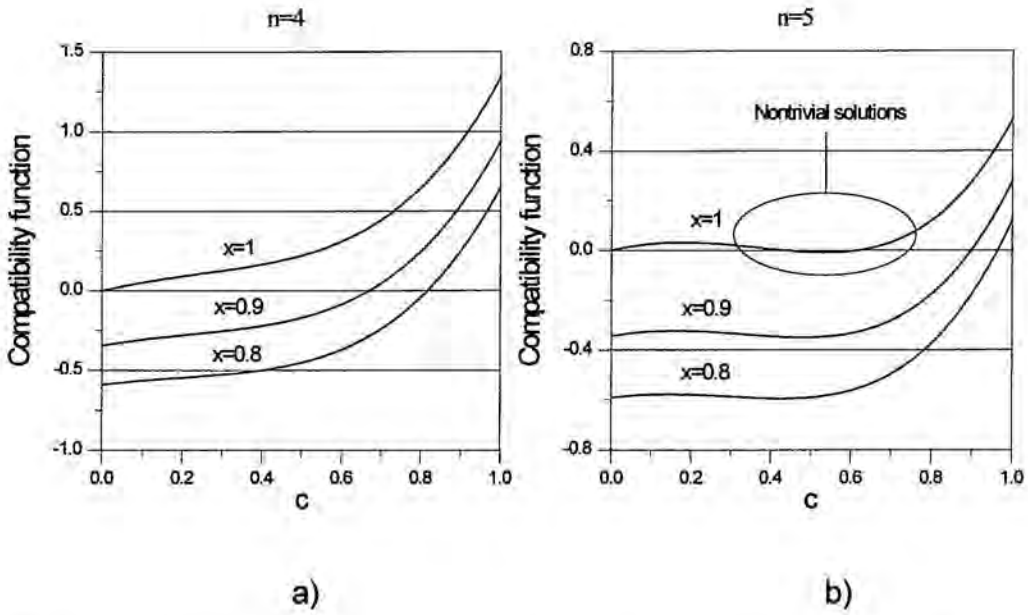
An important issue of the method is the maximum sidelobe that can be eliminated, which will depend on the number of filters used. The limit is due to the constraints of  $c$  that we have deduced in the preceding section. We obtain the expression for the value of the maximum eliminable sidelobe in what follows. Two cases must be treated separately:

- a) The number of filters is lower than five,  $n < 5$ .

When the number of filters is small, the method cannot correct sidelobes higher than the central peak because equations 5.3 and 5.4 become contradictory. This is due to that in the worst case, the sidelobes may be almost orthogonal to the correcting correlations and thus they have little effect in decreasing their height. To make them lower than a given threshold we need a large value for the constant  $c$ , which may produce the miss of the central correlation when the sidelobes are increased. A good example is  $n=2$  for which the method makes no sense since, in the worst case, it appears a sidelobe orthogonal to both correcting correlations as depicted in Fig. 5.4. Thus the correcting filters can only increase the height of the sidelobe, it must be suppressed by the threshold alone and therefore the sidelobes to be eliminated must be lower than the central peak.



*Figure 5.4- Two filters can only increase the magnitude of an orthogonal sidelobe.*



**Figure 5.5-** Compatibility function  $S(x,c,n)$ . a) If  $n=4$ , sidelobes equal to the central peak can not be eliminated. b)  $n=5$  represents the limit for which sidelobes higher than the central correlation can be suppressed.

In this situation the maximum sidelobe is suppressed when  $c=0$  and  $\theta=1$ , which makes the method unnecessary although it is still of use for sidelobes lower than the central peak as will be further commented. Eqs. 5.3 and 5.4 are mutually consistent when:

$$(x - c \cos \frac{\pi}{n})^2 + c^2 \sin^2 \frac{\pi}{n} < \theta < \frac{1}{(x+c)^2} \quad (5.16)$$

or equivalently while:

$$S(x,c,n) = (x+c)^4 - 2xc(x+c)^2(1 + \cos \frac{\pi}{n}) - 1 < 0 \quad (5.17)$$

Figure 5.5 represents the compatibility function  $S(x,c,n)$  for  $n=4$  and  $n=5$ . As can be observed, while there is a wide range of values for  $c$  allowing the

elimination of sidelobes lower than the central peak, sidelobes equal to it cannot be suppressed with  $n=4$  -Fig. 5.5.a)-, because there is no value for  $c$  making  $S(x=1, c, n=4)$  negative. The maximum sidelobe is thus  $x=1$ , the constant  $c$  must be zero and the method becomes unnecessary.  $n=5$  -Fig. 5.b)-, represents the limit in the appearance of nontrivial solutions.

b) The number of filters is higher or equals five,  $n \geq 5$ .

Equation 5.10 limits the height of the maximum sidelobe to be suppressed by the procedure since the discriminant of the equation must be positive because  $c$  is a real value:

$$-x^2 \sin^2 \frac{\pi}{n} + \theta > 0 \Rightarrow x_{M1} = \frac{\theta^{1/2}}{\sin \frac{\pi}{n}} \quad (5.18)$$

A second restriction in the height of the sidelobes arises from equations 5.9 and 5.12. On one hand, when the magnitude of the sidelobe to be eliminated increases, the lower limit for  $c_2$  -called  $c_2^-$ - also increases but on the contrary the upper limit for  $c_1$  decreases. The maximum sidelobe that is compatible with both constraints will be that for which  $c_2^- = c_1$ , namely:

$$x \cos \frac{\pi}{n} - [\theta - x^2 \sin^2 \frac{\pi}{n}]^{1/2} = -x + \theta^{-1/2} \quad (5.19)$$

whence after a few calculations we get:

$$x_{M2} = \frac{1}{2} \theta^{-1/2} \pm \left[ \frac{1}{4} \theta^{-1} - \frac{1}{2} \frac{(\theta^{-1} - \theta)}{(1 + \cos \frac{\pi}{n})} \right]^{1/2} \quad (5.20)$$



In this equation the minus sign must be discarded because it can be shown that for that solution the expression  $(\theta - x^2)$  becomes positive and therefore  $c_2^-$  is no longer a constraint over  $c - c_2^-$  becomes negative as deduced in eq. 5.14-.

On the other hand, applying the same argument to  $c_3$  and  $c_2^-$  the maximum sidelobe compatible with both constraints will be:

$$x \cos \frac{\pi}{n} - [\theta - x^2 \sin^2 \frac{\pi}{n}]^{1/2} = \theta^{1/2} \quad (5.21)$$

whence

$$x_{M_3} = 2 \theta^{1/2} \cos \frac{\pi}{n} \quad (5.22)$$

and the maximum sidelobe compatible with all the conditions can be written as:

$$x_{\max} = \min(x_{M_1}, x_{M_2}, x_{M_3}) \quad (5.23)$$

It is easy to determine the minimum between  $x_{M_1}$  and  $x_{M_3}$  :

$$\sin \frac{2\pi}{n} \leq 1 \Leftrightarrow 2 \cos \frac{\pi}{n} \sin \frac{\pi}{n} \leq 1 \Leftrightarrow 2 \cos \frac{\pi}{n} \leq \frac{1}{\sin \frac{\pi}{n}} \quad (5.24)$$

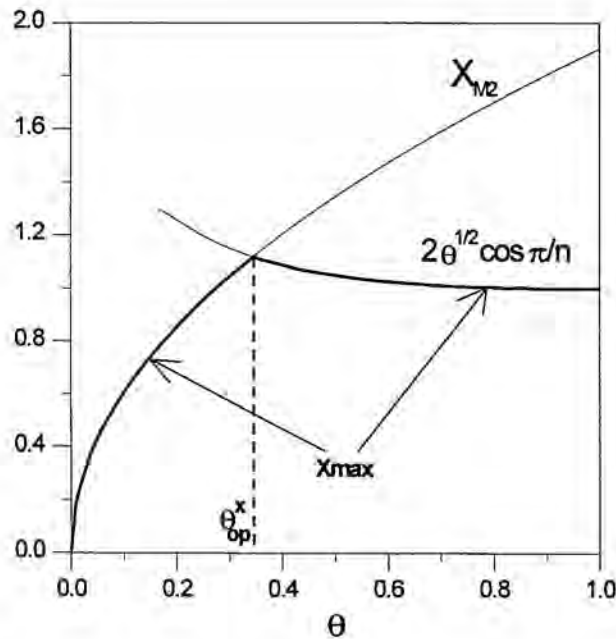
$$\min \left( \frac{\theta^{1/2}}{\sin \frac{\pi}{n}}, 2\theta^{1/2} \cos \frac{\pi}{n} \right) = 2\theta^{1/2} \cos \frac{\pi}{n} \quad (5.25)$$

and so we have:

$$x_{\max} = \min \left( 2\theta^{1/2} \cos \frac{\pi}{n}, \frac{1}{2}\theta^{-1/2} + \left[ \theta^{-1} \left( \frac{1}{4} - \frac{1}{2(1 + \cos \frac{\pi}{n})} \right) + \frac{\theta}{2(1 + \cos \frac{\pi}{n})} \right]^{1/2} \right) \quad (5.26)$$

As can be seen the maximum sidelobe depends on the number of filters used and on the threshold applied in the binarization. For a given number of filters, different thresholds will give different maximum sidelobes. It is possible to determine the threshold that enables us to eliminate the highest sidelobe allowed by the method with a prefixed number of filters.

In figure 5.6 we have represented the two expressions that determine the maximum sidelobe in eq. 5.25 as a function of the applied threshold. Notice that



*Figure 5.6- Determination of the threshold enabling the elimination of the maximum sidelobe.*

the squared root appearing in eq. 5.20 give rise to a condition over the threshold:

$$\frac{1}{4}\theta^{-1} - \frac{1}{2} \frac{(\theta^{-1} - \theta)}{(1 + \cos \frac{\pi}{n})} > 0 \Rightarrow \theta^2 > \frac{(1 - \cos \frac{\pi}{n})}{(1 + \cos \frac{\pi}{n})} \quad (5.27)$$

which is a lower limit that has to be taken into consideration in the election of this parameter. As can be observed, the optimum threshold -i.e. that which gives the highest sidelobe- corresponds to the intersection point between the two plots due to the increasing behaviour of the first constraint and the decreasing behaviour of the second. Therefore the optimum threshold will be reached when:

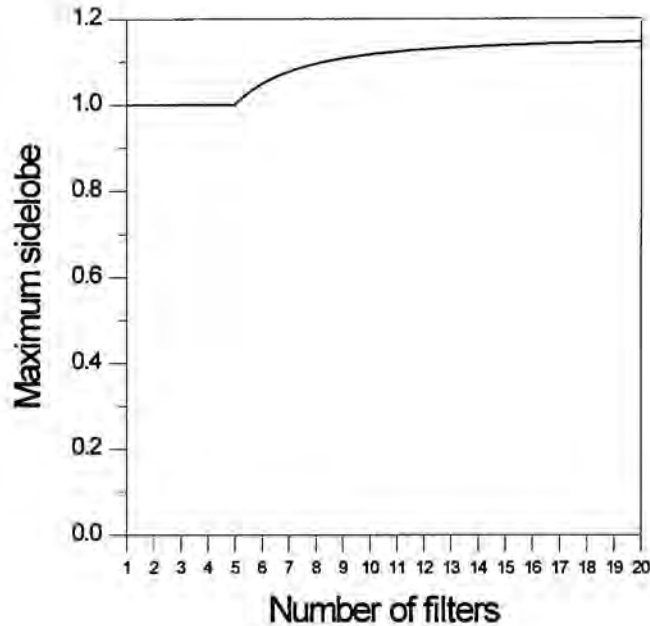
$$2\theta^{1/2} \cos \frac{\pi}{n} = \frac{1}{2}\theta^{-1/2} + \left[ \theta^{-1} \left( \frac{1}{4} - \frac{1}{2(1 + \cos \frac{\pi}{n})} \right) + \frac{\theta}{2(1 + \cos \frac{\pi}{n})} \right]^{1/2} \quad (5.28)$$

whence

$$\theta_{op}^x = \frac{2 \cos \frac{\pi}{n} (1 + \cos \frac{\pi}{n}) + \cos \frac{2\pi}{n}}{8 \cos^2 \frac{\pi}{n} (1 + \cos \frac{\pi}{n}) - 1} \quad (5.29)$$

and thus the maximum eliminable sidelobe is:

$$x_{\max}(n) = A(\theta_{op}^x)^{1/2} \quad (5.30)$$



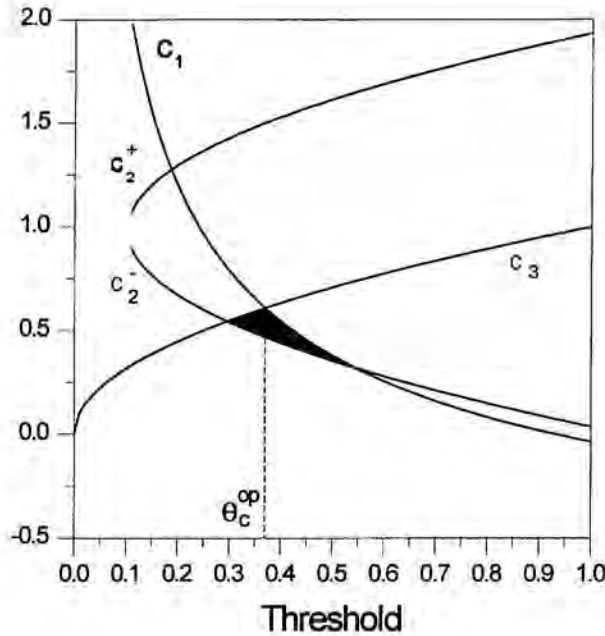
*Figure 5.7- Maximum sidelobe versus the number of filters used.*

Figure 5.7 represents this maximum sidelobe as a function of the number of filters used,  $n$ . For  $n \rightarrow \infty$  we have  $x_{max}(\infty) = 2/\sqrt{3}$  which is the value obtained for the real case.

Equation 5.29 can be interpreted in a more useful way for our purposes, as the expression which gives the minimum number of filters needed to eliminate a given height of sidelobe because it is an increasing function; if  $x$  represents the maximum sidelobe that can be suppressed with  $k$  filters we need at least  $k$  filters to eliminate  $x$ .

### 5.2.3 Optimal settings for constant $c$ and threshold $\theta$

In practice we will need a larger  $n$  than that given by eq. 5.29 because with such election of the threshold and the number of filters the possible choices for the constant  $c$  reduce to a single value, that is,  $c$  is completely determined. Because of the filter design, we have in actual correlations small fluctuations with respect to the desired plane which have not been considered until now. In



*Figure 5.8- Permitted range of variation for  $c$  (shaded area) as a function of the applied threshold ( $n=10, x=1,035$ ).*

order to allow this variation we need that a given sidelobe can be suppressed with different values of  $c$ , that is, we need a range of values for  $c$  being compatible with the constraints.

By selecting a number of filters larger than the minimum we have some freedom in the election of  $c$ . Furthermore, we are not restricted to use the threshold given by eq. 5.28 and we can select that which gives the maximum range of variation for the constant  $c$ .

For example, in Fig. 5.8 the allowed variation for  $c$  has been plotted versus the threshold for a problem in which, requiring six filters, ten filters has been used. Because of the monotonic behaviour of all the constraints of  $c$ , the maximum allowed variation is obtained when:

$$\theta^{1/2} = -x + \theta^{-1/2} \quad (5.31)$$

whence we obtain:

$$\theta_{op}^c = \frac{(2 + x^2) - x \sqrt{x^2 + 4}}{2} \quad (5.32)$$

Finally, the value for  $c$  that must be used in the filter design will be the medium value of the permitted range to account for positive and negative variations of the constant, namely:

$$c = \frac{1}{2} \left[ (\theta_{op}^c)^{1/2} + x \cos \frac{\pi}{n} - \left( \theta_{op}^c - x^2 \sin^2 \frac{\pi}{n} \right)^{1/2} \right] \quad (5.33)$$

The method allows the elimination of sidelobes higher than the central peak as commented before and as will be illustrated in the next section. However, this procedure may also be used when we have lower lateral peaks to obtain a more reliable process than the simple utilization of a threshold function. Large sidelobes imply large thresholds and thus a small variation of the peaks due to any degradation of the input image may easily cause either a detection miss or a false alarm. In these situations the multichannel procedure may be used to increase the range of variation for the peaks to obtain a more dependable detection.

On the other hand it should be pointed out that the the control of sidelobes is not the only desirable ability of a correlation procedure. When only this is considered, the procedure is in general overspecialized and gives little performance in other interesting quality criteria such as noise resistance. The filter design -the MSE-SDF filter- we are using as correcting filter, focuses only in minimizing the error in the shape of the correlation plane with respect to a constant plane and thus has not built-in capabilities to defeat noise. The problem is analogous to that found in the standard antisidelobe design, the MACE filter that is known to present a low signal-to-noise ratio. This analogy could be used to adapt the solutions proposed to increase the noise robustness of the MACE filter to the MSE-SDF filter. This solution would consist in trading-off antisidelobe power for noise robustness by introducing the noise induced peak variance as an additional term to be minimized jointly with the shape error. Such

approach is called optimum trade-off filter design and is known to give good results as commented in chapter 2.

### 5.3 Algorithm.

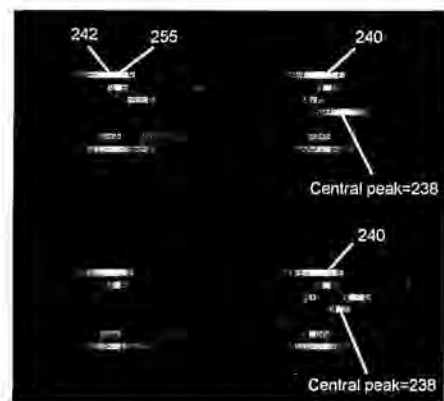
In order to summarize the method we are giving an algorithm to compute all the parameters envisaged by the procedure. In addition a practical example is analyzed step by step in this section to clarify the process.

Let us consider the four letters of Figure 5.9 as our training set. The values of the central correlations were chosen to be (0.0, 0.0) for **c** and **o**, (1.0, 0.0) for **e** and (-0.866, -0.5) for **s** in designing a composite filter. The magnitude of the response for the two target images is the same, namely one. With such values several sidelobes appear, the largest of which is 107% higher than the central correlations with the target images -Fig. 5.10-. To overcome the problem the following steps must be followed:

- Step 1: Select the magnitude of the maximum sidelobe to be eliminated. In our case this value is set to  $x=\sqrt{1.07}=1.035$ .



*Figure 5.9- Training set. The specified values are (0,0) for **c** and **o**, (1,0) for **e** and (-0.86,-0.5) for **s**.*



*Figure 5.10- Intensity of the correlation between the four letters of Figure 5.9 and a composite filter.*



- Step 2: Determine the minimum number of filters required to suppress the sidelobe by means of equation 5.29, i.e.:

$$x_{\max}(n) = A(\theta_{op}^x)^{1/2}$$

In the example  $n=6$  gives a maximum sidelobe of  $x=1.048$  and  $n=5$  gives  $x=1.000$ . Therefore the minimum number of filters is six.

- Step 3: Choose a value for  $n$  larger than the minimum. The larger the number of filters used the larger the permitted range of variation for  $c$  -and thus the more reliable the method is- but also the more complex the procedure becomes. We have chosen ten filters for our problem.

- Step 4: Select the threshold by means of equation 5.31:

$$\theta_{op}^c = \frac{(2 + x^2) - x \sqrt{x^2 + 4}}{2}$$

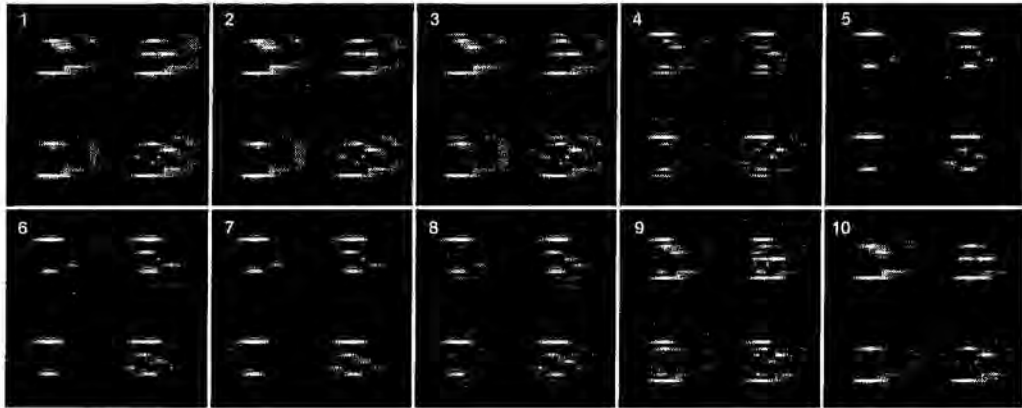
This election will permit a maximum variation for the constant. In the example the threshold must be  $\theta_{op}^c=0.370$ .

- Step 5: Select the constant  $c$  using equation 5.32:

$$c = \frac{1}{2} \left[ (\theta_{op}^c)^{1/2} + x \cos \frac{\pi}{n} - \left( \theta_{op}^c - x^2 \sin^2 \frac{\pi}{n} \right)^{1/2} \right]$$

In our case we have  $c=0.538$  with an allowed range of  $\Delta c=\pm 0.07$ , which represents a possible variation of 13% above and below the specified value.

The intensity of the correlations between the ten filters and the scene containing the four letters is shown in Fig. 5.11. By binarizing these correlations with the threshold calculated in step 4, figure 5.12 is obtained. The final result after multiplying pixel by pixel the ten binarized correlations is depicted in

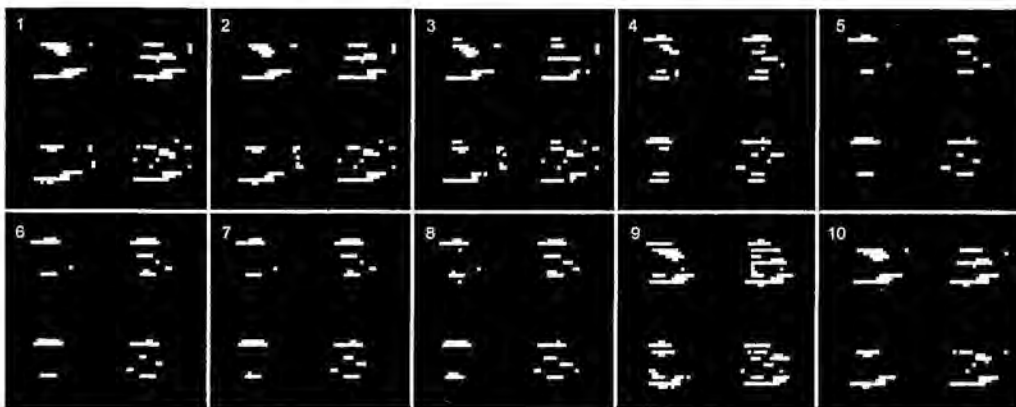


*Figure 5.11- Intensity of the correlation between the four letters of Figure 5.9 and the set of ten filters obtained by the method.*

Figure 5.13 where a perfect detection of the target patterns is now possible in contrast to the original correlations -Fig. 5.10-.

In conclusion, the appearance of lateral peaks is one of the difficulties encountered in certain type of filters. These sidelobes may be even higher than the central correlation thus turning the filter unable to detect the target patterns. In such cases a single threshold is not enough to eliminate them without also hiding the central peak. Even if this is not the case a high threshold may be necessary lowering the reliability of the recognition process.

We have presented a new method to handle this problem which is capable



*Figure 5.12- Binarization of the correlations of Figure 5.11 with  $\theta=0.370$ .*



*Figure 5.13- Pixel by pixel multiplication of the ten binarized correlations of Figure 5.12. A perfect detection of letters e and s is now possible.*

of eliminating sidelobes higher than the central peak. The process consist of a multichannel correlation followed by a binarization and a pointwise multiplication of the outputs and can be used with filters with no symmetry constraints.

A complete mathematical description of the process has been developed and an algorithm has been reported to design the filters. Computer simulations show the validity of the approach.

## **Chapter six. Arbitrarily constrained filters**

### **Introduction.**

We have seen how the multifilter setups provide the means to overcome the intrinsic limitations of single filter correlations. The use of several filters is therefore advantageous in some problems and absolutely necessary in some others, as when we deal with three-dimensional objects and multiple invariances are required. This is the theory but what about practice?. In other words, multichannel correlators are necessary but are they feasible? How can this systems be built and what problems do we encounter in practice?. Of course, the answer to these questions is neither easy nor general. The problems, as the filters, are multiple. Starting from an abstract level, we need to consider the architectural and organizing issues. For example, the SPOTR correlator mentioned in Chapter one is divided into three functional modules: preprocessing system, the correlator itself and the postprocessing system. Each of these subsystems need their own algorithms -see for instance [Bau93]- and control devices. Focusing our attention on the correlator we have in turn to handle organization issues, perhaps adopting a hierarchical decomposition of the identification problem as proposed by Casasent [Cas94]. Once this has been decided, the filter database has to be designed and ordered, determining which filter should start the process and which ones should be used next based on the information given by the first. The possibilities proposed in [Jar86] and [Cal94] may be of help. There are also technical problems such as the implementation of the filters and the procedure to change them with rapidity and accuracy. A mechanical system could be utilized for a small number of filters but this would probably lead to stability and alignment difficulties. A better choice could be the spatially multiplexed setup proposed in [Cas94] where a matrix of laser diodes select the proper filter.

However, for a large database the use of spatial light modulators to render the filters is surely the best solution. This in turn raises new problems in filter design because such devices frequently present a finite modulation capability, namely that the possible values of amplitude and phase that can be codified are limited. The theoretical designs and procedures presented in the preceding chapters are based on the assumption of unrestricted modulation and consequently they must be adapted to the new constraints.

This field have deserved little attention until recently and therefore it is not as mature as its full modulation counterpart. Juday [Jud93] and recently Laude and Réfrégier [Lau94] have addressed the problem of the optimum projection of single image filters to the domain allowed by a modulator. For the single image designs and for a wide class of metrics -the quality criteria according to the terminology of these authors- the optimal solution is more or less as follows. First compute the optimal filter assuming there is no modulation restrictions following the procedures summarized in chapter two. Second, provided the available complex values are known, for every component of this filter find the realizable value placed at the minimum Euclidean distance. This ensures the optimality among all constrained filters.

For SDF filters with several images in the training set, the problem is worse, since a simple projection of the values of the fully complex filter will, in most cases, dramatically modify the desired outputs. Therefore the constraints imposed by the filter plane modulator must be considered in the design procedure. Notice the difference with single image filters. Here we are not trying to determine the optimal among several solutions but we are in a previous stage: we find problems even in fulfilling the SDF conditions. Optimality will be a later concern.

The majority of the efforts has been directed at obtaining phase-only and binary phase-only SDF filters, although the general case of arbitrary constraints have also been studied. The first attempt to design a Phase-Only Synthetic Discriminant filter was reported by Horner and Gianino [Hor85b] whose solution consisted simply of using the phase of a conventional composite filter. Although the first tests gave good results, the approach is not appropriate because the SDF constraints are no longer met [Cas86]. Since then a variety of recipes have appeared in the literature but none of them seems to give the ultimate answer:

Kallman's algorithm [Kal86] is computationally expensive and gives little control over the correlation peaks -according to [Jar89]-. The procedures proposed by Jared and Ennis [Jar89] and Bahri and Kumar [Bah89] limit the number of possible solutions by supposing the filter to be a linear combination of the training images, thus affecting the probability of convergence. Jared and Ennis algorithm has also been applied to design arbitrarily constrained filters in [Carl92]. The entropy-optimized filter proposed by Mahlab and Shamir [Mahl89] uses the simulated annealing algorithm, inheriting its drawbacks: heuristic selection of several parameters -initial temperature, number of iterations until thermal equilibrium, etc.- and a high computational load -Réfrégier in [Réf90b] reports 7h. 30 m. on a VAX8200 computer for 10 images of 64x64 pixels-.

In this chapter a new algorithm for computing constrained SDF filters is proposed. It is based on a new filter design that we call minimum Euclidean distance synthetic discriminant function (MED-SDF) filter, which enables us to obtain the closest SDF filter to a given non-SDF filter in the sense of Euclidean distance. This design is then used in an iterative algorithm which leads to a solution in a few steps.

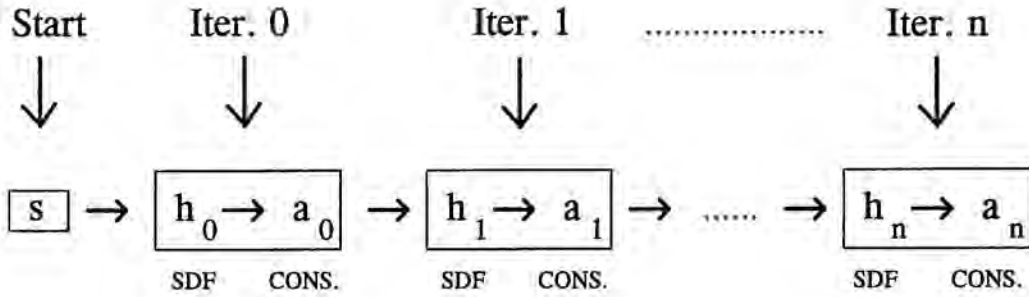
## 6.1 The structure of the algorithm.

We designed an iterative algorithm with a structure very similar to that used in the Successive Forcing Algorithm [Bah89], that is an algorithm in which, at each iteration, the filter is forced to fulfill the conditions for the central correlations and subsequently to take values on the allowed domain - Fig. 6.1-.

This technique of successive projections is widely used as a basis of different algorithms in filter synthesis or in image restoration, see for example [Ros91] and [You82]. The answer to:

I) How to project the SDF filter  $h_k$  onto the domain allowed by the modulator to obtain filter  $a_k$ , and





*Figure 6.1- Structure of the algorithm.*

II) How to force the constrained filter  $a_k$  to give the desired central correlations, that is to obtain the SDF filter  $h_{k+1}$

will completely shape the algorithm.

Among all the possibilities, we are interested in those leading to convergent procedures. Since we are looking for the SDF filters that take values on the specified subset of the complex unit circle, such convergence should be expressed as:

$$\sum_{i=1}^N |h_{k+1}^i - a_{k+1}^i|^2 < \sum_{i=1}^N |h_k^i - a_k^i|^2 \quad (6.1)$$

where  $h$  and  $a$  are the SDF and the constrained filters respectively. The subscripts  $k$  and  $k+1$  indicate the iteration and the superscript  $i$  the pixel; finally  $N$  represents the number of pixels of  $h$  and  $a$ . Equation 6.1 demands that some measure of similarity between the SDF and the constrained filter be a decreasing quantity with the number of iterations. Although it has been written using the Euclidean distance as a measure of similarity, this is not the only possibility and other metrics might be of use. The solution will be reached when the distance between the two filters drop under some limit depending on the desired precision for the correlation values.



However, imposition of eq. 6.1 implies that one solution always exists and that such solution can be reached by means of an algorithm with the chosen structure. We are not able to ensure this so we had to use the less ambitious condition:

$$\sum_{i=1}^N |h_{k+1}^i - a_{k+1}^i|^2 \leq \sum_{i=1}^N |h_k^i - a_k^i|^2 \quad (6.2)$$

which is very similar to the former but where the equality sometimes may hold.

By splitting the latter inequality as:

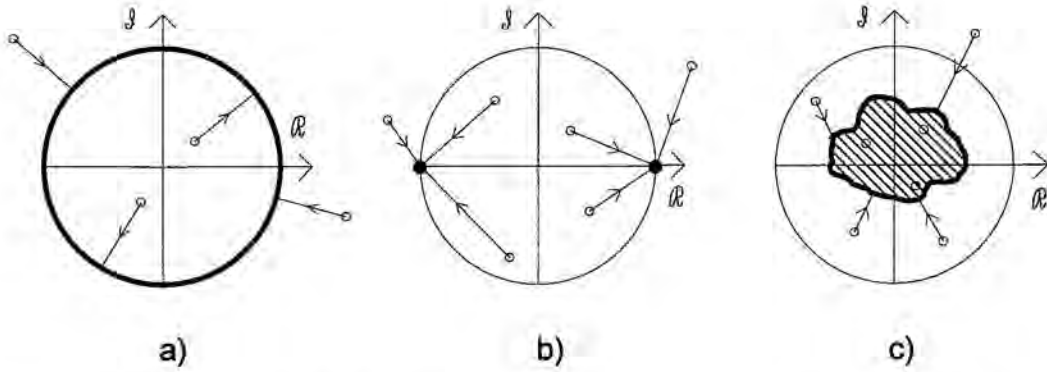
$$\sum_{i=1}^N |h_{k+1}^i - a_{k+1}^i|^2 \leq \sum_{i=1}^N |h_{k+1}^i - a_k^i|^2 \leq \sum_{i=1}^N |h_k^i - a_k^i|^2 \quad (6.3)$$

which obviously imply eq. 6.2, we can answer the two questions stated above in an unique way.

The first inequality:

$$\sum_{i=1}^N |h_{k+1}^i - a_{k+1}^i|^2 \leq \sum_{i=1}^N |h_{k+1}^i - a_k^i|^2 \quad (6.4)$$

is used to project the SDF filter onto the allowed domain of the complex plane. At this point of the process we know everything except the filter  $a_{k+1}$ . Eq. 6.4 compares the distance of  $a_k$  and  $a_{k+1}$  with respect to the SDF filter at iteration  $k+1$ ,  $h_{k+1}$ . Since  $a_k$  is a constrained filter if we choose  $a_{k+1}$  as the constrained filter that makes minimum:



**Figure 6.2-** Projection of the filter values onto the allowed domain for a) a phase-only filter, b) a binary phase-only filter and c) for an arbitrary domain.

$$E(\mathbf{a}) = \sum_{i=1}^N |h_{k+1}^i - a^i|^2 \quad (6.5)$$

eq. 6.4 is automatically satisfied. Eq. 6.5 is a sum of positive and independent terms and therefore the minimum is reached by minimizing each addend. Since the coding domain is assumed to be known, the process merely reduces to a search of the closest domain value to each component of  $\mathbf{h}_{k+1}$ , a process sketched in Figure 6.2. For example, for a phase-only filter,  $\mathbf{a}_{k+1}$  is obtained by extracting the phase of the SDF filter  $\mathbf{h}_{k+1}$ . It is worth pointing out that this process is used to project in an optimum way, a single image filter [Jud93][Lau94].

Equivalently, the second inequality

$$\sum_{i=1}^N |h_{k+1}^i - a_k^i|^2 \leq \sum_{i=1}^N |h_k^i - a_k^i|^2 \quad (6.6)$$

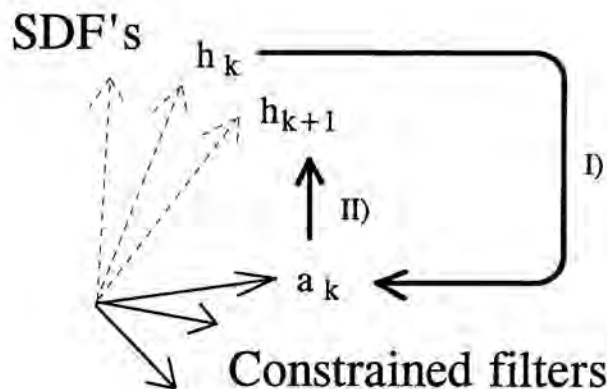
tells us how to derive the SDF filter at the next iteration,  $\mathbf{h}_{k+1}$  from the constrained filter at the previous iteration  $\mathbf{a}_k$ . At this point we know everything except  $\mathbf{h}_{k+1}$ . Eq. 6.6 compares the distance of  $\mathbf{h}_k$  and  $\mathbf{h}_{k+1}$  with respect to the

constrained filter at iteration  $k$ ,  $\mathbf{a}_k$ . Since  $\mathbf{h}_k$  is an SDF filter, by selecting  $\mathbf{h}_{k+1}$  as the SDF filter which makes minimum:

$$E(\mathbf{h}) = \sum_{i=1}^N |h^i - a_k^i|^2 \quad (6.7)$$

the inequality in eq. 6.6 is automatically met. The process for obtaining such a filter is not as evident as before because now the terms of the sum are not independent; they are linked together by the conditions imposed on the correlations with the training images. The appropriate conditions and the resulting expression for the MED-SDF filter will be given in the next section.

Figure 6.3 summarizes the information contained in this section. Starting from the SDF filter at iteration  $k$ ,  $\mathbf{h}_k$ , we obtain the constrained filter  $\mathbf{a}_k$  by looking, among all possible constrained filters, for the closest to  $\mathbf{h}_k$ . This filter is forced to be an SDF design -to obtain  $\mathbf{h}_{k+1}$ - by looking, among all possible SDF filters for the closest to  $\mathbf{a}_k$ . Notice that  $\mathbf{h}_{k+1}$  is not necessarily equal to the original filter  $\mathbf{h}_k$  and when they coincide the algorithm stops. This issue will be discussed in section 6.3. The latter step is carried out by computing the MED-SDF filter.



*Figure 6.3- Sketch of the two successive projections that form an iteration of the algorithm.*

## 6.2 The minimum Euclidean distance synthetic discriminant function (MED-SDF).

Let us suppose we have a correlation filter  $a$  that we wish to modify to give some prespecified values for the central correlations with  $M$  images. The shape of the filter is important and so we want to change it as little as possible. The question as to which filter enables us to obtain the desired correlations by preserving the original filter  $a$  as far as possible is answered in this section. We call this design minimum Euclidean distance synthetic discriminant function (MED-SDF). The problem can be stated in the following terms:

Let  $\mathbf{x}_1, \dots, \mathbf{x}_M$  denote the Fourier transforms of the  $M$  images of  $N$  components for which we wish to obtain the values  $c_1, \dots, c_M$  at the center of the correlation plane. Let  $\mathbf{a}$  be the filter we need to modify. We are looking for the filter  $\mathbf{h}$  so that

$$\mathbf{h}^+ \mathbf{X} = \mathbf{c}^t \quad (6.8)$$

where

$$E(\mathbf{h}) = \sum_{i=1}^N |h^i - a^i|^2 \quad (6.9)$$

is a minimum.

In the above expressions  $\mathbf{X}$  is the  $N \times M$  matrix whose columns are the images  $\mathbf{x}_i$ ,  $\mathbf{c}$  is the column vector of  $M$  components containing the values  $c_i$  and the superscripts  $t$  and  $+$  mean transpose and conjugate transpose. Finally,  $h^i$  and  $a^i$  represent the component number  $i$  of  $\mathbf{h}$  and  $\mathbf{a}$  respectively. Equation 6.9 can be rewritten using vector notation in the following way:

$$\begin{aligned}
E(\mathbf{h}) &= \sum_{i=1}^N [(h^i - a^i)^* (h^i - a^i)] = \sum_{i=1}^N [(h^i)^* h^i - (h^i)^* a^i - (a^i)^* h^i + (a^i)^* a^i] = \\
&= \mathbf{h}^+ \mathbf{h} - \mathbf{h}^+ \mathbf{a} - \mathbf{a}^+ \mathbf{h} + \mathbf{a}^+ \mathbf{a}
\end{aligned} \tag{6.10}$$

By making explicit the real and imaginary parts of all the quantities, namely  $\mathbf{h} = \mathbf{h}_R + j\mathbf{h}_I$ ,  $\mathbf{a} = \mathbf{a}_R + j\mathbf{a}_I$ ,  $\mathbf{X} = \mathbf{X}_R + j\mathbf{X}_I$  and  $\mathbf{c} = \mathbf{c}_R + j\mathbf{c}_I$  where  $j$  is the imaginary unit, eqs. 6.8 and 6.10 become:

$$\begin{aligned}
\mathbf{h}_R^t \mathbf{X}_R + \mathbf{h}_I^t \mathbf{X}_I &= \mathbf{c}_R^t \\
\mathbf{h}_R^t \mathbf{X}_I - \mathbf{h}_I^t \mathbf{X}_R &= \mathbf{c}_I^t
\end{aligned} \tag{6.11}$$

and

$$E(\mathbf{h}_R, \mathbf{h}_I) = \mathbf{h}_R^t \mathbf{h}_R + \mathbf{h}_I^t \mathbf{h}_I - 2\mathbf{h}_R^t \mathbf{a}_R - 2\mathbf{h}_I^t \mathbf{a}_I + \mathbf{a}_R^t \mathbf{a}_R + \mathbf{a}_I^t \mathbf{a}_I \tag{6.12}$$

The solution can be found by setting the gradients of the Lagrange function  $L(\mathbf{h}_R, \mathbf{h}_I)$  to zero with respect to the filter components, where:

$$\begin{aligned}
L(\mathbf{h}_R, \mathbf{h}_I) &= \mathbf{h}_R^t \mathbf{h}_R + \mathbf{h}_I^t \mathbf{h}_I - 2\mathbf{h}_R^t \mathbf{a}_R - 2\mathbf{h}_I^t \mathbf{a}_I + \mathbf{a}_R^t \mathbf{a}_R + \mathbf{a}_I^t \mathbf{a}_I - \\
&\quad - 2(\mathbf{h}_R^t \mathbf{X}_R + \mathbf{h}_I^t \mathbf{X}_I - \mathbf{c}_R^t) \mathbf{u} - 2(\mathbf{h}_R^t \mathbf{X}_I - \mathbf{h}_I^t \mathbf{X}_R - \mathbf{c}_I^t) \mathbf{v}
\end{aligned} \tag{6.13}$$

and  $\mathbf{u}$  and  $\mathbf{v}$  are  $M$ -dimensional column vectors containing the Lagrange multipliers. Such a solution can be written in vector notation as:

$$\mathbf{h} = \mathbf{a} + \mathbf{X}\mathbf{w}^* \tag{6.14}$$

with  $\mathbf{w} = \mathbf{u} + j\mathbf{v}$ . Substitution of eq. 6.14 into eq. 6.8 leads to:

$$(\mathbf{a}^+ + \mathbf{w}^t \mathbf{X}^+) \mathbf{X} = \mathbf{c}^t \tag{6.15}$$

whence

$$\mathbf{w}^* = (\mathbf{X}^+ \mathbf{X})^{-1} (\mathbf{c}^* - \mathbf{X}^+ \mathbf{a}) \quad (6.16)$$

and substituting into eq. 6.14 we finally obtain:

$$\mathbf{h} = \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^* + [\mathbf{I}_N - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+] \mathbf{a} = \mathbf{Comp} + \mathbf{Pa} \quad (6.17)$$

where  $\mathbf{I}_N$  is the  $N \times N$  identity matrix. This equation can be rewritten as:

$$\mathbf{h} = \mathbf{a} + \Delta \mathbf{a} = \mathbf{a} + \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} (\mathbf{c}^* - \mathbf{X}^+ \mathbf{a}) \quad (6.18)$$

that is, the modification of filter  $\mathbf{a}$  is a composite filter which complements the central correlations in the exact amount needed. Since the composite filter is, among the SDF's, that with minimum modulus -it minimizes  $\mathbf{h}^+ \mathbf{h}$ -, it changes as little as possible the original filter  $\mathbf{a}$ . The expression in eq. 6.17 admits a potentially useful interpretation. When  $\mathbf{a}$  is an arbitrarily chosen vector, eq. 6.17 represents the most general solution to the SDF problem -see chapter 2- where the term  $\mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{c}^*$  is the classical composite filter and  $\mathbf{I}_N - \mathbf{X}(\mathbf{X}^+ \mathbf{X})^{-1} \mathbf{X}^+$  is the projection operator over the subspace spanned by the  $N-M$  orthogonal vectors to the training images. Every SDF design can be expressed in the above form by properly choosing the vector  $\mathbf{a}$ , which can now be interpreted as the filter to which  $\mathbf{h}$  most approximates thus establishing an interesting link between SDF and non-SDF filters.

### 6.3 The algorithm.

The whole process to compute constrained SDF filters is depicted in Fig. 6.4 and can be sketched as follows:

- Step 1: Choose an initial vector  $\mathbf{a}_0$  of  $N$  components.

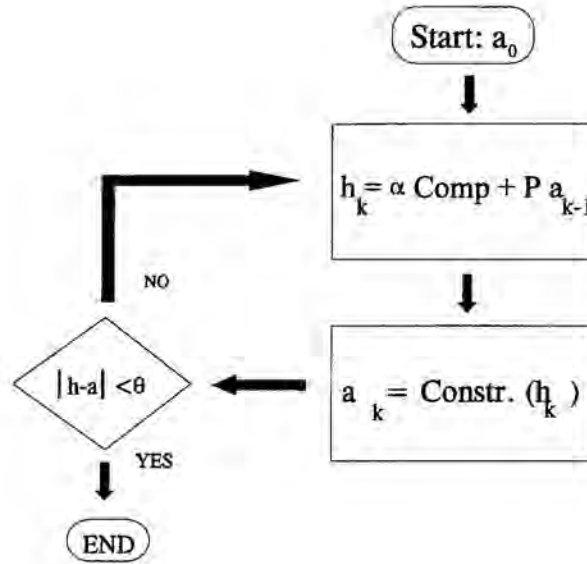


Figure 6.4- Block diagram of the proposed algorithm.

- Step 2: Start iteration  $k$ : compute

$$\mathbf{h}_k = \alpha \text{Comp} + P \mathbf{a}_{k-1} \quad (6.19)$$

by means of equation 6.17, where  $\alpha$  is a scaling constant that will be later justified.

- Step 3: Compute for  $i=1$  to  $N$

$$\mathbf{a}_k^i = \text{ArgMin}_{\forall s \in D} (|\mathbf{h}_k^i - s|^2) \quad (6.20)$$

where  $D$  represents the coding domain.

- Step 4: If the difference between the constrained and the SDF filter is small or if the algorithm stops its convergence, i.e. if



$$E_k = \sum_{i=1}^N |h_k^i - a_k^i|^2 < \theta_f \quad (6.21)$$

or

$$|E_k - E_{k-1}| < \theta_m \quad (6.22)$$

then exit;  $\theta_f$  and  $\theta_m$  are arbitrarily chosen small numbers.

- Step 5: If the condition in step 4 is not satisfied, finish iteration  $k$  by going to step 2.

The computation of the MED-SDF filter is carried out in step 2. It involves a matrix vector multiplication to project the constrained filter onto the orthogonal subspace to the training images and the addition of the resulting vector to the composite filter. Note that both this filter and the projection matrix are fixed and can be precomputed and stored. However, the projection operator is a matrix of  $N \times N$  components and would require a huge amount of memory -for 128x128 images it needs over 1 Gigabyte-. A good compromise between memory requirements and computation complexity is to precalculate only  $S = X(X^+X)^{-1}$ , which is a  $N \times M$  matrix -for 20 images of 128x128 pixels it needs about 1 Megabyte-. Then the computation is completed at each iteration by multiplying the stored matrix by the  $M$ -dimensional vector  $X^+a$  and by adding the result to the  $N$ -dimensional vector  $a$ :

$$\left[ I - X(X^+X)^{-1} X^+ \right] a = a - X(X^+X)^{-1} X^+ a = a - S(X^+ a) \quad (6.23)$$

The scaling constant in step 2 provides an additional degree of freedom. It accounts for the fact that the specified values for the central correlations can be rescaled to obtain a better matching between the computed filter and the values

available in the coding domain. We wish a bright correlation spot with the target images and a dim one with the nontarget patterns but the exact value to be imposed at the center of the correlation plane to achieve this, depends a good deal on the characteristics of the filter plane modulator. For the sake of clarity, we give the derivation of the scaling constant leading to a minimum distance between  $h_{k+1}$  and  $a_k$  in a separated appendix. Note that we have two additional sources of degrees of freedom:

- The phases of the central correlations with the  $M$  training images can be used to further minimize the difference between the SDF filter  $h_{k+1}$  and the constrained filter  $a_k$ . The same procedure used in phase optimization for previous SDF designs can be used [Vij88][Réf90c].

- The SDF filter  $h_{k+1}$  can be rescaled again by a complex constant to obtain  $a_{k+1}$  with minimal error as proposed in [Jud93] for single image filters.

We did not exploit these two possibilities in order not to overly complicate the algorithm. Although the nondivergent behavior of the algorithm will not be affected by not performing these additional operations the probability of convergence is lower.

The process stops only when:

- a) One solution is reached; the constrained filter  $a_k$  fulfills the SDF conditions, i.e.

$$a_k = \text{Comp} + P\mathbf{v} \quad (6.24)$$

where  $\mathbf{v}$  is some  $N$ -dimensional vector. Filter  $h_{k+1}$  is then

$$h_{k+1} = \text{Comp} + P(\text{Comp} + P\mathbf{v}) = \text{Comp} + P(\text{Comp}) + P^2\mathbf{v} = \text{Comp} + \mathbf{0} + P\mathbf{v} = a_k \quad (6.25)$$

and the output is constant from this point on. Eq. 6.25 has been derived using that  $P$  is a projection operator and therefore  $P^2=P$  and that the filter  $Comp$  is a linear combination of the training images and thus its projection  $P(Comp)$  is null.

b) If  $a_k$  is not an SDF filter but can be written as:

$$a_k = a_{k-1} + v \quad (6.26)$$

where  $v$  belongs to the kernel of the projection matrix, i.e.  $v \in \text{Ker}(P)$ . Filter  $h_{k+1}$  is then

$$h_{k+1} = Comp + Pa_k = Comp + P(a_{k-1} + v) = Comp + Pa_{k-1} = h_k \quad (6.27)$$

whence

$$\begin{aligned} a_k &= a_{k+1} = a_{k+2} = \dots \\ h_k &= h_{k+1} = h_{k+2} = \dots \end{aligned} \quad (6.28)$$

and the distance between the SDF and the constrained filter remains constant. This possibility is unlikely since the dimension of the kernel of the projection matrix is  $M$ , the number of training images, which is in general much smaller than the dimension of the space, the bandwidth product  $N$ .

c) If  $a_k$  is not an SDF filter but

$$\begin{aligned} h_k &\neq h_{k+1} \\ a_k &= a_{k+1} \end{aligned} \quad (6.29)$$

then

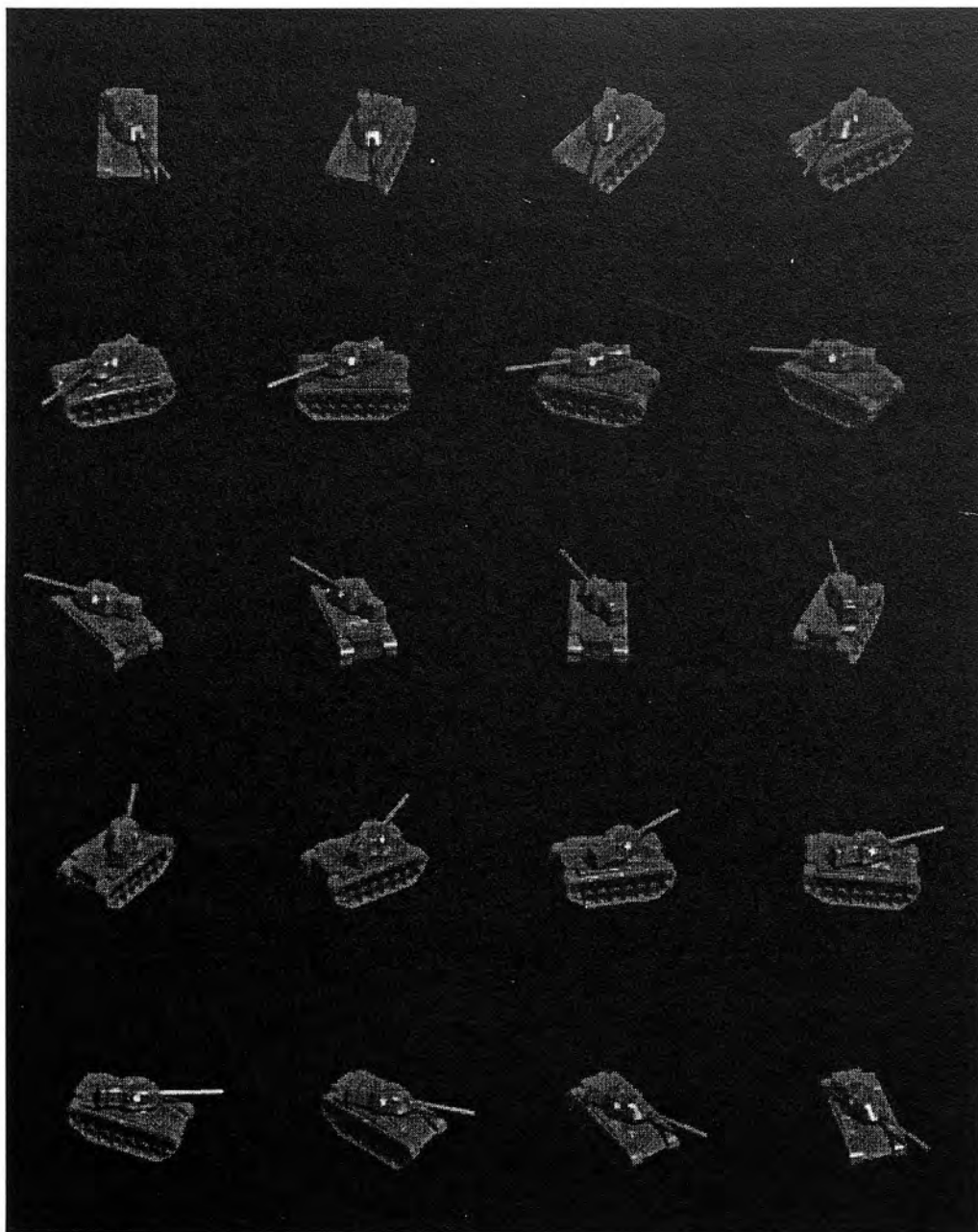
$$\begin{aligned} a_k &= a_{k+1} = a_{k+2} = \dots \\ h_{k+1} &= h_{k+2} = \dots \end{aligned} \quad (6.30)$$

i.e. when the projection of two consecutive and different SDF filters is the same, the algorithm stops its convergence. This possibility is difficult to analyze and depends on the coding domain. It represents the intuitive notion that the smaller the number of coding values the smaller the probability to find a solution. When there are no restrictions and the entire complex plane is available, the SDF and the constrained filters are always equal and equation 6.29 is self-contradictory and never holds. When only one coding value is allowed, all the constrained filters are the same and the process stops at the first iteration. No solution is possible. Binary modulators permit the coding of  $2^N$  different filters, which for 128x128 images is over  $10^{4900}$ . In spite of this seemingly large number, since two consecutive SDF filters may be very similar, specially when we are near the solution, their binarization may be equal with relative ease. Although we found this problem with binary phase-only filters we show in the next section that the algorithm can still produce usable filters.

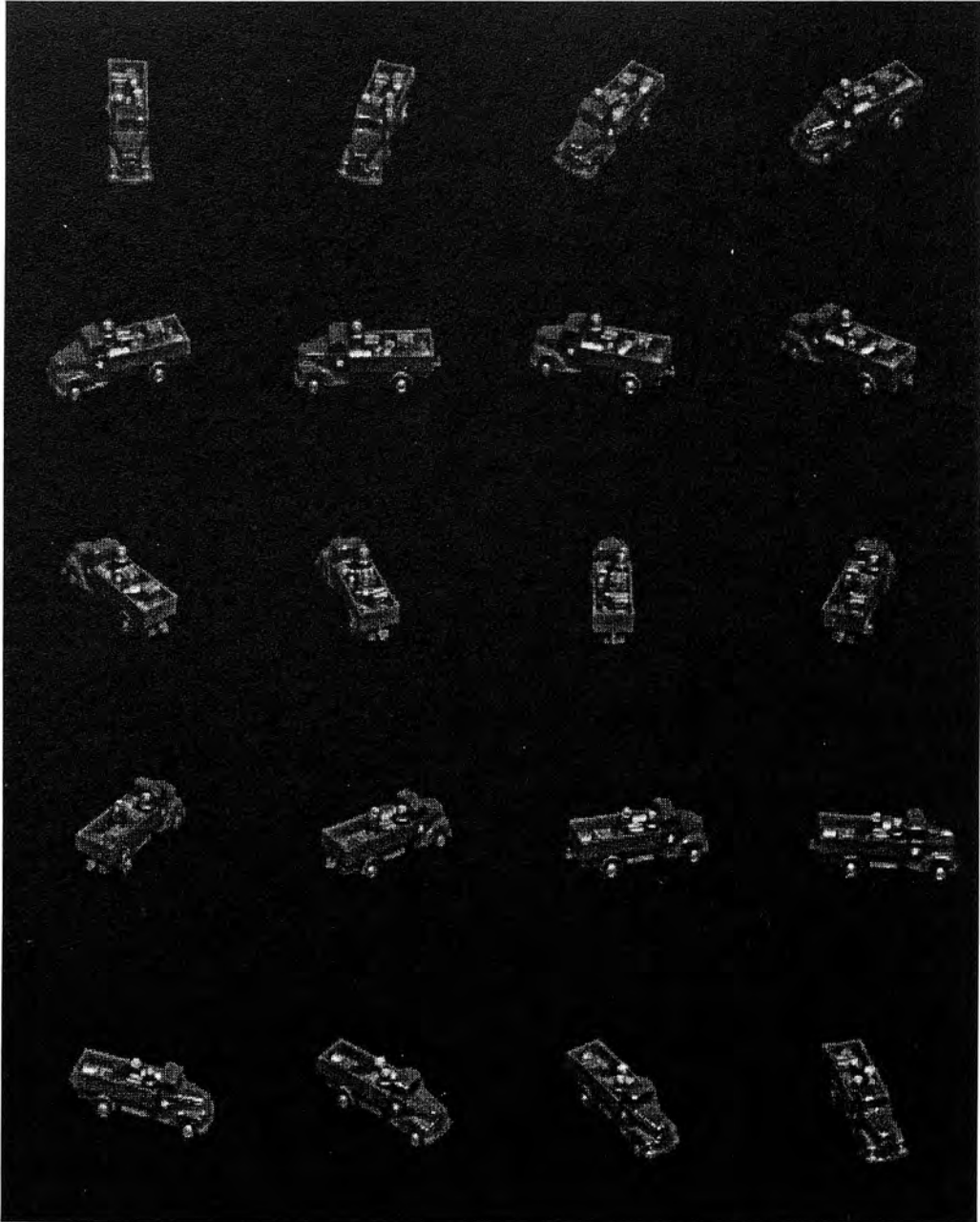
#### 6.4 Results.

The algorithm was tested by means of a computer simulation of the optical correlation process. Toward this end we designed several filters to solve a two-class problem involving different views of out-of-plane rotated objects. The true class -see Figure 6.5- was formed by twenty images of a tank captured each 18 degrees. The false class -see Figure 6.6- contains twenty images of a truck obtained in the same conditions. All the images are of 128x128 pixels and no special preprocessing such as edge enhancement was carried out. The training set, that is the set of images used in designing the filters, for all the examples that follow is composed of the ten samples of each object taken at  $0^\circ$ ,  $36^\circ$ , ...,  $324^\circ$  angles.

The performance of the algorithm was studied for four different coding domains -Fig. 6.7-: a phase only, a binary phase-only, a spiral coupling between amplitude and phase and an arbitrary domain. Phase-only filters are attractive designs because they provide a good trade-off between noise resistance and peak sharpness together with optimum light efficiency.

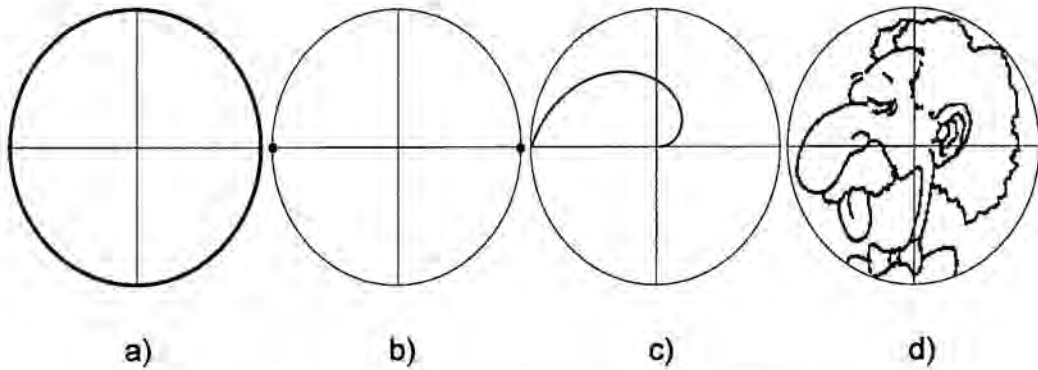


*Figure 6.5- True class images.*



*Figure 6.6- False class images.*





*Figure 6.7- Different coding domains used to test the method. a) phase-only b) binary phase-only c) spiral coupling between amplitude and phase d) arbitrary domain.*

Binary phase-only filters retain to a large extent the properties of the latter design but they can be implemented in actual devices such as Magneto-Optic Spatial Light Modulators (MOSLM), which in addition are very fast. The spiral domain is typical for liquid crystal displays (LCD). Finally, Fig. 6.7.d) shows a rather arbitrary modulator characteristic for which the algorithm will work.

The first issue addressed was the choice of the initial filter  $\mathbf{a}_0$ . The algorithm was found to be capable of producing SDF filters using a wide variety of starting points. In particular we tested:

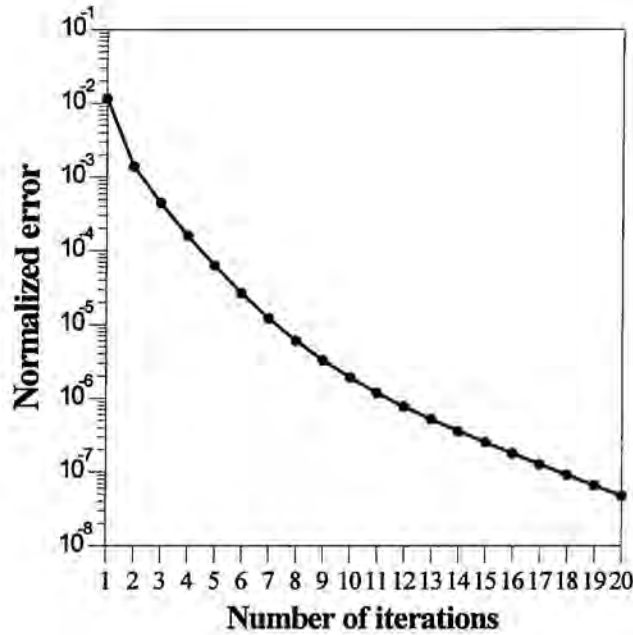
a) Full complex SDF filters designed to solve the same problem. We used them because sometimes the simple projection of an SDF filter is a good solution. For example the phase of a composite filter is sometimes a good phase-only SDF [Hor85] and thus would require only small modifications.

b) Random complex vectors. In contrast to case a) they contain no information about the problem.

c) The same starting point used in the Jared and Ennis algorithm[Jar89]:

$$\mathbf{a}_0 = \sum_{i=1}^M c_i \mathbf{x}_i \quad (6.31)$$





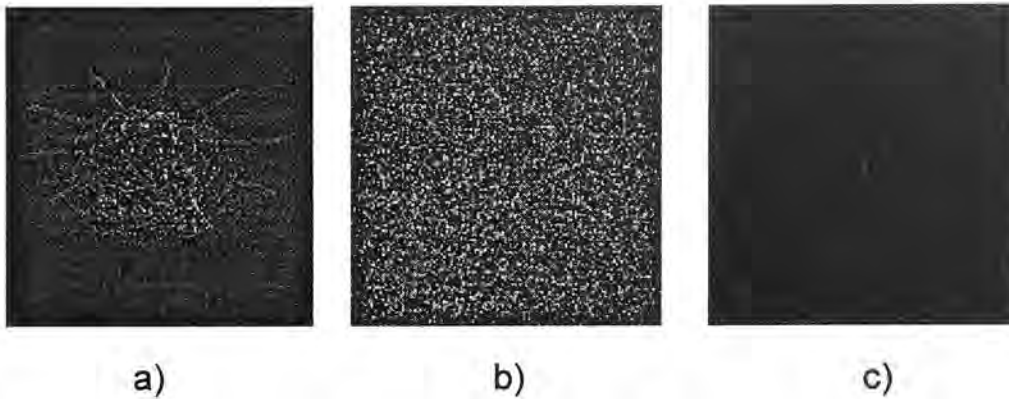
*Figure 6.8-* The plot shows the convergence of the algorithm for a phase-only SDF.

where  $x_i$  is the  $N$ -dimensional vector representing the  $i$ th training image,  $c_i$  its desired output and  $M$  the number of training images.

Figure 6.8 shows the convergence of the algorithm when designing a phase-only SDF with a MACE filter as a starting point. The  $Y$ -axis, which has a logarithmic scale, represents the error between the SDF and the constrained filter at a given iteration divided by the sum of the squared magnitude of the components of the SDF filter, i.e.

$$E_{\text{NORM}}(\mathbf{h}) = \left( \sum_{i=1}^N |h_k^i - a_k^i|^2 \right) / \left( \sum_{i=1}^N |h_k^i|^2 \right) \quad (6.32)$$

where  $\mathbf{h}$  is the SDF filter and  $\mathbf{a}$  the phase-only version of  $\mathbf{h}$ . The superscript  $i$  indicates the component and the subscript  $k$  the iteration. For an ideal phase-only filter this error function is zero. The graph shows an exponential decay with the number of iterations meaning a fast approach to the desired phase-only filter.

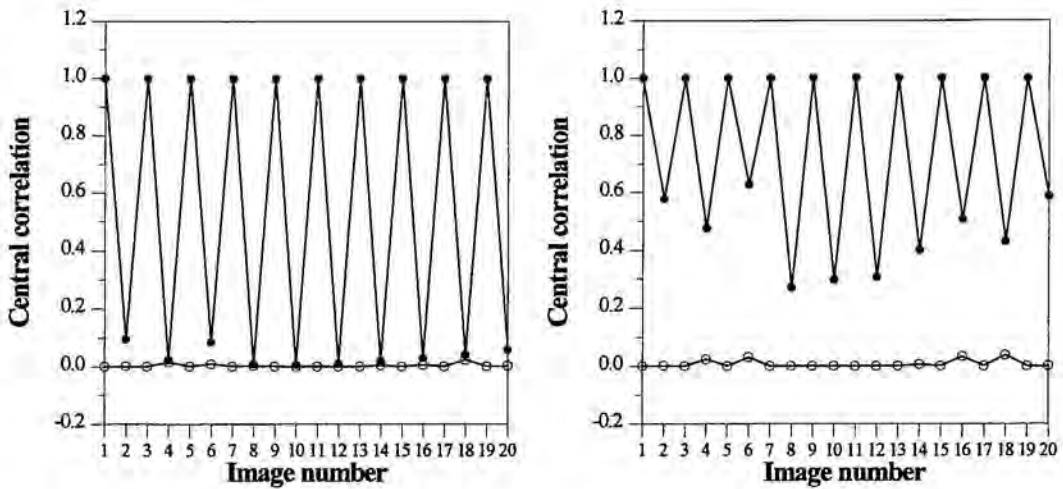


*Figure 6.9- Impulse response of three phase-only SDF designed with the same training set but with different starting points: a) the phase of a MACE filter b) a phase-only random vector c) a constant plane in Fourier space.*

The accuracy attained for the SDF conditions was found to be only slightly dependent upon the initial point although different number of iterations - from ten to twenty- were required for different points. However, depending on  $a_0$ , the behavior of the final filters may be very different because the algorithm seems to find a solution easily, without modifying so much the starting point. Since the solution is close to the initial filter, it preserves to some extent its characteristics.

We give two examples of this feature. In the first one, three phase-only SDF's were computed using the phase of a MACE filter, a phase-only random vector and a constant plane in Fourier space -a delta function in object space- as initial points. Figure 6.9 shows the impulse responses of the obtained filters. All three meet the SDF conditions with an accuracy of 96% and are very similar to their respective starting filters. For instance, filter 6.9.b) is clearly random or filter 6.9.c) is almost a delta function. The rest of the correlation plane is of course very different and this property enables an indirect control over the characteristics of the final filters. The second example illustrates more clearly this point.

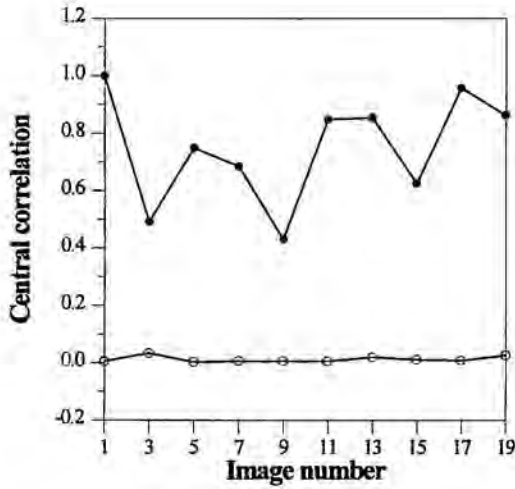
Figure 6.10.a) shows the central correlations between the whole set of 40 images -20 tanks and 20 trucks- and a phase-only SDF designed with only the 20 intermediate views. The starting point was the sum of the ten target images -the ten tanks-. Note that while there is a perfect control over the central correlations with the training images and small sidelobes -Fig. 6.15-, the correlations with the



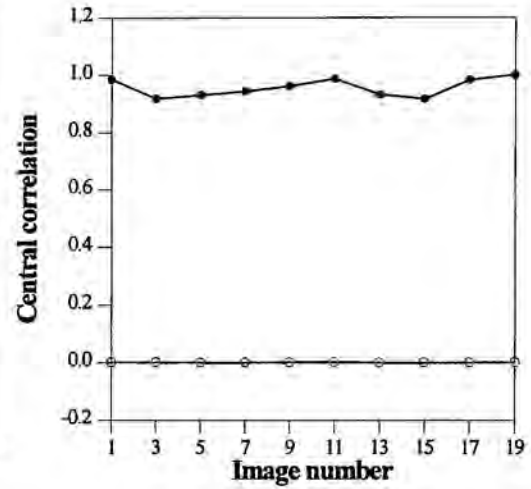
**Figure 6.10.-** Central correlations obtained with two different phase-only SDF: a) with a starting point formed by the sum of the tenk tanks of the training set, b) with a starting point formed by the sum of the whole set of 20 tanks. Open circles represent the correlations with the trucks. Filled circles represent those of the tanks.

tanks not included in the training set are too small to be separated from those of the trucks. The true class images can be separated from those belonging to the false class, if the initial point is formed by the sum of the whole set of true-class images -20 tanks-, as shown in Fig. 6.10.b). Thus, the election of an initial vector  $\mathbf{a}_0$  that includes information about the intermediate views leads to a filter with enhanced generalization capabilities.

The type of coding domain is the most influential factor with respect to the control of the central correlations. Figures 6.11, 6.12 and 6.13 show respectively the central correlations between the images of the training set and the binary phase-only, the spiral and the arbitrarily constrained filter whose domains are represented in Fig. 6.7. The correlations with the nontraining images -the intermediate views- are not shown since they strongly depend on the filter  $\mathbf{a}_0$  as commented. The two latter designs accurately meet the SDF constraints. The binary phase-only filter presents more difficulties because although the values for the true class images are significantly higher than those of the false class they show the most marked variation. This is due to the stop of the algorithm at about four iterations with all the initial points we used. Finally Figs. 6.15, 6.16, 6.17 and 6.18 present 3-D plots as well as a front and a lateral view of the intensity of

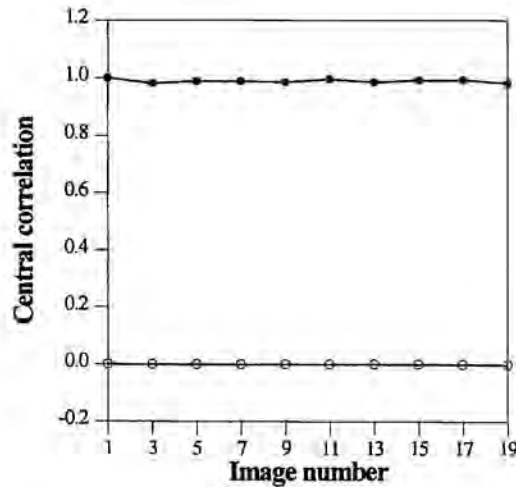


**Figure 6.11-** Central correlations between a binary phase-only SDF and the images of the training set. Open circles represent the correlations with the trucks. Filled circles represent those of the tanks.



**Figure 6.12-** Central correlations between a spirally constrained SDF filter and the images of the training set. Open circles represent the correlations with the trucks. Filled circles represent those of the tanks.

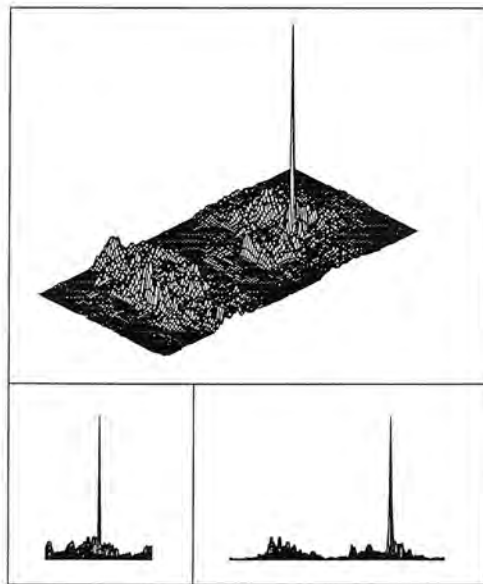
the correlation between the test scene of Fig. 6.14 and the four filters. A good detection of the tank is possible in all cases.



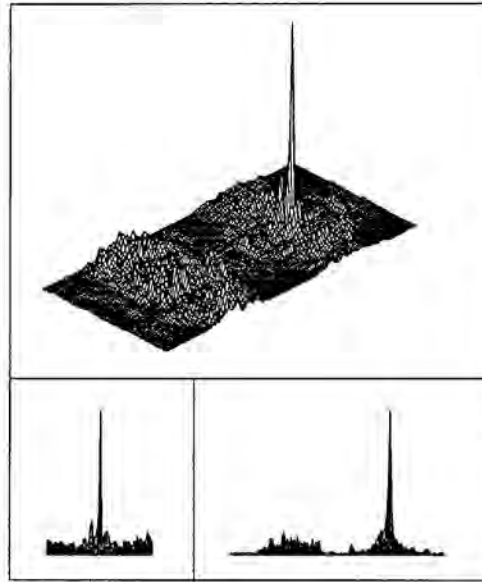
**Figure 6.13-** Central correlations between an SDF with values on the domain of Fig. 5.d) and the images of the training set. Open circles represent the correlations with the trucks. Filled circles represent those of the tanks.



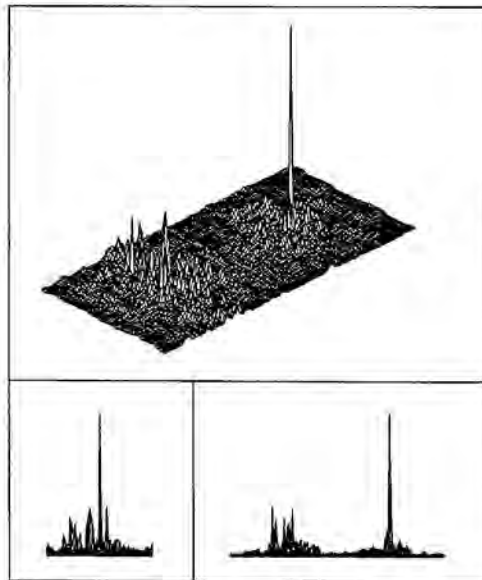
*Figure 6.14- Test input scene. Both images belong to the training set.*



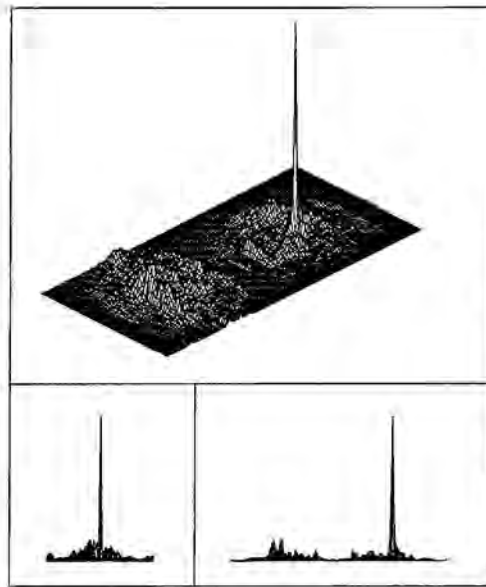
*Figure 6.15- Correlation between the phase-only SDF and input scene of Fig. 6.14.*



*Figure 6.16- Correlation between the binary phase-only SDF and input scene of Fig. 6.14.*



*Figure 6.17- Correlation between the spirally constrained SDF and input scene of Fig. 6.14.*



*Figure 6.18- Correlation between the arbitrarily constrained SDF and input scene of Fig. 6.14.*

In conclusion, a new algorithm for computing synthetic discriminant functions adapted to the restrictive modulation characteristics of present-day devices has been developed. In contrast to other previously proposed methods our procedure can be proved to be non-divergent. Furthermore it has a solid mathematical background that enables the analysis of the cases not leading to a solution. The algorithm needs only a few iterations, ranging from ten to twenty, to obtain the desired filter so the computational load is moderate. Finally, no special assumption over the shape of the filter was made, such as the imposition for the filter to be a linear combination of the training images and therefore multiple solutions can be reached by changing the initial point. This property enables an indirect control over the characteristics of the final filter as indicated by the results of the simulation. Nevertheless a systematic approach for the selection of the initial filter to take full profit of this feature must still be devised.

There are other possibilities worth exploring such as the use of different metrics to measure the similarity between the SDF and the constrained filter. The



change of the similarity criterion might permit to obtain a solution when this is not feasible using the Euclidean distance.

## 6.5 Appendix: scaling constant leading to minimum error.

We derive an expression for the scaling of the central correlations leading to minimum error as mentioned in section 6.3.

The expression for the error function to be minimized can be written in vector notation as -see eq. 6.10-:

$$E = \mathbf{h}^+ \mathbf{h} - \mathbf{a}^+ \mathbf{h} - \mathbf{h}^+ \mathbf{a} + \mathbf{a}^+ \mathbf{a} \quad (6.33)$$

where  $\mathbf{h}$  is the minimum Euclidean distance synthetic discriminant function (MED-SDF) filter:

$$\mathbf{h} = \alpha \mathbf{Comp} + \mathbf{Pa} \quad (6.34)$$

and  $\alpha$  is a real constant that scales the central correlations. By substituting eq. 6.34 into eq. 6.33 we get:

$$\begin{aligned} E(\alpha) &= (\alpha(\mathbf{Comp})^+ + \mathbf{a}^+ \mathbf{P})(\alpha \mathbf{Comp} + \mathbf{Pa}) - \mathbf{a}^+ (\alpha \mathbf{Comp} + \mathbf{Pa}) - \\ & - (\alpha(\mathbf{Comp})^+ + \mathbf{a}^+ \mathbf{P})\mathbf{a} + \mathbf{a}^+ \mathbf{a} = \alpha^2 (\mathbf{Comp})^+ \mathbf{Comp} + \alpha (\mathbf{Comp})^+ \mathbf{Pa} + \\ & + \alpha \mathbf{a}^+ \mathbf{P}(\mathbf{Comp}) + \mathbf{a}^+ \mathbf{P}^2 \mathbf{a} - \alpha \mathbf{a}^+ \mathbf{Comp} - \mathbf{a}^+ \mathbf{Pa} - \alpha (\mathbf{Comp})^+ \mathbf{a} - \mathbf{a}^+ \mathbf{Pa} + \mathbf{a}^+ \mathbf{a} \end{aligned} \quad (6.35)$$

Owing to the properties of the projection operator  $\mathbf{P}$ :

$$\begin{aligned} (\mathbf{Comp})^+ \mathbf{Pa} &= 0 \\ \mathbf{a}^+ \mathbf{P}(\mathbf{Comp}) &= 0 \\ \mathbf{a}^+ \mathbf{P}^2 \mathbf{a} &= \mathbf{a}^+ \mathbf{Pa} \Rightarrow \mathbf{a}^+ \mathbf{P}^2 \mathbf{a} - \mathbf{a}^+ \mathbf{Pa} = 0 \end{aligned} \quad (6.36)$$

we can write:

$$E(\alpha) = \alpha^2 (\mathbf{Comp})^+ \mathbf{Comp} - 2\alpha \mathbf{a}^+ \mathbf{Comp} - \mathbf{a}^+ \mathbf{Pa} + \mathbf{a}^+ \mathbf{a} \quad (6.37)$$

and finally by setting the derivative of  $E$  to zero we find the expression for  $\alpha$  leading to minimum error:

$$\frac{\partial E}{\partial \alpha} = 0 \Leftrightarrow 2\alpha (\mathbf{Comp})^+ \mathbf{Comp} - 2 \mathbf{a}^+ \mathbf{Comp} = 0 \quad (6.38)$$

whence

$$\alpha = \frac{\mathbf{a}^+ \mathbf{Comp}}{(\mathbf{Comp})^+ \mathbf{Comp}} \quad (6.39)$$

## **Chapter seven. The alternative: iterative correlators**

### **Introduction.**

In chapter 3 the inherent limitations of single-channel correlators were pointed out. The considerations about the constraints imposed by the decision boundaries lead to the conclusion that arbitrary classification of a set of patterns is not always possible. By using more filters and a proper combination of the results, for example using pointwise logic operations, the decision boundaries become more complex and those difficulties can be solved. A direct implementation of this idea for sidelobe elimination was the topic of chapters 4 and 5. However depending on the problem such procedure may be complex, mainly by the large amount of data produced by the correlations and therefore an alternative would be desirable. We propose in this section such an alternative, by showing that an iterative process involving a single-channel correlation processed by a simple nonlinear function is also free of restrictions. To do so we need to travel through the world of cellular automata and universal computers.

### **7.1 Cellular automata.**

Cellular automata were introduced by J. von Neumann as theoretical idealizations of biological systems to investigate the possibility of developing self-reproducing machines. Since then, they have been used for very different purposes and are considered a very promising framework to modelize complex physical systems -see for example [Wol83]-. They are usually described as a regular infinite array, a rectangular grid for example, of elements called cells that

can take one of several permitted states. The automaton evolves in discrete time steps -called generations- governed by deterministic rules. The state of a given cell in the generation  $k+1$  is a function of its state as well as of the states of its neighbors at generation  $k$ , where the neighborhood comprises the closer cells, usually in all directions, up to a given distance. An important feature of these systems is the synchronous dynamics; the cells are all updated at the same time by applying the rules over the former generation.

Some of these cellular machines exhibit a behavior complex enough to support universal computation [Lan91]. Loosely speaking, an universal computer is a machine with a fixed structure that following precisely determined instructions, which all the time direct its behavior, is capable of executing any algorithm no matter its complexity. The question of whether it is possible to build a more powerful machine than an universal computer is a controversial issue -see [Hof79] and [Pen89] for confronted opinions- but, at least at present, universal computation is the most powerful known form of information processing. Cellular automata with such capacities to handle information are called Class-IV automata being the most popular the Conway's Game of Life [Ber82]. We will use it to show our approach.

## 7.2 Optical Life.

Life consists of a bidimensional infinite rectangular array of cells each one with two possible states: dead or alive. The neighborhood is composed by the eight sites that lie at the  $3 \times 3$  square centered at the cell. The game has carefully chosen rules that control the death and birth of the cells in the successive generations. These rules can be summarized in the following way:

- Death rules: If a living cell at generation  $k$  has less than two living neighbors it dies at generation  $k+1$  owing to isolation. If a cell has more than three living neighbors it dies owing to overcrowding.

- **Birth rules:** If a dead cell has exactly three living neighbors at generation  $k$  it becomes alive at generation  $k+1$ . If a living cell has two or three living neighbors it continues alive at generation  $k+1$ .

Three features of the automaton are relevant for our discussion:

- Life is an universal computer.
- As customary in cellular automata, the updating of the state of all cells is carried out synchronously.
- The rules are totalistic. The state of a cell depends only on the sum of the states of its neighbors, assuming one for live cells and zero for the dead ones, and not in the particular configurations of live and dead cells.

The latter two characteristics allow the implementation of Life by means of correlation, which basically computes the alive neighbors and a simple nonlinear processing of such correlations, which performs the decision of which cells are dead and alive at the next time step. The filter and the nonlinearity we use are displayed in Figs. 7.1 and 7.2. This is neither the only possibility nor the optimum in any sense. The filter in object space is represented in fig. 7.1.a). As can be seen it is of size 3x3 pixels and symmetrical, which gives the real filter in

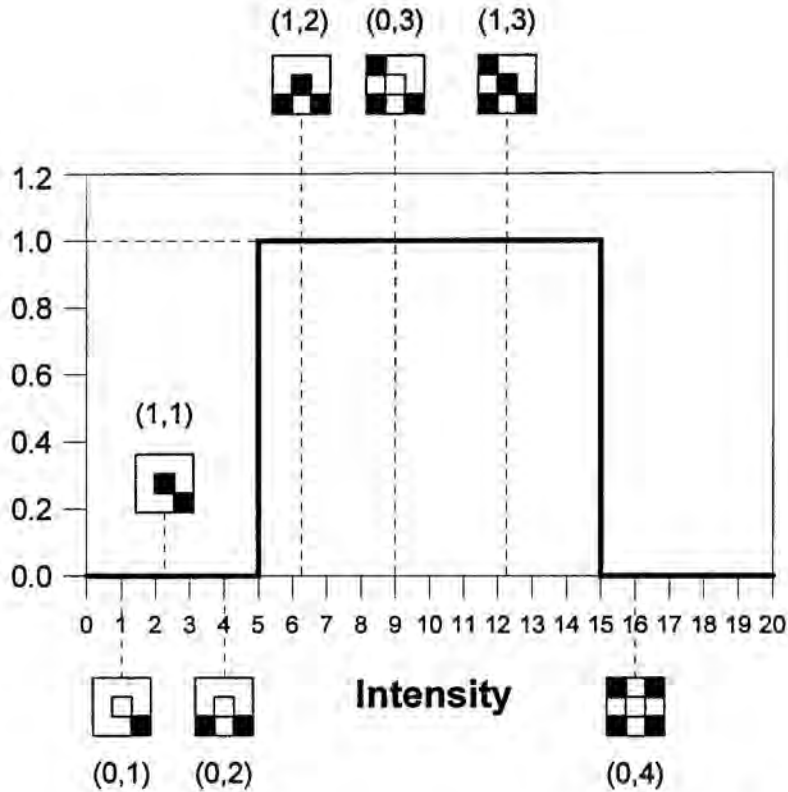
1	1	1
1	0.5	1
1	1	1

a)



b)

*Figure 7.1- a) Filter in object space. b) filter in Fourier space.*



**Figure 7.2-** Nonlinear function used to binarize the intensity. The figure shows several configurations of dead and alive cells and their corresponding correlation intensities. The alive neighbors are represented by filled squares. A gray square means that the considered cell is alive and a white one that it is dead.

Fourier space represented in Figure 7.1.b). The off-center values are all one because the effect of all neighbors over the state of the cell is the same. A different value for the center is needed to codify the distinct rules applied when the cell is dead or alive.

The possible values for the intensity of the correlation given by such a filter are summarized in Table 7.1. The column labeled as Output, gives the result after the nonlinear function  $f(x)=rect(x/10-1)$  shown in Fig. 7.2. It is easy to verify that the rules are fulfilled.

The extremely simple rules that govern the evolution of an initial configuration of cells produce, however, a very rich behavior: some initial patterns decay rapidly until they become extinguished and some turn into

Cell state	No of neighbors	Correlation	Output
0	0	0	0
	1	1	0
	2	4	0
	3	9	1
	4	16	0
	5	25	0
	6	36	0
	7	49	0
	8	64	0
1	0	0.25	0
	1	2.25	0
	2	6.25	1
	3	12.25	1
	4	20.25	0
	5	30.25	0
	6	42.25	0
	7	56.25	0
	8	72.25	0

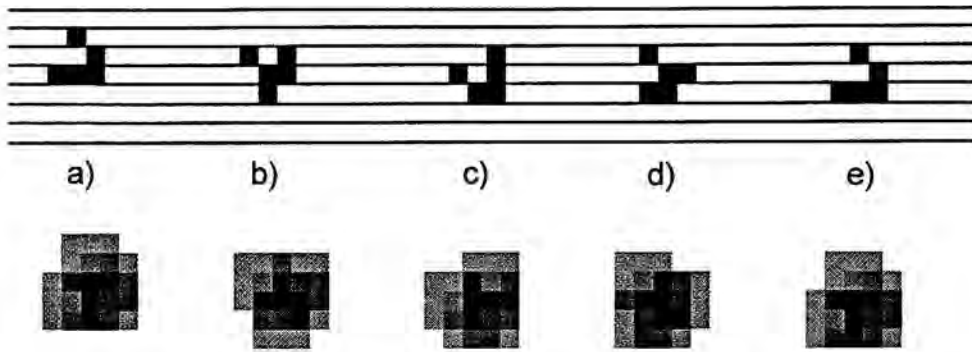
*Table 7.1- The first column indicates the state of the cell. Depending on the number of alive neighbors, listed in column two, the different correlation intensity values of column three are obtained. The fourth column shows the result of these intensity values after being processed by the nonlinear function of Figure 7.2.*

periodic configurations but all of them evolve in highly unpredictable ways. The most interesting of these patterns because they are the building blocks of the universal computer are:

- The glider: it is a periodic structure composed of five living cells. After four generations the initial pattern of cells is recovered but centered at the next diagonal site. This traveling structure enables the interchange of signals between the different parts of the Life universe. Typically the presence or absence of a glider encodes one bit of information. The pattern is represented in figure 7.3.

When two gliders traveling in opposite directions meet each other, different reactions can occur depending upon their particular phases inside the



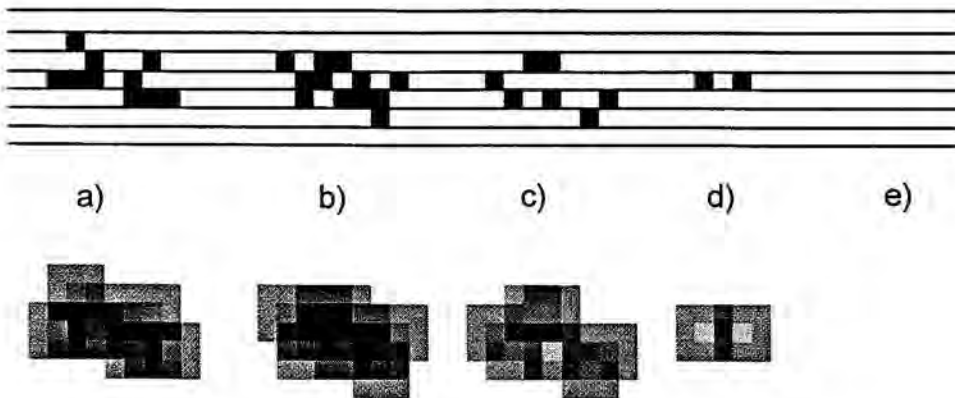


**Figure 7.3-** Up: Life structures. Down: Correlation intensities; their binarization gives the pattern at the next generation.

a), b), c), d) Glider at generations 0, 1, 2, 3 respectively. e) Generation 4; it is the same as that in a) but displaced one site right and one down.

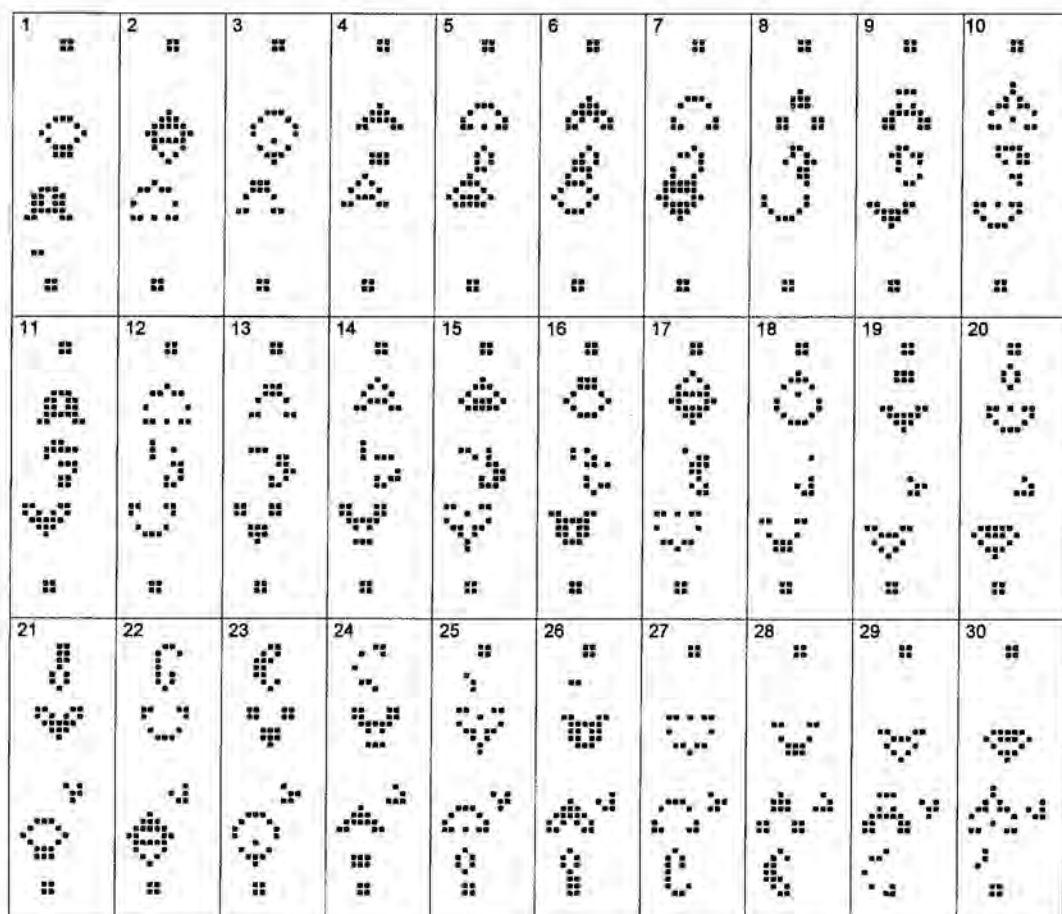
four-generation cycle. The most interesting are the vanishing reactions by which both gliders are destroyed in a variable number of steps -Fig. 7.4-.

- The glider-gun: it is a structure of period 30 composed by 45 living cells in its first generation. After 30 time steps it emits a glider and recovers its initial shape. The pattern is used to produce the constant stream of gliders necessary in



**Figure 7.4-** Up: Life structures. Down: Correlation intensities; their binarization gives the pattern at the next generation.

A vanishing reaction. a) Gliders traveling in opposite directions. b) Collision. c), d) and e) The resulting reaction annihilates both gliders in four steps.

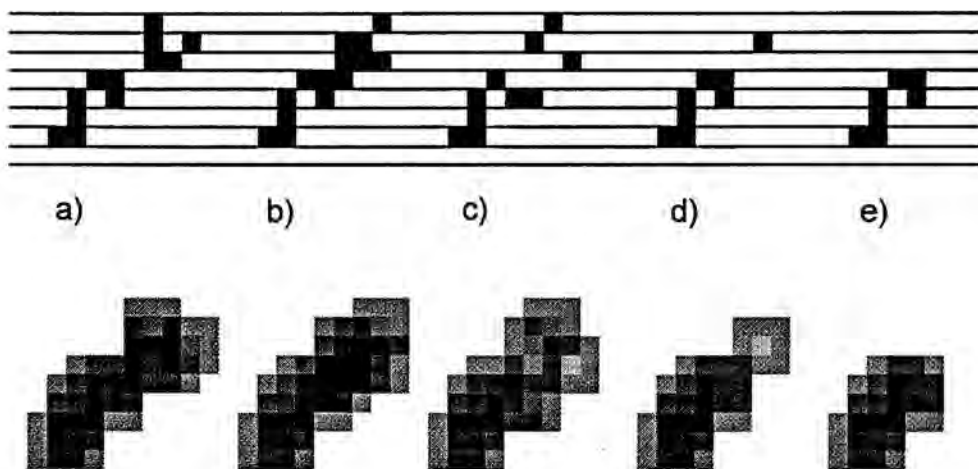


*Figure 7.5-* The 30 different generations of a glider-gun; note the new glider between the two large structures at generation 18.

the construction of logic gates. Figure 7.5 displays the 30 generations of a glider-gun.

and finally,

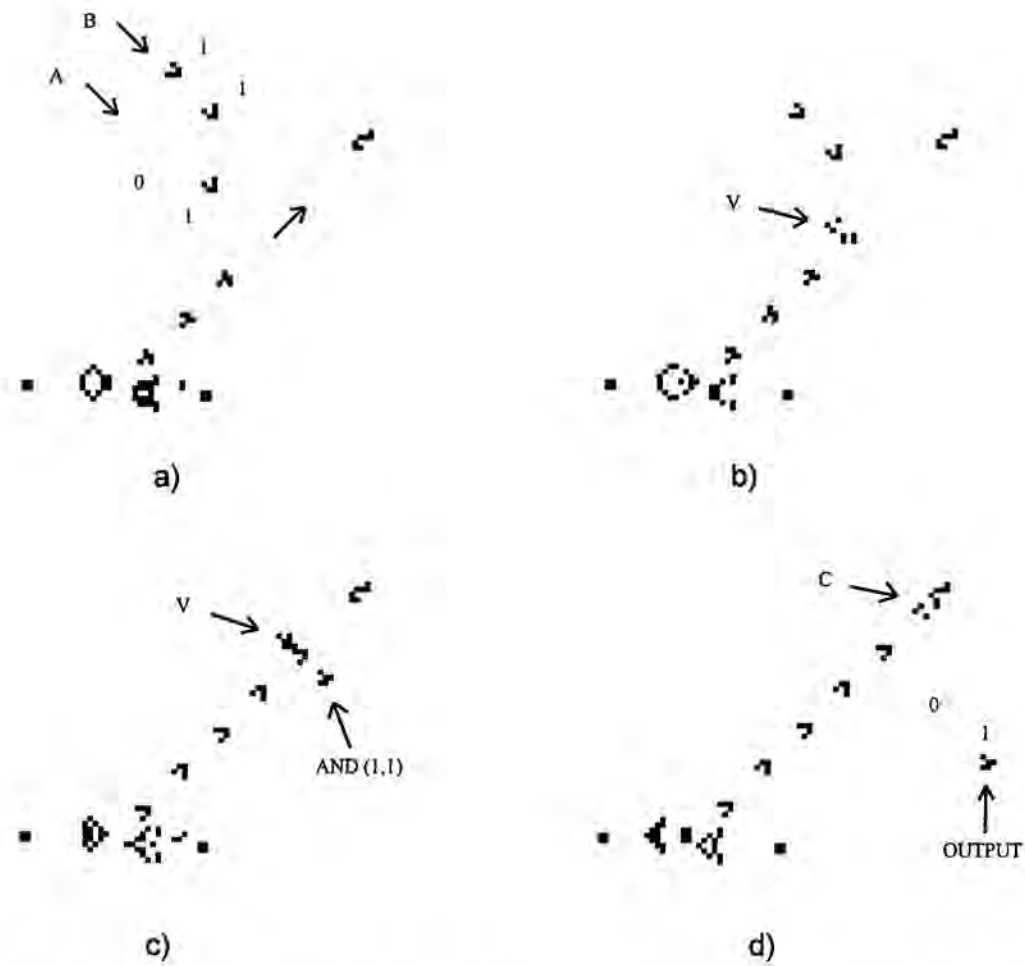
- The eater: it is a stable pattern that is capable of annihilating a glider in four generations without being altered in the collision -Figure 7.6-. It is commonly used in logic gates to eliminate the unwanted gliders produced by a glider-gun, and to stabilize different reactions.



*Figure 7. 6- Up: Life structures. Down: Correlation intensities; their binarization gives the pattern at the next generation.*

As an example of the computing capabilities of the automaton we show the construction of an AND gate. The input pattern is that shown in Figure 7.7.a) where the two inputs are labeled by the letters **A** and **B**. The signals are carried by gliders. The presence of a glider encodes the binary digit one. Analogously, its absence encodes the bit zero. The glider-gun at the bottom left corner emits a continuous stream of gliders which will eventually collide with the data as can be observed in Fig. 7.7. b).

The logic operation is performed by these collisions. When there is a glider in both channels -**A** and **B**-, one of them opens a hole in the stream by means of a vanishing reaction -letter *V* in Fig. 7.7.b)- that is profited by the second glider to pass through as sketched in Figure 7.7.c). A single glider in either channel is destroyed in the collision -Fig. 7.7.c)- and therefore by detecting the presence of a glider at the point indicated as output in figure 7.7.d) we obtain the AND operation. The non used gliders of the stream are eliminated by an eater to avoid interferences with other structures -the collision **C** marked in fig. 7.7.d)-. It is worth noting that if we use as output the rest of the stream produced by the glider-gun the resulting operation is a NOR gate. This latter operation is



**Figure 5.7.-** Life AND gate. a) The desired operations are  $AND(A=1,B=1)$  and  $AND(A=0,B=1)$ . b) the first glider of channel A opens a hole in the stream. c) The first glider of channel B pass while the solitary second glider of channel B is destroyed. d) The eater eliminates the residual gliders.

important because every logical function can be carried out by means of a suitable combination of NOR gates.

Based mainly on these patterns we can perform a complete imitation of an universal computer -see [Ber82] for details-. As a consequence, the iterative correlation process by which we have implemented Life, can be used to solve arbitrarily complex mathematical problems such as those involved in pattern recognition.

A few comments about this surprising result are necessary. First of all, the proof for Life to be universal is carried out by showing the possibility to build inside the automaton, a replica of an electronic computer. However computation neither necessarily imply logic operations nor the particular architecture of such machines. What we can ensure is that there exists, for any describable process, a Life pattern that encodes the problem, which after some iterations produces a different pattern that codifies the solution. The encoding and the decoding must be simple in such a way that these steps do not hide an universal computer. It might be possible to find a more natural way to perform complex calculations inside Life. The same argumentation can be applied to our optical process. It is not necessary to codify pattern recognition problems as Life structures to solve them. We have used the game to show that the proposed architecture is free of the a priori restrictions of single-filter correlations. There is only a little difficulty in this approach. How do we must design the filters to use the iterative architecture the natural way?

## Conclusions

The scientific objectives and the corresponding original results mentioned in the introduction and developed throughout the dissertation can be summarized in the following way:

- Chapter one and two:

These are basically introductory chapters and give theoretical support to the rest of the work. However a compilation of results, some of them little known, and a unification work has been carried out, so that they should be mentioned in the conclusions. In particular, the following may be considered interesting contributions:

\* The deduction of the capability of lenses to produce Fourier transforms at finite distance, based upon considerations of geometrical optics and by means of methods similar to those proposed by Keller and his geometrical theory of diffraction. This ad hoc demonstration could be more intuitive.

\* The detailed derivation of the necessity of using a filter with the same space bandwidth as the input scene, and not necessarily equal to that of the target, a result already known but seldom made explicit.

\* The detailed calculation of the computational capacity of an actual optical correlator in terms of floating point operations by second. This kind of equivalence with the digital systems is often mentioned, but the number varies

depending on the author. Finally, such estimation enables us to compare with off-the-shelf digital hardware, from supercomputers to specialized image processing cards.

\* Presentation of the filter design techniques as a natural evolution based on two opposing needs: the flexibility of adaptation to applications of growing complexity and the reliability of the detections. The conditions that ensure the detection of an image have been made explicit. Also the conditions in which single-image filters can be used with confidence are pointed out. Later, multiple-image filters are presented and the reasons for the appearance of sidelobes are clarified. The section about synthetic discriminant functions filters is presented in a novel way, by means of geometrical interpretations based on the concept of generalized metric. Finally, the equivalence between circular harmonic filters and synthetic discriminant functions is proved through an original derivation. This result, first obtained by Réfrégier, is little known and of great theoretical interest.

- Chapter three:

In this chapter we adapt the formalism of the decision regions to the particular case of the optical correlation.

\* We obtain that the decision boundaries of an optical correlator are a pair of hyperplanes, whose equation is determined by the filter and the binarization threshold.

\* This formalism allows us to point the inability of single filter correlations to handle complex pattern recognition problems.

\* Finally it is shown how multichannel correlations allow to attain more complex decision boundaries therefore providing a solution to the above problem.



- Chapter four and five:

As an example of the later statement and as a method of practical interest we develop a procedure to eliminate lateral peaks -sidelobes-. Chapter four is devoted to correlations giving only real values and chapter five deals with the generalization to the case of complex correlations.

The results of the real case can be summarized as follows:

- \* It is shown how the combined action of two properly designed filters ensures the suppression of the false alarms caused by the spurious peaks, even if they are higher than the central peak.

- \* The mathematical expression of the necessary correcting filter is obtained. This is a particular case of a previous design proposed by Kumar.

- \* We deduce the necessary conditions to have a working procedure.

- \* Based on these conditions we derive the optimal values for the parameters envisaged by the procedure.

- \* It is mathematically proved that the correcting power of the method is higher when larger sidelobes are expected. We also check this property by means of a computer simulation.

- \* The superiority of this process with respect to antisidelobe solutions based on a single filter is illustrated. In particular we give an example in which the method allows the total elimination of the sidelobes not removed by a MACE filter.

- \* Finally, we carry out an experimental implementation of the procedure that gives satisfactory results and in agreement with previous numerical simulations.

The complex case is treated with less detail. In particular we have

- \* An analysis of the conditions under which the method ensures the solution to the problem.

- \* The derivation of the number of filters required to suppress a sidelobe of a given height.

- \* The deduction of the optimal values for the parameters.

- \* An algorithm summarizing the procedure.

- \* The verification by computer simulation of the method performance for a typical problem of character recognition.

- Chapter six:

The practical possibility to carry out multichannel correlations in a reliable, fast and convenient way involves necessarily the use of spatial light modulators. This tackles a difficulty owing to the limited modulation capability of currently available devices, which makes necessary the development of filters taking values only on a small portion of the complex plane. Whereas the case of single-image filters is satisfactorily solved thanks to the work of Juday and Réfrégier, the algorithms proposed for the case of SDF filters are deficient and present several difficulties. Taking care of this necessity

- \* We have proposed an iterative algorithm for the computation of arbitrarily constrained synthetic discriminant functions.

- \* The algorithm is nondivergent, involves a moderate computational load and is capable of providing multiples solutions.

\* This later characteristic is of special importance and represents a difference with earlier methods. The possibility of achieving multiple solutions permits us to select the most appropriate one depending on the application.

\* The influence of the initial point of the iterative procedure over the properties and final behavior of the filter has been studied. The results show that suitable choice of this initial point enables us to control, in an indirect way, the generalization capabilities of the filter.

\* We have studied the convergence of the algorithm and the results obtained are highly satisfactory. The most problematic restriction was that of the binary filters, although, even in this case, the procedure produces usable solutions.

\* Finally we have obtained simulation results for four different domains: phase-only filters, binary phase-only, with spiral coupling and for an arbitrary domain, in a two-class problem requiring three-dimensional rotation invariance. The obtained filters show in all cases a suitable performance.

#### - Chapter seven:

In this chapter we propose an alternative to multichannel correlations as the means to overcome the limitations of single filter correlators. In particular

\* We show that an iterative correlator where the output has been processed by means of a simple nonlinear function -similar to the usual binarization procedure- has universal computation capabilities.

\* This proof, which involves an argumentation based on the cellular automaton called Life, is a theoretical result and not necessarily indicates how to solve particular problems with this architecture.

\* Since universal computation is the most complex form of information processing, the development of iterative procedures -an open question- would allow the solution of general pattern recognition problems.

## Bibliography

[Alm94] - G. S. Almasi and A. Gottlieb, Highly Parallel Computing, Benjamin/Cummings Publishing Company, Redwood City (1994).

[Aww90] - A. A. S. Awwal, M. A. Karim, and S. R. Jahan, "Improved correlation discrimination using an amplitude-modulated phase-only filter," *Appl. Opt.* **29**, 233-236 (1990).

[Bah89] - Z. Bahri and B. V. K. Vijaya Kumar, "Algorithms for designing phase-only synthetic discriminant functions," *Proc. of SPIE* **1151**, 138-147 (1989).

[Bau93] - K. A. Bauchert, "Real-time hardware implementation of an optical correlation image preprocessing algorithm using an off-the-shelf image processing board," *Proc. of SPIE* **1959**, 44-47 (1993).

[Ber82] - E. Berlekamp, J. Conway, and R. Guy, "What is Life?," in Winning Ways for Your Mathematical Plays, Academic Press, New York (1982).

[Bor75] - M. Born and E. Wolf, Principles of Optics, Pergamon Press, Oxford (1975).

[Bra65] - R. Bracewell, The Fourier Transform and its Applications, McGraw-Hill, New York (1965).

[Bur70] - C. B. Burckhardt, "A simplification of Lee's method of generating holograms by computer," *Appl. Opt.* **9**, 1949 (1970).

[Cal94] - D. L. Calloway, "Tanks in trees: A case study of ternary phase-amplitude filter classification trees," *Proc. of SPIE* **2297**, 250-259 (1994).

[Cam91] - J. Campos, S. Bosch, and J. Sallent, "Optical pattern recognition in defocused images using correlation filters," *Opt. Comm.* **82**, 370-379 (1991).

[Car93] - A. Carnicer, I. Juvells, and S. Vallmitjana, "Design of inverse filter for pattern recognition. Analysis of the performance of an amplitude-compensated matched filter," *J. Mod. Opt.* **40**, 391-400 (1993).

- [**Carl92**] - D. W. Carlson and B. V. K. Vijaya Kumar, "Synthetic discriminant functions for implementation on arbitrarily constrained devices," Proc. of SPIE **1772**, 10-20 (1992).
- [**Cas84**] - D. Casasent, "Unified synthetic discriminant function computational formulation," Appl. Opt. **23**, 1620-1627 (1984).
- [**Cas86**] - D. Casasent and W. A. Rozzi, "Computer-generated and phase-only synthetic discriminant function filters," Appl. Opt. **25**, 3767-3772 (1986).
- [**Cas91**] - D. Casasent, G. Ravichandran, and S. Bollapraggada, "Gaussian MACE correlation filters," Appl. Opt. **30**, 5176-5181 (1991).
- [**Cas92**] - D. Casasent, R. Schaefer, and R. Sturgill, "Optical hit-miss morphological transform," Appl. Opt. **31**, 6255-6263 (1992).
- [**Cas94**] - D. Casasent, "General-purpose optical pattern recognition image processors," Proc. IEEE **82**, 1724-1734 (1994).
- [**Caw89**] - P. S. Cawte, I. R. Cooper, G. C. Gibbons, S. C. Webster, F. Dubois, M. Büchel, J. M. Brodin, J. P. Schnell, and B. Loiseaux, "An optical image correlator for robotic applications," Proc. of SPIE **1134**, 196-203 (1989).
- [**Cla91**] - N. Clark, C. M. Crandall, and M. K. Giles, "Using liquid crystal TVs in Vander Lugt optical correlators," Proc. of SPIE **1564**, 439-451 (1991).
- [**Cov65**] - T. M. Cover, "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition," IEEE Trans. Electron. Comput., **EC-14**, 326-324 (1965).
- [**Cra91**] - C. M. Crandall, M. K. Giles, and N. Clark, "Performance limitations of miniature optical correlators," Proc. of SPIE **1564**, 98-109 (1991).
- [**Cro93**] - D. G. Crowe, J. Shamir, and T. W. Ryan, "Sidelobe reduction in optical signal processing," Appl. Opt. **32**, 179-183 (1993).
- [**Dic89a**] - F. M. Dickey and B. D. Hansche, "Quad-phase correlation filters for pattern recognition," Appl. Opt. **28**, 1611-1613 (1989).
- [**Dic89b**] - F. M. Dickey and L. A. Romero, "Dual optimality of the phase-only filter," Opt. Lett. **14**, 4-5 (1989).

- [Dic90] - F. M. Dickey, B. V. K. Vijaya Kumar, L. A. Romero, and J. M. Connelly, "Complex ternary matched filters yielding high signal-to-noise ratios," *Opt. Eng.* **29**, 994-1001 (1990).
- [Dud73] - R. O. Duda, P. E. Hart, Pattern classification and scene analysis, Wiley, New York (1973).
- [Fle90] - M. Fleisher, U. Mahlab, and J. Shamir, "Entropy optimized filter for pattern recognition," *Appl. Opt.* **29**, 2091-2098 (1990).
- [Fla89] - D. L. Flannery and J. L. Horner, "Fourier optical signal processors," *Proc. IEEE* **77**, 1511-1527 (1989).
- [Geb91] - R. Gebelein, S. Connely, and L. Foo, "Advances in the optical design of miniaturized optical correlators," *Proc. of SPIE* **1564**, 452-463 (1991).
- [Ghe89] - G. Gheen, "Maximum mean square projection filters," *Proc. of SPIE* **1151**, 278-283 (1989).
- [Ghe92] - G. Gheen, E. Washwell, and C. Huang, "Problems facing optical correlators," *Proc. of SPIE* **1772**, 96-103 (1992).
- [Gon93] - R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison Wesley, Reading (1993).
- [Goo68] - J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill, San Francisco (1968).
- [Gua93] - O. Gualdron and H. Arsenault, "Locally nonlinear matched filters," *Opt. Lett.* **18**, 1949-1951 (1993).
- [Hed92] - S. Heddle, D. G. Vass, and R. M. Sillitto, "Reduction of aliasing in correlation using a pixelated spatial light modulator," *Proc. of SPIE* **1772**, 116-127 (1992).
- [Hes80] - C. F. Hester and D. Casasent, "Multivariant technique for multiclass pattern recognition," *Appl. Opt.* **19**, 1758-1761 (1980).
- [Hof79] - D. R. Hofstadter, Gödel, Escher, Bach: an Eternal Golden Braid, Harvester Press, Hassocks, Sussex (1979).



- [Hor84] - J. L. Horner and P. D. Gianino, "Phase-only matched filtering," *Appl. Opt.* **23**, 812-816 (1984).
- [Hor85a] - J. L. Horner and J. R. Leger, "Pattern recognition with binary phase-only filters," *Appl. Opt.* **24**, 609-611 (1985).
- [Hor85b] - J. L. Horner and P. D. Gianino, "Applying the phase-only filter concept to the synthetic discriminant function correlation filter," *Appl. Opt.* **24**, 851-855 (1985).
- [Hua91] - C. H. Huang, G. Gheen, and E. Washwell, "Throughput comparison of optical and digital correlators for automatic target recognition," *Proc. of SPIE* **1564**, 427-438 (1991).
- [Jar86] - D. A. Jared and D. J. Ennis, "Learned distortion-invariant pattern recognition using synthetic discriminant functions," *Proc. of SPIE* **638**, 91-101 (1986).
- [Jar89] - D. A. Jared and D. J. Ennis, "Inclusion of filter modulation in synthetic-discriminant-function construction," *Appl. Opt.* **28**, 232-239 (1989).
- [Jav89] - B. Javidi, "Nonlinear joint power spectrum based optical correlation," *Appl. Opt.* **28**, 2358-2367 (1989).
- [Jen87] - E. L. Jenkins, Jr., J. W. Morris, and S. R. F. Sims, "Synthetic discriminant functions for target correlation," *Proc. of SPIE* **781**, 140-147 (1987).
- [Jud89] - R. D. Juday, "Correlation with a spatial light modulator having phase and amplitude cross-coupling," *Appl. Opt.* **28**, 4865-4869 (1989).
- [Jud93] - R. D. Juday, "Optimal realizable filters and the minimum Euclidean distance principle," *Appl. Opt.* **32**, 5100-5111 (1993).
- [Juv91] - I. Juvells, S. Vallmitjana, A. Carnicer, and J. Campos, "The role of amplitude and phase of the Fourier transform in the digital image processing," *Am. J. Phys.* **59**, 744-748 (1991).
- [Kal86] - R. R. Kallman, "Optimal low noise phase-only and binary phase-only optical correlation filters for threshold detectors," *Appl. Opt.* **25**, 4216-4217 (1986).

- [Kas87] - S. A. Kassam, Detection of Signals in NonGaussian Noise, Springer-Verlag, New York (1987).
- [Kast91] - B. A. Kast and F. M. Dickey, "Normalization of correlations," Proc. of SPIE **1564**, 34-42 (1991).
- [Kau90] - M. A. Kaura and W. T. Rhodes, "Optical correlator performance using a phase-with-constrained-magnitude complex spatial filter," Appl. Opt. **29**, 2587-2593 (1990).
- [Lan91] - C. Langton, "Life at the Edge of Chaos," in Artificial Life II: Proceedings of the Second Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems. C. Langton et al. Eds, Santa Fe Institute Studies in the Sciences of Complexity Series **10**, 41-91 (1991).
- [Lau94] - V. Laude and Ph. Réfrégier, "Multicriteria characterization of coding domains with optimal Fourier spatial light modulators filters," Appl. Opt. **33**, 4465-4471 (1994).
- [Lin94] - S. D. Lindell, "Transfer of optical processing to systems: optical pattern recognition program overview," Optical Processing and Computing SPIE's Working Group Newsletter, May 1994 issue, 6-7 (1994).
- [Mah87] - A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," Appl. Opt. **26**, 3633-3640 (1987).
- [Mahl89] - U. Mahlab and J. Shamir, "Phase-only entropy-optimized filter generated by simulated annealing," Opt. Lett. **14**, 1168-1170 (1989).
- [Met89] - A. Metioui, H. H. Arsenault, and L. Leclerc, "Methods for reducing sidelobes associated with composite filters," Opt. Comm. **71**, 332-336 (1989).
- [Met90] - A. Metioui and L. Leclerc, "Sidelobe reduction methods in optical pattern recognition," J. Opt. (Paris) **21**, 161-170 (1990).
- [Mol91] - P. A. Molley, W. C. Sweatt, and D. S. Strong, "A miniature acousto-optic image correlator," Proc. of SPIE **1564**, 610-616 (1991).
- [Mu88] - G. Mu, X. Wang, and Z. Wang, "Amplitude-compensated matched filtering," Appl. Opt. **27**, 3461-3463 (1988).

- [Nus82] - H. J. Nussbaumer, Fast Fourier Transform and Convolution Algorithms, *Springer Series in Information Sciences*, Springer-Verlag, Berlin (1982).
- [Opp81] - A. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE* **69**, 529-541 (1981).
- [Pen89] - R. Penrose, The Emperor's New Mind, Oxford University Press, Oxford (1989).
- [Pic93] - B. Picinbono, Random Signals and Systems, *Prentice-Hall Signal Processing Series*, Prentice-Hall International, Englewood Cliffs (1993).
- [Pre88] - W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, Numerical Recipes in C. The Art of Scientific Computing, Cambridge University Press, Cambridge (1988).
- [Psa82] - D. Psaltis, "Optical image correlation using acousto-optic and charge-coupled devices," *Appl. Opt.* **21**, 491-495 (1982).
- [Raj92] - H. Rajbenbach, S. Bann, P. Réfrégier, P. Joffre, J. Huignard, H. Buchkremer, A. Skov Jensen, E. Rasmussen, K. Brenner, and G. Lohman, "Compact photorefractive correlator for robotic applications," *Appl. Opt.* **31**, 5666-5674 (1992).
- [Rav91] - G. Ravichandran and D. Casasent, "Generalized in-plane rotation-invariant minimum average correlation energy filter," *Opt. Eng.* **30**, 1601-1607 (1991).
- [Réf90a] - Ph. Réfrégier, "Filter design for optical pattern recognition: multicriteria optimization approach," *Opt. Lett.* **15**, 854-856 (1990).
- [Réf90b] - Ph. Réfrégier, "Méthodes de reconnaissance des formes pour la corrélation optique," *Revue Technique THOMSON-CSF* **22**, 649-734 (1990).
- [Réf90c] - Ph. Réfrégier and J. P. Huignard, "Phase selection of synthetic discriminant functions filters," *Appl. Opt.* **29**, 4772-4778 (1990).
- [Réf91a] - Ph. Réfrégier, "Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency," *Opt. Lett.* **16**, 829-831 (1991).

[Réf91b] - Ph. Réfrégier and J. Figue, "Optimal trade-off filters for pattern recognition and comparison with Wiener approach," *Optical Computing & Processing* **1**, 245-266 (1991).

[Rog91] - S. K. Rogers and M. Kabrisky, An Introduction to Biological and Artificial Neural Networks, *SPIE Tutorial Texts Series* Vol. TT4, SPIE Press, Bellingham (1991).

[Ros91] - J. Rosen and J. Shamir, "Application of the projection-onto-constraint-sets algorithm for optical pattern recognition," *Opt. Lett.* **16**, 752-754 (1991).

[Sal95] - J. Sallent, M. Montes-Usategui, J. Campos, and S. Bosch, "A filter design to improve the recognition of defocused images," *Inst. Phys. Conf. Ser.* **139**: Part III, 349-352 (1995).

[Son94] - M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis and Machine Vision, *Chapman & Hall Computing Series*, Chapman & Hall, London (1994).

[Van64] - A. VanderLugt, "Signal detection by complex spatial filtering," *IEEE Trans. Info. Theory* **IT-10**, 139-145 (1964).

[Van92] - A. VanderLugt, Optical Signal Processing, *Wiley Series in Pure and Applied Optics*, J. W. Goodman Ed., Wiley, New York (1992).

[Vij86] - B. V. K. Vijaya Kumar, "Minimum variance synthetic discriminant functions," *J. Opt. Soc. Am. A* **3**, 1579-1584 (1986).

[Vij88] - B. V. K. Vijaya Kumar, Z. Bahri and A. Mahalanobis, "Constraint phase optimization in minimum variance synthetic discriminant functions," *Appl. Opt.* **27**, 409-413 (1988).

[Vij89a] - B. V. K. Vijaya Kumar, D. Casasent, and A. Mahalanobis, "Correlation filters for target detection in a Markov model background clutter," *Appl. Opt.* **29**, 2997-3006 (1990).

[Vij89b] - B. V. K. Vijaya Kumar and Z. Bahri, "Phase-only filters with improved signal to noise ratio," *Appl. Opt.* **28**, 250-257 (1989).

[Vij90a] - B. V. K. Vijaya Kumar and L. Hassebrook, "Performance measures for correlation filters," *Appl. Opt.* **29**, 2997-3006 (1990).

[Vij90b] - B. V. K. Vijaya Kumar, W. Shi, and C. Hendrix, "Phase-only filters with maximally sharp correlation peaks," *Opt. Lett.* **15**, 807-809 (1990).

[Vij92a] - B. V. K. Vijaya Kumar, A. Mahalanobis, S. Song, S. R. F. Sims, and J. F. Epperson, "Minimum squared error synthetic discriminant functions," *Opt. Eng.* **31**, 915-922 (1992).

[Vij92b] - B. V. K. Vijaya Kumar, F. M. Dickey, and J. M. DeLaurentis, "Correlation filters minimizing peak location errors," *J. Opt. Soc. Am. A* **9**, 678-682 (1992).

[Vij92c] - B. V. K. Vijaya Kumar, "Tutorial survey of composite filter designs for optical correlators," *Appl. Opt.* **31**, 4773-4801 (1992).

[Wol83] - S. Wolfram, "Statistical mechanics of cellular automata," *Rev. Mod. Phys.* **55**, 601-644 (1983).

[Yar86] - L. P. Yaroslavsky, "Applied problems of digital optics," in Advances in Electronics and Electron Physics, Vol. 66, P. W. Hawkes Ed., Academic Press, Orlando (1986).

[You82] - D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1-Theory," *IEEE Trans. Med. Imaging* **MI-1**, 81-94 (1982).

[Yu92] - F. T. S. Yu and S. Jutamulia, Optical Signal Processing, Computing and Neural Networks, *Wiley Series in Microwave and Optical Engineering*, K. Chang Ed., Wiley, New York (1992).

UNIVERSIDAD DE CAROLINA

Linda esta Memoria el día 30 de marzo de 1995 en la Facultad de Física, ante el siguiente Tribunal:

PRESIDENTE  
José R. de la Cruz

VOCALES

[Signature] [Signature] [Signature]

fue calificada de APTO CON CALDE POR  
UNANIMIDAD

-198-