Basic properties of nonlinear global filtering techniques and optimal discriminant solutions

Philippe Réfrégier, Vincent Laude, and Bahram Javidi

The basic properties of nonlinear global filtering techniques are analyzed. A nonlinear processor for pattern recognition that is optimum in terms of discrimination and that is tolerant of variations of the object to be recognized is presented. We compare this processor with power-law and nonlinear joint transform correlators.

1. Introduction

Nonlinear joint transform correlators[1] (JTC's) have been shown to be attractive for pattern-recognition applications. However, their basic properties in terms of signal processing and pattern recognition are still an intensive subject of investigations.

For optical correlation, different criteria have been proposed to characterize the filter performances.[2] Among them, it has been shown[2] that some of the most interesting are related to noise robustness of the filter and sharpness of the correlation function. Furthermore, the importance of finding trade-offs among different criteria is now well established.[2,3] It has been shown that this approach[3] leads to useful filters and figures of merit. However, up to now, the discrimination capabilities of linear filters were optimized indirectly by the minimization of the sharpness of the correlation function,[3] the energy of the correlation function with false objects (that is, objects to be rejected), or background models to be discriminated against.[4]

Alternatively, nonlinear JTC's have been shown to be very discriminant with good correlation performance.[1] Optimal methods for discrimination capabilities of the processor for which *a priori* knowledge of the false objects or of the background[4,5] is not needed result in nonlinear filtering techniques. Furthermore, the optimal processor introduced in Ref. 5 presents strong analogies with nonlinear JTC's.[1]

In this paper we analyze the basic properties of nonlinear global filtering (NGF) techniques. We introduce new definitions that allow us to derive naturally the different solutions introduced previously for a nonlinear JTC.[1] We thus can emphasize the properties that are satisfied in each case and then generalize and enhance the solution given in Ref. 5. We also provide numerical simulations results that emphasize analogies with and differences from other nonlinear JTC's.

In Section 2, we introduce our notations and recall the main results of the basic problem of discrimination. We introduce the basic definitions and properties of NGF techniques in Section 3. In Section 4, we derive the optimal nonlinear filtering method for discrimination. In Section 5, we illustrate the relevance of this approach with numerical simulations. In Section 6, we summarize our results.

2. Background

In the analysis, monodimensional notations are used for simplicity with no loss of generality. Let $\mathbf{r}$ and $\mathbf{s}$ denote, respectively, the reference and the input images. The output of a processor is denoted as $\mathbf{c}$. All these images are assumed to be sampled on N pixels, and their values at location t are, respectively, $r(t)$, $s(t)$, and $c(t)$, where t varies between 0 and $N-1$. $\mathbf{r}$, $\mathbf{s}$, and $\mathbf{c}$ are thus vectors in $\mathbf{C}^N$, where $\mathbf{C}$ is the set of complex numbers. Their Fourier transforms are denoted, respectively, as $\hat{\mathbf{r}}$, $\hat{\mathbf{s}}$, and $\hat{\mathbf{c}}$, or $\hat{r}(k)$, $\hat{s}(k)$, and $\hat{c}(k)$ at frequency k. When the output of the processor, $\mathbf{c}$, can be written as a correlation between a filter $\mathbf{h}$ and the input image $\mathbf{s}$, we write

$$c(t) = [\mathbf{h} \circledast \mathbf{s}](t) = \sum_{t'} h^*(t')s(t+t'), \tag{1}$$

P. Réfrégier is with ENSPM, Signal and Image Laboratory, Domaine Universitaire de Saint-Jérôme, 13 397 Marseille cedex 20, France. V. Laude is with the Laboratoire Central de Recherches, Thomson-CSF 91404 Orsay (cedex) France. B. Javidi is with the Department of Electrical Engineering, University of Connecticut, Storrs, Connecticut 06269-3157.

Received 3 October 1994; revised manuscript received 6 March 1995.
0003-6935/95/203915-09$06.00/0.

or, in the Fourier domain, we write

$$\hat{c}(k) = \hat{h}^*(k)\hat{s}(k). \tag{2}$$

For simplicity, discrete notations are used below for both the image and the reference. Generalization to continuous Fourier transforms and notations can be obtained by the use of integrals instead of discrete summations. The translation operation T_τ applied to an image is defined by

$$\mathbf{g} = T_\tau \mathbf{f} \Leftrightarrow g(t) = f[\mathrm{Mod}_N(t - \tau)], \tag{3}$$

where Mod_N is the modulo N operation defined by

$$\mathrm{Mod}_N(t - \tau) = t - \tau + mN, \tag{4}$$

and where the integer m is defined such that

$$0 \leq t - \tau + mN < N.$$

When written in two-dimensional notation, Eq. (4) should be clearly understood as cyclic boundary conditions in both x and y directions.

It can be important to improve the discrimination capabilities of a filter between the object to be recognized and the objects to be rejected. For this purpose, a method discussed in Ref. 4 can be generalized. Let $\mathbf{y}^l$ denote (with $l = 1, \ldots, P$) P objects to be rejected. For improving the discrimination capabilities of a filter $\mathbf{h}$, a possible method consists of minimizing the energy of the correlation with the images $\mathbf{y}^l$, with the constraint that the correlation with the object $\mathbf{r}$ to be recognized is equal to a given value. In this case, the criterion to minimize is

$$D_E = \sum_{l=1}^{P} \sum_{k} |\hat{h}(k)|^2 |\hat{y}^l(k)|^2. \tag{5}$$

Let $\hat{S}^{(rj)}(k) = \Sigma_{l=1}^{P} |\hat{y}^l(k)|^2$ be the average spectral density of patterns to be rejected. An optimal filter for this purpose is given by[4]

$$\hat{h}(k) = \frac{\hat{r}(k)}{\hat{S}^{(rj)}(k)} \cdot \tag{6}$$

Then, optimizing the discrimination capabilities leads to the consideration of a matched filter with a spectral density of noise equal to $\hat{S}^{(rj)}(k)$. Indeed, realizations of noise can be considered as images to be rejected. In that case, $\Sigma_{l=1}^{P} |\hat{y}^l(k)|^2$ will converge to an ensemble average and then to the true spectral density of the noise.

However, the general question is how to infer the appropriate images $\hat{\mathbf{y}}^l$ that correspond to objects to be rejected is not obvious. The same problem arises for the matched filter for which it is assumed that the spectral density of the noise $\hat{S}$ is known. This is not the case in general in image processing, in which, in contrast to radar processing, it is difficult to estimate the spectral density of the noise. Then an important question in the context of pattern recognition is the determination of an appropriate model for the spectral density $\hat{S}$. Furthermore, there is no reason to consider that the realizations of noise (or of the images to be rejected) are obtained with a temporal stationary density probability law. For example, this is clear if images of the ground obtained from an airplane are considered. If the background is modeled by a noise, its spectral density can be very different for the sea, mountains, or fields.

A new approach to overcome this drawback was recently proposed in Ref. 5. The main idea is to process the input image adaptively. Indeed, a discriminant filter for the input image is obtained by the minimization of the energy of the correlation function between the filter and the input image:

$$E_S = \sum_{k} |\hat{h}(k)|^2 |\hat{s}(k)|^2. \tag{7}$$

Of course, minimization of the criterion of Eq. (7) alone leads to the null filter. However, if this minimization is performed under the constraint that there still is a correlation peak at the location of the target, then it is shown below that this approach can be attractive if correctly regularized. The new point here is that E_S is not a fixed value as, for example, D_E but is dependent on the input image $\mathbf{s}$. As we show in Section 4, this leads to a nonlinear processor.

3. Definitions and Properties of Global Filtering Operations

To clarify the nonlinear filtering operations that have been analyzed in the past or that are considered below, we review and propose an analysis of the properties of such techniques.

Let us consider a general filtering processor. The output that is denoted as $\mathbf{c}$ is a function of both the reference $\mathbf{r}$ and the input image $\mathbf{s}$. We do not consider a multireference problem here. Then, formally, we can write

$$\mathbf{c} = \mathscr{H}[\mathbf{r}, \mathbf{s}], \tag{8}$$

or, equivalently,

$$c(t) = \mathscr{H}[\mathbf{r}, \mathbf{s}](t). \tag{9}$$

Equation (9) means that, for every location t, the complex number $c(t)$ is a function of vectors $\mathbf{r}$ and $\mathbf{s}$. The input is $\mathbf{s}$, the output is $\mathbf{c}$, and $\mathbf{r}$ is a model of the target that is being looked for. Equation (8) can be alternatively written in the Fourier domain:

$$\hat{\mathbf{c}} = \hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}], \tag{10}$$

or, equivalently,

$$\hat{c}(k) = \hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k), \tag{11}$$

where k is the spatial frequency, and $\hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$ are, respectively, the Fourier transforms of $\mathbf{r}$ and $\mathbf{s}$. The equivalence of Eqs. (8) and (11) is obvious, given that the same information is contained in the image and its Fourier transform, and is merely a coordinate transformation in phase space.

A global filtering is said to be linear if, for every $\mathbf{r}$, $\mathbf{s}$, $\mathbf{s}'$, and complex number λ, the following properties hold:

$$\mathscr{H}[\mathbf{r}, \lambda \mathbf{s}](t) = \lambda \mathscr{H}[\mathbf{r}, \mathbf{s}](t), \tag{12a}$$

$$\mathscr{H}[\mathbf{r}, \mathbf{s} + \mathbf{s}'](t) = \mathscr{H}[\mathbf{r}, \mathbf{s}](t) + \mathscr{H}[\mathbf{r}, \mathbf{s}'](t), \tag{12b}$$

that is, the global filtering is linear with the input scene. This is the case with classical convolution filtering that is widespread in optical processing. It is clear that, in general, only the squared modulus of the output of a coherent optical correlator can be obtained, so that we have to consider linearity before the squaring operation in that case. The linearity conditions of Eqs. (12) imply that the global filtering can be written as

$$c(t) = \sum_{t'} A[\mathbf{r}](t, t')s(t'), \tag{13}$$

where $A[\mathbf{r}](t, t')$ is a matrix element.

It is interesting to note that we do not require that the global filtering be linear with the reference image, which we would then write as

$$\mathscr{H}[\lambda \mathbf{r}, \mathbf{s}](t) = \lambda \mathscr{H}[\mathbf{r}, \mathbf{s}](t), \tag{14a}$$

$$\mathscr{H}[\mathbf{r} + \mathbf{r}', \mathbf{s}](t) = \mathscr{H}[\mathbf{r}, \mathbf{s}](t) + \mathscr{H}[\mathbf{r}', \mathbf{s}](t), \tag{14b}$$

for every $\mathbf{r}$, $\mathbf{r}'$, and $\mathbf{s}$. This has been a confusion in the past in denoting filters that do not satisfy Eqs. (14) as nonlinear filters.

For some applications, although linearity is not necessary, it can be useful to satisfy a weaker condition, that is, for every $\mathbf{r}$, $\mathbf{s}$, and complex number λ,

$$\mathscr{H}[\mathbf{r}, \lambda \mathbf{s}](t) = g(\lambda) \mathscr{H}[\mathbf{r}, \mathbf{s}](t), \tag{15}$$

where $g(\lambda)$ is a function of λ. This means that a variation of the illumination of the input image does not alter the shape of the output of the global filtering, even though it modifies the absolute values. It is then easy to show that $g(\lambda)$ must be of the form

$$g(\lambda) = m^{\alpha} \exp(in\phi), \tag{16}$$

where m is the modulus of λ and ϕ is its phase, and where α is a real number and n is an integer. The proof of this point is given in appendix A.

In the same way, one could require that

$$\mathscr{H}[\lambda \mathbf{r}, \mathbf{s}](t) = g(\lambda) \mathscr{H}[\mathbf{r}, \mathbf{s}](t), \tag{17}$$

whatever the complex number λ, or one could require a even more stringent property,

$$\mathscr{H}[\lambda_r \mathbf{r}, \lambda_s \mathbf{s}](t) = g_r(\lambda_r) g_s(\lambda_s) \mathscr{H}[\mathbf{r}, \mathbf{s}](t), \tag{18}$$

whatever the complex numbers λ_r and λ_s. For example, the normalized correlation introduced in Ref. 6 corresponds to $g_r = g_s = 1$.

Let us now analyze the property of stationarity. A global filtering is said to be cyclostationary if, for every $\mathbf{r}$, $\mathbf{s}$, t, and τ, the following property holds:

$$\mathscr{H}[T_{-\tau}\mathbf{r}, \mathbf{s}](t) = \mathscr{H}[\mathbf{r}, T_{\tau}\mathbf{s}](t) = T_{\tau} \mathscr{H}[\mathbf{r}, \mathbf{s}](t). \tag{19}$$

If the global filtering is linear and cyclostationary, then Eq. (13) becomes

$$c(t) = \sum_{t'} A[\mathbf{r}][\mathrm{Mod}_N(t - t')]s(t'), \tag{20}$$

which is nothing but a convolution of periodic signals. This clearly corresponds to the case of classical linear correlation filters.

It is clear from these definitions that nonlinearity and noncyclostationarity do not allow one to characterize global filtering techniques. For this reason, we now define some properties that allow us to obtain better insight into NGF operations. Below, we analyze nonlinear but cyclostationary global filtering.

Among the nonlinear cyclostationary global filtering there is a special class that we call local Fourier cyclostationary global filtering. It is particularly interesting to introduce this definition because most cases of NGF that have been implemented optically until now satisfy this property. This class of global filtering is easily defined in the Fourier domain. A global filtering is said to be local Fourier if the following property holds:

$$\hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k) = \hat{H}[\hat{r}(k), \hat{s}(k)]. \tag{21}$$

In other words, $\hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}]$ is no longer a general operator on vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$ that depends on $2N$ variables, but for every spatial frequency k the output of the processor is a function of only two complex numbers $\hat{r}(k)$ and $\hat{s}(k)$. Note that the letter H is then used in place of $\mathscr{H}$. Nonlinear joint Fourier correlation systems achieve, in general, local Fourier cyclostationary NGF's. One can note that the hypothesis of Fourier locality is not necessary from a pure signal point of view, but is often true in optical correlation when the processing is done in a Fourier plane.

Let us now analyze the example of the power-law correlator/nonlinear JTC introduced in Ref. 1. In that case, the output of the system is given by

$$\hat{c}(k) = |\hat{r}(k)\hat{s}(k)|^{\beta-1}\hat{r}(k)^*\hat{s}(k), \tag{22}$$

where β is a real number between 0 and 1. The binary JTC, or pure phase correlation, corresponds to $\beta = 0$. It is easy to check that this NGF is local Fourier, cyclostationary, and satisfies the weak linearity condition of Eq. (18).

But it is possible to derive the power-law JTC from some of the definitions that we have given above. Let us first consider that the global filtering satisfies the weak linearity condition of Eq. (18); we then have

$$\hat{\mathscr{H}}[\lambda_r \hat{\mathbf{r}}, \lambda_s \hat{\mathbf{s}}](k) = g_r(\lambda_r) g_s(\lambda_s) \hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k), \tag{23}$$

and, taking Eq. (16) into account,

$$\hat{\mathscr{H}}[\lambda_r \hat{\mathbf{r}}, \lambda_s \hat{\mathbf{s}}](k) = (m_r)^{\alpha_r}(m_s)^{\alpha_s}\exp(in_r\phi_r) \times \exp(in_s\phi_s) \hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k), \tag{24}$$

where m_r (m_s) is the modulus of λ_r (λ_s) and ϕ_r (ϕ_s) is the phase of λ_r (λ_s). α_r and α_s are real numbers, and n_r and n_s are integers.

Now let us include the condition of Fourier locality, which yields

$$\hat{H}[\lambda_r\hat{r}(k), \lambda_s\hat{s}(k)] = (m_r)^{\alpha_r}(m_s)^{\alpha_s}\exp(in_r\phi_r) \times \exp(in_s\phi_s)\hat{H}[\hat{r}(k), \hat{s}(k)]. \tag{25}$$

Because Eq. (25) must be satisfied for every λ_r, λ_s, $\hat{r}(k)$, and $\hat{s}(k)$, we can consider the following transformations:

$$\lambda_r \rightarrow \hat{r}(k),$$
$$\lambda_s \rightarrow \hat{s}(k),$$
$$\hat{r}(k) \rightarrow 1,$$
$$\hat{s}(k) \rightarrow 1, \tag{26}$$

from which we end up with

$$\hat{c}(k) = \hat{H}[\hat{r}(k), \hat{s}(k)] = B|\hat{r}(k)|^{\alpha_r}|\hat{s}(k)|^{\alpha_s} \times \exp(in_r\phi_{\hat{r}(k)})\exp(in_s\phi_{\hat{s}(k)}), \tag{27}$$

where $B = \hat{H}[1, 1]$ is a constant. In addition to the weak linearity condition of Eq. (18) and the assumption of Fourier locality that we have used, let us add the condition of stationarity. In the Fourier plane a translation operation of shift τ results in the modulation of the Fourier transform by an exponential factor $\exp(-2i\pi k\tau/N)$. For expression (27) to be cyclostationary, it is then required that

$$n_r = -1,$$
$$n_s = 1, \tag{28}$$

which, in turn, yields

$$\hat{c}(k) = B|\hat{r}(k)|^{\alpha_r-1}|\hat{s}(k)|^{\alpha_s-1}\hat{r}(k)^*\hat{s}(k). \tag{29}$$

This expression is similar to the definition of the power-law nonlinear JTC of Eq. (22), except that the exponent on the absolute values of the reference and the input image can be different. If we want them to be equal, we need a further property that can be written as

$$\mathscr{H}[\mathbf{s}, \mathbf{r}](t) = \mathscr{H}[\mathbf{r}, \mathbf{s}]^*(t), \tag{30}$$

that is, the global filtering is invariant, apart from a complex conjugation, under the exchange of the reference and the input images. We denote as joint Fourier such NGF's that are also cyclostationary. This property can be interesting for some applications. Indeed, let us consider, for example, a tracking problem. In this case $\mathbf{r}$ can be the image at time $n - 1$ and $\mathbf{s}$ can be the image at time n but one can wish that the result be invariant by the permutation of Eq. (30).

We have then shown that if a NGF satisfies the hypotheses of Fourier locality of Eq. (21), the weak linearity of Eq. (18), and is joint Fourier [as defined by Eq. (30)] and cyclostationary [Eq. (19)], then it is necessarily a power-law nonlinear JTC as given by Eq. (22).

4. Optimal Adaptive Discriminant Processors

We now derive a nonlinear processor for pattern recognition that is optimum in terms of discrimination and that is tolerant to variations of the object to be recognized.

We have seen that the output of a nonlinear processor can be written as

$$\hat{c}(k) = \hat{\mathscr{H}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k). \tag{31}$$

We assume that we can write this expression as

$$\hat{c}(k) = \hat{h}[\hat{\mathbf{r}}, \hat{\mathbf{s}}]^*(k)\hat{s}(k), \tag{32}$$

as this is equivalent to assuming that $\hat{c}(k) = 0$ if $\hat{s}(k) = 0$. Equation (32) can be written as a correlation in the object domain:

$$c(t) = \sum_{t'} h[\mathbf{r}, \mathbf{s}]^*(t')s(t + t') = (\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{s})(t). \tag{33}$$

To be precise we need to define the problem in more details. We consider that in the input image there is an object analog to the reference image that we denote as $\mathbf{r}'$ or $\hat{\mathbf{r}}'$ in the Fourier domain and that is translated in the input image at location τ. We can then write

$$\mathbf{s} = T_\tau\mathbf{r}' + \mathbf{b}, \tag{34}$$

which defines $\mathbf{b}$ as the background image, that is, everything in the input image that is not $\mathbf{r}'$. Furthermore, it is always possible to write

$$\mathbf{r}' = \mathbf{r} + \delta\mathbf{r}, \tag{35}$$

where $\delta\mathbf{r}$ appears as a perturbation of the reference image, and thus

$$\mathbf{s} = T_\tau(\mathbf{r} + \delta\mathbf{r}) + \mathbf{b}. \tag{36}$$

In Eq. (36) $\mathbf{r}$ and $\mathbf{s}$ are images that are known, but $\delta\mathbf{r}$ and $\mathbf{b}$ are unknown, together with the actual location τ of the target. However, we consider that we know more about $\delta\mathbf{r}$ than we do about $\mathbf{b}$. Indeed, the uncertainty $\delta\mathbf{r}$ can arise in practice from a distortion of the reference image or an acquisition noise. In any case, we consider that we have a model of the spectral density of $\delta\mathbf{r}$ that we denote as $\sigma^2(k)$. Note that we do not require a similar condition for the background image $\mathbf{b}$.

We want to derive a processor that has the following desired properties:

(i) It should yield a correlation peak at location τ that is close to a Dirac function,
(ii) It should be robust to the perturbation $\delta\mathbf{r}$,
(iii) It should be robust to the background $\mathbf{b}$.

With the model of Eq. (36), we can write the correlation of Eq. (33) as

$$(\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{s})(t) = (\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{r})(t - \tau) + (\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \delta\mathbf{r})(t - \tau) + (\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{b})(t). \quad (37)$$

Let us define the notation $E(\mathbf{f})$ as the total energy of the correlation of the filter $\mathbf{h}[\mathbf{r}, \mathbf{s}]$ with the image $\mathbf{f}$:

$$E(\mathbf{f}) = \sum_k |\hat{h}[\mathbf{r}, \mathbf{s}](k)|^2 |\hat{f}(k)|^2. \quad (38)$$

Let us start with condition (i). We can minimize the energy of the correlation function of the filter with the input scene $E(\mathbf{s})$:

$$E(\mathbf{s}) = \sum_k |\hat{h}[\mathbf{r}, \mathbf{s}](k)|^2 |\hat{s}(k)|^2, \quad (39)$$

under the constraint that

$$c(\tau) = (\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{s})(\tau) = c_o, \quad (40)$$

where c_o is a given constant. But obviously this constraint is unusable because the location τ of the target is unknown. However, using conditions (ii) and (iii) in Eq. (37), we should have approximately

$$(\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{s})(\tau) \approx (\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{r})(0). \quad (41)$$

For this to be true, conditions (ii) and (iii) must make the contributions to the correlation function at location τ of the perturbations $\delta\mathbf{r}$ and of the background $\mathbf{b}$ negligible. We then replace the constraint of Eq. (40) by

$$(\mathbf{h}[\mathbf{r}, \mathbf{s}] \circledast \mathbf{r})(0) = c_o, \quad (42)$$

which no longer involves the unknown location τ.

For condition (ii) to hold, we can minimize the energy of the correlation function with the perturbations $E(\delta\mathbf{r})$:

$$E(\delta\mathbf{r}) = \sum_k |\hat{h}[\mathbf{r}, \mathbf{s}](k)|^2 |\widehat{\delta r}(k)|^2, \quad (43)$$

This term accounts for the regularization of the solution, as discussed in Ref. 7. Instead of the exact power spectral density $|\widehat{\delta r}(k)|^2$ that might not be known exactly, we can still use a model $\sigma^2(k)$. For example, we could use $\sigma^2(k) = \langle |\widehat{\delta r}(k)|^2 \rangle$ where $\langle . \rangle$ represents the ensemble average over possible situations, or any of the stabilizing functionals discussed in Ref. 7. We then write Eq. (43) in the modified form:

$$E(\delta\mathbf{r}) = \sum_k |\hat{h}[\mathbf{r}, \mathbf{s}](k)|^2 \sigma^2(k). \quad (44)$$

Similarly, for condition (iii) to be true, we should require that the energy of the correlation function with the background $E(\mathbf{b})$ be minimized:

$$E(\mathbf{b}) = \sum_k |\hat{h}[\mathbf{r}, \mathbf{s}](k)|^2 |\hat{b}(k)|^2. \quad (45)$$

But we have no knowledge of the background image $\mathbf{b}$, so expression (45) cannot be actually used. However, remarking that

$$\mathbf{b} = \mathbf{s} - T_\tau(\mathbf{r} + \delta\mathbf{r}), \quad (46)$$

we can make use of the Minskowski inequality[8] because $\sqrt{E(\mathbf{f})}$ defines a norm of the image $\mathbf{f}$:

$$\sqrt{E(\mathbf{b})} \leq \sqrt{E(\mathbf{s})} + \sqrt{E(\mathbf{r})} + \sqrt{E(\delta\mathbf{r})}. \quad (47)$$

In this expression, $E(\mathbf{r})$ is the classical correlation-plane energy (CPE) criterion.[2]

We have shown that conditions (i), (ii), and (iii) can be fulfilled by the simultaneous minimization of $E(\mathbf{s})$, $E(\delta\mathbf{r})$, and $E(\mathbf{b})$, under the constraint of Eq. (42). We can perform this simultaneous minimization by finding the optimal trade-off[3,9] (OT) solutions for the previous criteria. But the problem of finding the OT solutions for $E(\mathbf{s})$, $E(\delta\mathbf{r})$, and $E(\mathbf{b})$ can be replaced by the problem of finding the OT solutions for $E(\mathbf{s})$, $E(\delta\mathbf{r})$, and $E(\mathbf{r})$, as is implied by the inequality of Eq. (47). We obtain these solutions by minimizing the following functional:

$$\Psi(\hat{\mathbf{h}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}]) = \alpha E(\mathbf{s}) + \beta E(\delta\mathbf{r}) + \gamma E(\mathbf{r}) - 2\lambda \sum_k \hat{h}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k)\hat{r}(k), \quad (48)$$

where α, β, and γ are real positive numbers to balance the trade-off between the three criteria,[10] and λ is a real number that has to be identified after minimization. To minimize $\Psi(\hat{\mathbf{h}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}])$, we impose

$$\frac{\partial \Psi(\hat{\mathbf{h}}[\hat{\mathbf{r}}, \hat{\mathbf{s}}])}{\partial \hat{h}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k)} = 0 \quad (49)$$

for each spatial frequency k. We thus have

$$\hat{h}[\hat{\mathbf{r}}, \hat{\mathbf{s}}](k)[\alpha|\hat{s}(k)|^2 + \beta\sigma^2(k) + \gamma|\hat{r}(k)|^2] = \lambda\hat{r}(k), \quad (50)$$

and then

$$\hat{c}(k) = \frac{\lambda \hat{r}^*(k)\hat{s}(k)}{\alpha|\hat{s}(k)|^2 + \beta\sigma^2(k) + \gamma|\hat{r}(k)|^2}. \quad (51)$$

Because a constant multiplicative factor is not relevant for our purpose, we can simplify Eq. (51) to

$$\hat{c}(k) = \frac{\hat{r}^*(k)\hat{s}(k)}{\sigma^2(k) + \alpha|\hat{s}(k)|^2 + \gamma|\hat{r}(k)|^2}. \quad (52)$$

It is obvious from Eq. (52) that this optimal processor is a global nonlinear filter, as it requires nonlinear transformation of the input-image Fourier transform. The nonlinearity clearly results from minimizing the criterion E_S defined by Eq. (39). This nonlinear processor is adaptive because the filter function is dependent on the input-image energy spectrum. If α is equal to 0, this filtering method becomes linear and is analogous to a linear OT filter.[3] Different but related results have also been obtained with different assumptions.[11]

It is also interesting to remark that this global nonlinear filtering satisfies the cyclostationarity and Fourier locality conditions. However, if we choose $\sigma^2(k)$ independently of $\hat{s}(k)$ it does not satisfy the weak linearity condition with the input scene of Eq. (15), although it is linear with the reference function. It is not joint Fourier either, except for the particular trade-off given by $\alpha = \gamma$:

$$\hat{c}(k) = \frac{\hat{r}(k)^*\hat{s}(k)}{\sigma^2(k) + \alpha|\hat{r}(k)|^2 + \alpha|\hat{s}(k)|^2} \cdot \qquad (53)$$

For the sake of simplicity, we consider this law below [the generalization to Eq. (52) is straightforward]. The NGF of Eq. (53) does not satisfy the weak linearity condition of Eq. (15). However, it is still possible to satisfy this property, for example, if both the reference and the input images are normalized with their energies:

$$\hat{c}(k) = \frac{\hat{r}(k)^*\hat{s}(k)}{\sigma^2(k) + \alpha \dfrac{|\hat{r}(k)|^2}{\sum_{k'} |\hat{r}(k')|^2} + \alpha \dfrac{|\hat{s}(k)|^2}{\sum_{k'} |\hat{s}(k')|^2}}, \qquad (54)$$

but now the Fourier locality condition is no longer valid.

5. Illustration with Numerical Simulations

Let us now illustrate the performance of the optimum nonlinear processor with numerical simulations performed on images of 256 × 256 pixels with gray levels. The reference image **r** is a car shown in Fig. 1 in an array of 64 × 64 pixels. The input image is shown in Fig. 2 and contains the reference object placed both on the top left-hand side and in the center of the input. On the bottom right-hand side, the reference object has been rotated by 7°. This composite image has been placed in the presence of a $1/f$ colored noise. The noise is additive except within the object in the center, where it is spatially disjoint or nonoverlapping.[12] As a result, the reference objects with overlapping noise are not clearly visible in Fig. 2 because the very low input signal-to-noise ratio (−7 dB). The convention of denomination for the different correlation peaks is shown in Fig. 3.

Numerical experiments were performed with the nonlinear filter given by expression (53). The spectral density $\sigma^2(k)$ is chosen to be a constant, equal to σ^2. It is clear that this choice for the spectral density of $\delta\mathbf{r}$, the variation added to the reference object, is a white envelope and is different from the actual additive $1/f$ noise spectral density, which also produces the background.

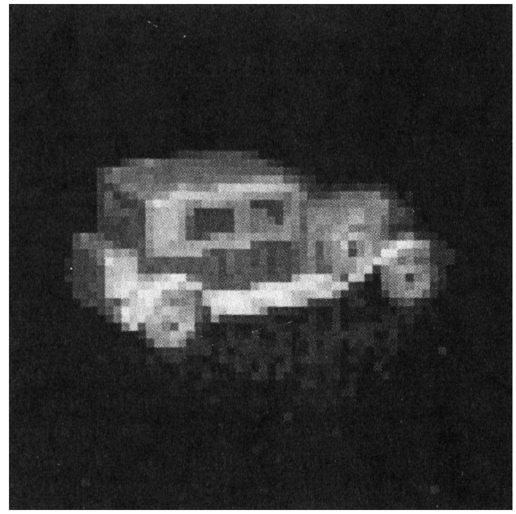

Fig. 1. 64 × 64 pixel reference image used for numerical simulations.

The projections of the normalized modulus squared of the correlation functions are shown in Fig. 4. More precisely, the following function is represented:

$$I(y) = \frac{\max_x |c(x, y)|^2}{\max_y \max_x |c(x, y)|^2}, \qquad (55)$$

where the maximum correlation intensity has been normalized to unity for all the plots. Figure 4 shows plots of $I(y)$ for values of σ^2 ranging from 10^{-4} to 10^7 and for $\alpha = 1$. For large values of σ^2, the NGF converges to the classical correlation $[\hat{c}(k) = \hat{r}^*(k)\hat{s}(k)/\sigma^2]$. The very low correlation-peak value for the reference object in the center of the input image is due to the low mean value of the object in comparison

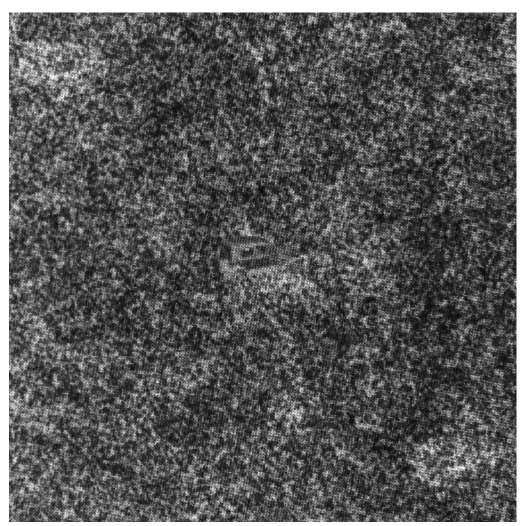

Fig. 2. 256 × 256 pixel input image used for numerical simulations that shows zero-mean additive $1/f$ noise that is disjoint to the central object.

with the mean value of the background noise, which is spatially disjoint with the object in the center.[12] The classical correlation is identical to the filtering with the matched filter for white noise. Its performance is very poor for the image in the context of spatially disjoint noise.[12]

As σ^2 increases, optimum detection changes between the three inputs as follows: (1) detection in disjoint noise appears optimal in Fig. 4(b), (2) detection in additive noise appears optimal in Figs. 4(b) and 4(c), and (3) detection for the rotated input in additive noise appears optimal in Figs. 4(d) and 4(e). In general, when σ^2 decreases, it can be seen that the correlation peak that corresponds to the car located at the center of the input image increases. The nonlinear operation allows one to be less sensitive to this situation of disjoint noise. When σ^2 becomes smaller, the correlation peaks become sharper. However, the background level increases as a consequence of the decreasing of the regularization term σ^2.

In Fig. 5 we show the results obtained with the power-law correlator[1] using the NGF of Eq. (22). When $\beta = 1$, the classical correlation is obtained. The pure phase correlation or binary nonlinear JTC is obtained with $\beta = 0$. One can see that the best results correspond to values of β approximately equal to 0.5, as shown in Fig. 5(d). We observe that this NGF leads to a higher level of background in the correlation plane. However, it is not clear whether that higher level of background can have a strong influence in the pattern-recognition or signal-processing task.

In order to understand better the differences and analogies between the different NGF techniques in

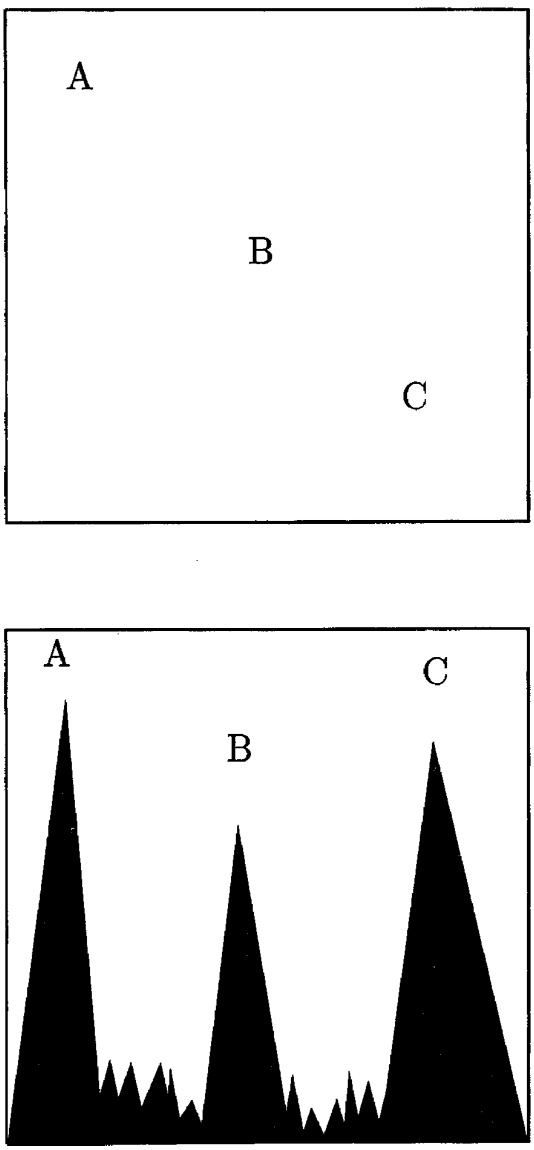

Fig. 3. Convention of denomination of the different objects present in the input image and the corresponding correlation peaks.

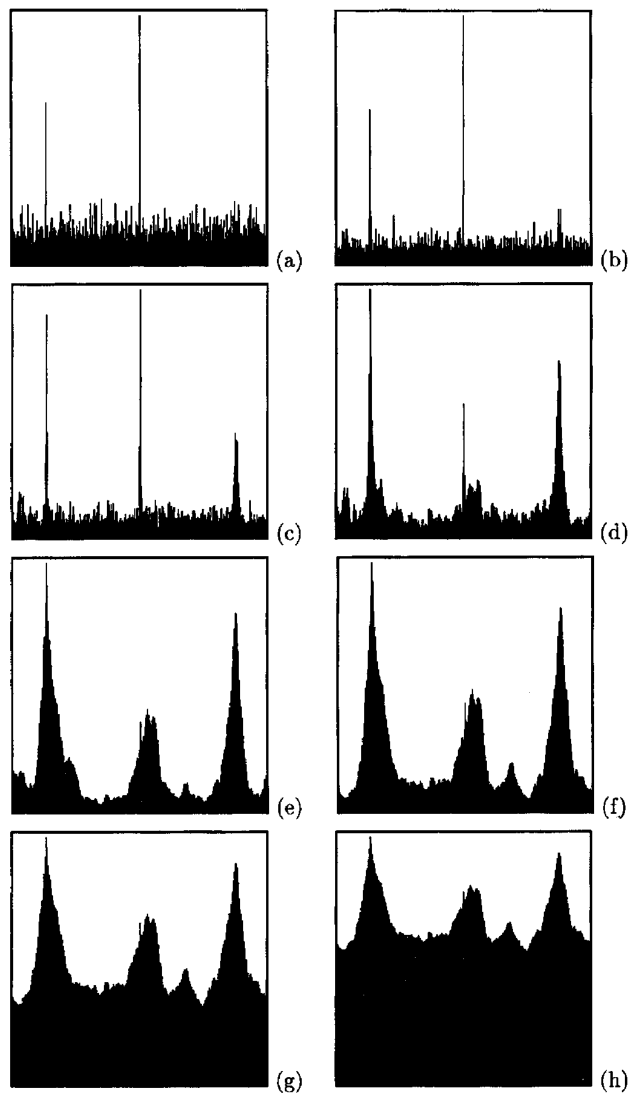

Fig. 4. Normalized projection of the modulus squared of the correlation functions for the NGF of Eq. (53) for $\alpha = 1$. This shows the performance trade-off in the selection of the following values of σ^2: (a) $\sigma^2 = 10^{-4}$, (b) $\sigma^2 = 10^{-2}$, (c) $\sigma^2 = 10^{-1}$, (d) $\sigma^2 = 1$, (e) $\sigma^2 = 10^2$, (f) $\sigma^2 = 10^3$, (g) $\sigma^2 = 10^4$, (h) $\sigma^2 = 10^7$.

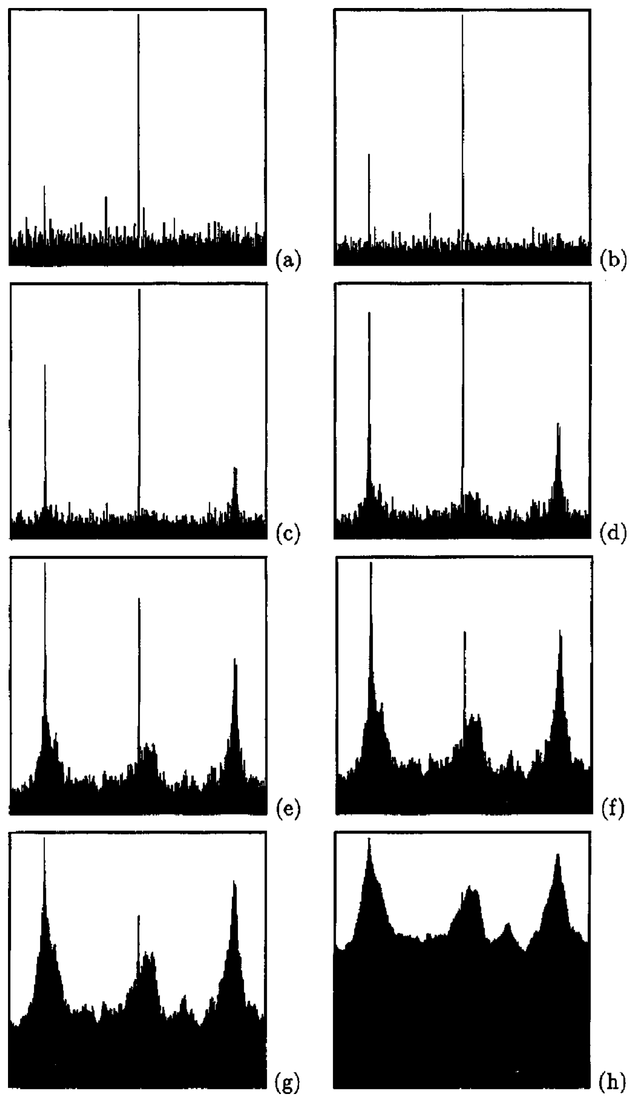

Fig. 5. Normalized projection of the modulus squared of the correlation functions with the power-law correlator/nonlinear JTC of Eq. (22).[1] This shows the performance trade-off in the selection of the following values of β: (a) $\beta = 0.0$, (b) $\beta = 0.2$, (c) $\beta = 0.4$, (d) $\beta = 0.5$, (e) $\beta = 0.6$, (f) $\beta = 0.7$, (g) $\beta = 0.8$, (h) $\beta = 1.0$.

the presence of noise, further theoretical investigations are needed.

6. Conclusion

We have analyzed the basic properties of NGL techniques. This analysis allowed us to designed a processor that is optimum in terms of discrimination and input-noise robustness and to understand better the basic assumptions that lead to the power-law correlator/nonlinear JTC previously introduced in the literature.[1] This optimum processor is a nonlinear filter that can be implemented by a nonlinear JTC and yields a new theoretical insight into obtaining optimal nonlinear transformations. Computer simulations have illustrated the performance of the processor for noisy and distorted objects in the presence of both overlapping and nonoverlapping input noise. Further studies are necessary to characterize the performance of NGF techniques quantitatively.

Appendix A.

The term $g(\lambda)$, as defined by Eq. (15), is a multiplicative function because it holds $g(\lambda\lambda') = g(\lambda)g(\lambda')$. Let us define the modulus m and the phase ϕ of the complex number λ. Then one can define

$$g[m \exp(i\phi)] = M(m)P(\phi).$$

One should have $P(x + y) = P(x)P(y)$, where x and y are real numbers, $P(0) = 1$ [as $P(0) = P(0)P(0)$]. If one defines $Q(x) = \ln[P(x)]$, this property can be written as $Q(x + y) = Q(x) + Q(y)$, from which it follows that $Q(x) = iax$, where a is *a priori* a complex number and then $P(x) = \exp(iax)$. However, one needs to have $P(x + 2\pi) = P(x)$, and then a must be an integer.

Let us now analyze the function $M(x)$. $M(x)$ satisfies the following property: $M(xy) = M(x)M(y)$. Let us introduce $N(x) = \ln[M(x)]$ and $H[\ln(x)] = N(x)$. It is possible to introduce $\ln(x)$ because the modulus x is positive. The property to satisfy is now $H[\ln(x) + \ln(y)] = H[\ln(x)] + H[\ln(y)]$ and thus $H[\ln(x)] = b \ln(x)$, where b is *a priori* a complex number (let us introduce $b = c + id$). Thus one has $N[x] = b \ln(x)$ and then

$$M(x) = \exp[c \ln(x)]\exp[id \ln(x)] = x^c \exp[id \ln(x)].$$

Let us now come back to the physical problem. The phase of $g[m \exp(i\phi)]$ should exist when m goes to 0 and thus the limit of the phase of $M(x)$ should exist when x goes to 0. This is the case only if $d = 0$. One thus finds that the general solution is

$$g[m \exp(i\phi)] = m^c \exp(in\phi),$$

where c is a real number and n is an integer.

The authors are grateful to J.-P. Huignard for his support in this work, and acknowledge fruitful discussions with S. Formont, B. Granier, and S. Tonda.

References

1. B. Javidi, "Nonlinear joint power spectrum based optical correlation," Appl. Opt. **28,** 2358–2367 (1989).
2. B. V. K. Vijaya Kumar and L. Hassebrook, "Performance measures for correlation filters," Appl. Opt. **29,** 2997–3006 (1990).
3. Ph. Réfrégier, "Filter design for optical pattern recognition: multi-criteria optimization approach," Opt. Lett. **15,** 854–856 (1990).
4. L. P. Yaroslavsky, "Is the phase-only filter and its modifications optimal in terms of the discrimination capability in pattern recognition?" Appl. Opt. **31,** 1677–1679 (1992).
5. Ph. Réfrégier, B. Javidi, and V. Laude, "Nonlinear joint Fourier transform correlation: an optimal solution for adaptive image discrimination and input noise robustness," Opt. Lett. **19,** 405–407 (1994).
6. F. M. Dickey and L. A. Romero, "Normalized correlation for pattern recognition," Opt. Lett. **16,** 1186–1188 (1991).

7. Ph. Réfrégier, "Application of the stabilizing functional approach to pattern recognition filters," J. Opt. Soc. Am. A **11,** 1243–1251 (1994).
8. N. Boccara, *Functional Analysis* (Academic, Boston, 1990), pp. 72–73.
9. Ph. Réfrégier, "Optimal trade-off filters for noise robustness, sharpness of the correlation peak and Horner efficiency," Opt. Lett. **16,** 829–831 (1991).
10. V. Laude and Ph. Réfrégier, "Multicriteria characterization of optimal Fourier spatial light modulator filters," Appl. Opt. **33,** 4465–4471 (1994).
11. L. P. Yaroslavsky, "The theory of optimal methods for localization of objects in images," in *Progress in Optics* E. Wolf, ed. (Elsevier, Amsterdam, 1993), Vol. 32, pp. 145–201.
12. B. Javidi and J. Wang, "Limitation of the classic definition of the correlation signal-to-noise ratio in optical pattern recognition with disjoint signal and scene noise," Appl. Opt. **31,** 6826–6829 (1992).