Optical Pattern Recognition

Bahram Javidi
Philippe Réfrégier
Chairs/Editors

14-17 June 1994
La Rochelle, France

Optical Pattern Recognition

Bahram Javidi
Philippe Réfrégier
Chairs/Editors

14-17 June 1994
La Rochelle, France

Conference Committee

Conference Chairs

Bahram Javidi, University of Connecticut (USA)
Philippe Réfrégier, Thomson-CSF/LCR (France)

Program Committee

P. Chavel, Institut d'Optique/Orsay (France)
B. Javidi, University of Connecticut (USA)
W. T. Rhodes, Georgia Institute of Technology (USA)
Ph. Réfrégier, Thomson-CSF/LCR (France)

Session Chairs

H. H. Arsenault, Université Laval (Canada)
D. P. Casasent, Carnegie Mellon University (USA)
P. Chavel, Institut d'Optique/Orsay (France)
J.-P. Huignard, Thomson-CSF (France)
A. D. Maruani, Ecole Nationale Superieure des Telecommunications (France)
D. Psaltis, California Institute of Technology (USA)
W. T. Rhodes, Georgia Institute of Technology (USA)

Introduction

Optical pattern recognition is a title given to a wide range of optical systems, technologies, and their algorithms that are aimed at providing a solution to the difficult task of image recognition, tracking, and parameter estimations. There are no general solutions for these tasks, and the practical systems used are often a compromise between hardware feasibility and algorithm performance. This is true for both numerical and optical realizations.

The ability of optics to make real-time correlation is now well known. However, it is only recently that the critical components for the interface between optics and electronics have become available. This new important step has led to two domains of intensive investigation. The first one is the improvement of the optical architectures for making more complex operations and more compact systems. The second one is related to algorithmic developments for achieving high-performance processors. From these studies it is shown that conventional matched filters and correlators do not perform as well as the newly developed linear and nonlinear filtering and neural networks techniques that have been demonstrated with success.

The goal of the workshop and of this book is to present and discuss this recent progress and to analyze the new relevant directions of research. In the first section, global filtering algorithms are analyzed. Next, applications and problems for image processing are discussed from both numerical and optical points of view. In the third section, optical architectures for pattern recognition with filtering techniques are covered; in the fourth, neural networks are presented. The volume is then concluded with analysis of the use of diffractive optics for optical pattern recognition.

We would like to thank the authors for their dedicated work and the SPIE staff for their help in preparing this volume.

Bahram Javidi
Philippe Réfrégier

Contents

Filtering Algorithms

Correlation, Functional Analysis and Optical Pattern Recognition

Fred M. Dickey
Mark L. Yee
K. Terry Stalker

Sandia National Laboratories
Albuquerque, NM 87185-0355

ABSTRACT

Correlation integrals have played a central role in optical pattern recognition. The success of correlation., however, has been limited. What is needed is a mathematical operation more complex than correlation. Suitably complex operations are the functionals defined on the Hilbert space of Lebesgue square integrable functions. Correlation is a linear functional of a parameter. In this paper, we develop a representation of functionals in terms of inner products or equivalently correlation functions. We also discuss the role of functionals in neural networks. Having established a broad relation of correlation to pattern recognition, we discuss the computation of correlation functions using acousto-optics.

1. INTRODUCTION

Pattern recognition is an extremely complex field. At present, a general solution to the problem of recognizing an arbitrary object in an arbitrary background does not exist. In fact, there are very few good solutions to restricted pattern recognition problems. Many approaches to the problem have been suggested and have been or are continuing to be investigated.[1]

The evaluation of a correlation integral is a popular approach to optical pattern recognition.[2] The initial interest in correlation probably stems from the applications of the matched filter in radar signal processing and communications theory. Further, there is a natural affinity between correlation and the Fourier transformation relation in optical systems.[3,4]

Considerable research effort has been directed toward the analysis of the performance of various correlation filters. There has also been significant effort directed at the development of optical correlation systems.

In normalized form, correlation does provide a rudimentary pattern recognition operation in that it measures mathematical "correlation" between objects.[2,4,5,6,7,8] However, one would readily recognize that correlation falls short of the pattern recognition capabilities of the human mind. This clearly implies that we need a mathematical operation more complex than correlation for general pattern recognition. Perhaps, a neural network can provide the required complexity.

One such mathematical operation is called a functional. We define a functional as a numerical (real or complex) function defined on a set of functions. A correlation or inner product is then just a linear functional. In general, a functional is a nonlinear mapping of a function into the real (or complex) numbers. Fortunately, a very broad class of functionals can be evaluated as a sum of products of inner products or correlations. Thus, correlation can be viewed as a fundamental computation in a broad sense.

In the following section, we review normalized correlation and introduce the concept of normalized quadratic filters. The concept of generalized correlation is introduced in terms of functionals in Section 3. Further, since perceptron neural networks are functionals on a vector space, we discuss the relation of correlation to neural networks.

Having argued the broad relation of correlation to pattern recognition, we discuss the implementation of the computation of correlation functions using an acousto-optic correlator in Section 4. Ultimately, the success of optical pattern recognition depends upon the efficiency and efficacy of the optical components and systems.

2. NORMALIZED CORRELATION AND QUADRATIC FILTERS

The pattern recognition problem can be said to be one of discriminating between different classes of objects or signals. Frequently, the problem is one of recognition (detection) of a specific object (signal) in the presence of all other possible objects (signals). An example is the detection of a specific vehicle type in the presence of other vehicle types, terrain features, and various man-made objects. This is a very difficult problem and a general solution may not be achievable. The major problem is one of suitably defining "all other objects." The problem is considerably simplified when the number of objects to be discriminated is small. In the case of a small number of well-defined objects, there may be several relatively simple approaches to the problem.

The interest in correlation filters is, in the main, associated with the problem of recognizing an object in an arbitrary background. The problem is, then, to determine a recognition system that will cause output values for all inputs to be recognized as the object of interest to fall within a specified range. For all other objects, the output values fall outside this range. This may sound simple, but the problem is one of defining all inputs that are equivalent to the object of interest. For example, how big must the tail on the letter "Q" be before it is discriminated from the letter "O"; is a white circle (or other shapes) contained in a white square?

In general, the literature on the technique of using correlation filters for optical pattern recognition assumes or implies that pattern recognition can be effected by computing the magnitude of the correlation between the object in question and a reference (filter) function. However, it is clear that a nonlinear operation is required to accomplish pattern recognition. A naive approach to the nonlinearity is peak detection.

A more sophisticated approach is to normalize the correlation function or inner product. Normalization achieves intensity invariance.

The treatment of normalization and intensity invariance in the literature is relatively sparse. Goodman[4] suggested the normalized matched filter as a means of achieving character recognition. Duda and Hart[9] suggested normalization to effect intensity invariance. Dickey and Romero[2,5] have discussed normalized correlation at length. In their papers, they use the normalized form to evaluate partial information filters and composite filters.

The use of correlation filters in pattern recognition is essentially an inner product between two functions: the object and reference functions. These functions may be considered vectors in a Hilbert space.[10] The normalization of the inner product (correlation integral) defines a unique angle between the reference and the object function. It is this angle that provides a measure of similarity between the object and reference functions.

$$c(y-x_o)=\left|\int h^*(x)s_h(x)f(x+y-x_o)dx\right|^2 \tag{1}$$

where

$$f(x)=\text{input (object) function}$$

$$h^*(-x)=\text{filter impulse response}$$

$$s_h(x)=\begin{cases}1, x\in \text{ support of } h(x)\\ 0, \text{ otherwise,}\end{cases}$$

$$x_o=\text{coordinate that maximizes the integral.}$$

The indicator function has the properties

$$s_h(x)h(x)=h(x), \tag{2}$$

$$s_h^2(x)=s_h(x). \tag{3}$$

Applying the Cauchy-Schwarz inequality to Eq. (1) gives

$$c(y-x_o)\le\int\left|f(x+y-x_o)\right|^2 s_h(x)dx\int\left|h(x)\right|^2 dx. \tag{4}$$

The preceding suggests a normalized correlation function given by

$$c(y-x_o)=\frac{\left|\int h^*(x)f(x+y-x_o)dx\right|^2}{\int\left|f(x+y-x_o)\right|^2 s_h(x)dx\int\left|h(x)\right|^2 dx}\leq 1. \tag{5}$$

It is a further property of the Cauchy-Schwarz inequality that the equality in Eq. (5) is obtained at $y - x_0 = 0$ and only if

$$h(x)=\lambda f(x)s_h(x). \tag{6}$$

For the case $y = x_0$, Eq. (5) is equivalent to

$$\hat{c}(0)=\cos^2(\theta) \tag{7}$$

where θ is the Hilbert space angle between h (x) and the restriction of f(x) to the support of h(x). It is interesting to note that the normalized inner product has been proposed[11] as a measure of similarity between vectors in a vector approach to automatic text retrieval.

It is the form of Eq. (5) and the λ in Eq. (6) that effects intensity invariance. Thus, it is the classical matched filter associated with white noise that maximizes the normalized correlation given by Eq. (5). If the filter used to identify an object function f(x) is not the matched filter, a normalized correlation value less than one is obtained. In this case, there are many functions different from f(x) that give the same correlation value. The difference (e.g., mean square difference) between these functions and the object function must approach zero as the normalized correlation function approaches one. Examples of normalized correlation are presented in Section 4.

For any filter other than the matched filter, the normalized correlation is less than one when the object function for which the filter was made is the input function. Further, there is always a function that gives a normalized correlation value of one. This function is just the filter impulse response. For an arbitrary filter, there are many functions that produce a correlation value equal to or greater than that produced by the object function. Gheen, Dickey, and DeLaurentis[8] discuss the arbitrary approximation of Bayes classifiers by performing a series of correlations. Javidi, Refregier and Willet[12] describe a constraint under which normalized correlation is optimum.

An extension of the correlation (inner product) approach to pattern recognition is the quadratic form (filter) given by

$$Q = |(\psi, A\psi)|^2. \tag{8}$$

where ψ is the input function restricted to the support of the target and A is a linear operator. Gheen [13, 14] has investigated the invariance properties of quadratic filters.

The quadratic form given by Eq. (8) can be normalized using the Cauchy-Schwarz inequality giving

$$\hat{Q} = \frac{|(\psi, A\psi)|^2}{\|\psi\|^2 \|A\psi\|^2} \leq 1, \tag{9}$$

with equality obtained iff

$$A\psi = \lambda\psi. \tag{10}$$

Thus, we can obtain the equality in Eq. (9) for a set of normal image functions ϕ_i, if we define the operator by

$$A\psi = \sum \lambda_i (\phi_i, \psi)\phi_i, \tag{11}$$

or

$$A = \sum \lambda_i (\phi_i, \cdot)\phi_i. \tag{12}$$

If the ϕ_i are also orthogonal and the λ_i are real, we have a Hermitian operator. However, if $\lambda_i = \lambda_j$ for some i, j, we have a multiple eigenvalue. Multiple eigenvalues are not desirable in that all linear combination of ϕ_i for which $\lambda_i = \lambda_j$ satisfy the equality in Eq. (9).

It appears that there is no requirement that the ϕ_i be orthogonal. However, we would want to retain the requirement that $\lambda_i \neq \lambda_j$ for all i, j. Thus, the quadratic filter defined by Eq. (9) and Eq. (12) defines a composite filter system with the major computations consisting of inner products (correlations). There, however, does not appear to be any increase in computational efficiency of the quadratic filter over the corresponding set of correlations. In fact, there is more information available in the set of individual correlations.

3. GENERALIZED CORRELATION AND FUNCTIONALS

Normalized correlation and normalized quadratic filters discussed in the last section are relatively simple functionals. Although these functionals provide a degree of pattern recognition, it is clear (see Reference 8) that considerably more complex functionals are required for general pattern recognition.

In this section, we assume that the ability to compute an arbitrary functional on the input function is sufficient for a large class of pattern recognition problems. A functional is the transformation of a vector space X into the real numbers R,

$$F: X \rightarrow R. \qquad (13)$$

The vector space X will typically be a function space. The Hilbert (function) space of Lebesgue square integrable functions (L_2) is an adequate space for modeling input (image) functions.

In the following, we show that arbitrary nonlinear functionals can be computed in terms of sums of products of linear functionals (inner products). That is, we can compute any pattern recognition functional as a sum of products of correlations. We develop this result from two approaches. One approach uses the theory of Volterra functionals. The other approach is based on the theory of perceptron neural networks.

3.1 Volterra Functionals

Volterra functionals (Volterra filters) have been applied to the nonlinear signal processing and nonlinear systems analysis.[15,16,17] Gheen[18] has investigated the invariant properties of Volterra filters for pattern recognition. Volterra[19] introduced regular homogeneous functionals of degree n,

$$F_n[f(x)] = \int\int\cdots\int k_n(x_1, x_2 \cdots x_n) f(x_1) f(x_2) \cdots f(x_n) dx_1 dx_2 \cdots dx_n. \qquad (14)$$

Analogous to polynomials, he then defines regular functionals of degree n as

$$G_n[f(x)] = k_o + F_1[f(x)] + F_2[f(x)] + \cdots F_n[f(x)]. \qquad (15)$$

The finite sum in Eq. (15) is readily extended to include an infinite number of terms. It is important to determine what class of functionals can be represented by infinite sums of terms given by Eq. (14). For our purposes, the answer is contained in a theorem given by Volterra.

Theorem: Every functional continuous in the field of continuous functions can be represented by

$$G[f(x)] = \lim_{n\to\infty}\Big[k_{n,o} + \sum_{1}^{r_n}\int\int\cdots\int k_{n,p}(x_1,x_2\cdots x_p)$$

$$f(x_1)f(x_2)\cdots f(x_p)dx_1dx_2\cdots dx_p\Big], \tag{16}$$

where the functions $k_{n,p}$ $(x_1, x_2, \ldots x_p)$ are continuous functions.[19]

We only need to consider continuous because continuous functions are dense in L_2. The limit in the theorem can be appreciated by considering the quadratic functional given by

$$H_2[f(x)] = \int f^2(x)dx. \tag{17}$$

This functional cannot be expressed by a second degree functional of Eq. (14) without the use of distributions.[15] The Dirac delta function allows one to write Eq. (17) as

$$H_2[f(x)] = \int\int \delta(x_1 - x_2)f(x_1)f(x_2)dx_1dx_2. \tag{18}$$

Palm and Poggio[15] give a theorem corresponding to the above theorem that eliminates the limit process by employing distributions.

In general, we can expand the $k_{n,p}$ functions in terms of one-dimensional basis functions giving

$$k_{n,p}(x_1,x_2\cdots x_p) = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_p}\alpha_{i_1,i_2-i_p}\,\psi_{i_1}(x_1) \tag{19}$$

$$\psi_{i_2}(x_2)\cdots\psi_{i_p}(x_p),$$

where the ψ_i are the appropriate basis functions. If we substitute Eq. (19) in Eq. (16), we obtain an expression for the functional in Eq. (16) as a sum of products of inner products (equivalently: linear functionals, correlations). Clearly, examining Eq. (18) and Eq. (16) leads one to the conclusion that the computation of a general functional would require a large number of correlations.

It should be noted that Eq. (16) may not be the most compact form for expressing an arbitrary functional. For example, normalized correlation given by Eq. (5) is a continuous functional and can be expressed in the form of Eq. (16). Analogous to ordinary functions, functionals can be expressed in a functional Taylor series.[20] The terms in the Taylor series are determined using the Frechet derivative.

3.2 Neural Networks

Perceptron neural networks compute functionals on an input vector. [21-27] Cybenko[21], Hornik, Stinchcomb and White[22], and Hornik[23] show that a perceptron neural network with a single hidden layer can approximate any continuous function with arbitrary precision. DeLaurentis and Dickey[27] show that a network with two hidden layers with nonnegative linear combinations and compositions of excitatory and inhibitory response functions uniformly approximate arbitrary nonnegative continuous functions.

Perceptron neural networks can also be represented as a sum of products of inner products (linear functionals, correlations). For simplicity, we consider single hidden layer networks. These networks compute functions (functionals) given by

$$G(x) = \sum_{i}^{N} \sigma\left[(y_i, x) + b_i\right], \tag{20}$$

where σ is an activation function, x is the input vector, y_i is a vector defining the weights to hidden node i, b_i is a bias and $(\cdot,\cdot)$ denotes an inner product. Based on the Weierstrass approximation theorem, we can uniformly approximate continuous functions with polynomials. Also, continuous functions are dense in L_2. Hence, we can approximate the activation function by a polynomial,

$$\sigma(z) = \sum^{Q} B_p z^p. \tag{21}$$

Substituting Eq. (21) in Eq. (20) gives

$$G(x) = \sum_{i}^{N} \sum_{p}^{Q} B_p\left[(y_i, x) + b_i\right]^p. \tag{22}$$

Thus, we can express arbitrary neural network functionals as sums of products of correlations. As in the previous subsection, it is clear that, in general, a large number of correlations are required. It should be noted that N in Eq. (20) and Eq. (22) and Q in Eq. (21) are determined by the complexity of the functional and the precision required. Generally, without other constraints, the size of a neural network (number of nodes) will be quite large. Attempts to bound neural networks with respect to a given problem have met with little success. Koiran [29] has derived a bound on the number of hidden layer nodes that is an exponential function of the dimension of the input vector.

3.3 Functionals and Pattern Recognition

In this section, we have shown a strong relation between correlation and pattern recognition. This is based on the assumption that the computation of an arbitrary functional should be sufficient for pattern recognition problems. However, these results indicate that the amount of computation required for an arbitrary functional is very large.

It is thus clear that the real problem in optical pattern recognition is in determining what functional to calculate. Once this is done, correlation may play a large role in the computation.

4. ACOUSTO-OPTIC CORRELATOR

In Section 3 we pointed out that large numbers of correlations can be instrumental in expressing arbitrary functionals. A natural concern then arises about the ability to compute large numbers of correlations in a short time, since correlation is computationally intensive. Historically this is where optical correlators have been employed, due to the ability of optical lenses to perform two-dimensional Fourier transformations.[4] (However, we note that digital electronics are advancing to comparable speeds.[30]) In this section we briefly review the method of acousto-optic correlation, which performs space-domain correlation instead of traditional frequency-domain correlation. We also present some examples of normalized correlation, both simulated and actual.

Current real-time frequency-domain optical correlators are unable to produce the full complex values necessary for Fourier-based correlation. Approximate values (partial information filters) are typically substituted for the full complex values. This introduces some error into the final correlation result. We do not wish to debate the implications of such errors for traditional approaches to pattern recognition using a correlation followed by a thresholding operation. Rather, it is of interest to consider the impact such errors would have for the computation of arbitrary functionals via sums of products of correlations as shown in Section 3. In this situation, it is obviously desirable to reduce the error in correlations to a minimum.

Unlike the frequency-domain approach, a space-domain correlation does not require complex values. (Correlation of real functions will only require real values. Correlation of complex functions can be performed by processing real and complex parts separately.) Thus a space-domain correlator can in principle produce correlation results with less error than a frequency-domain correlator, resulting in arbitrary functionals closer to the desired form. We now highlight the fundamentals of our working acousto-optic correlator, which has been detailed elsewhere.[31]