Optical pattern recognition based on pure phase correlation

KATARZyna CHałASIŃSKA-MACUKOW

Division of Information Optics, Institute of Geophysics, Faculty of Physics, Warsaw University, ul. Pasteura 7, 02-093 Warszawa, Poland.

JUAN CAMPOS, MARIA J. YZUEL

Group of Optics, Physics Department, Autonomous University of Barcelona, 08-193 Bellaterra, Barcelona, Spain.

In this paper, we present a review of recent results concerning the properties, performance, and optoelectronic realization of pure phase correlation, taking into consideration their application to pattern recognition.

1. Introduction

Pattern recognition has recently become a well established discipline of applied science with growing application in the robot vision, military area, industrial quality inspection, security control and identification, letter recognition, etc. The availability of low-cost semiconductor lasers, high resolution spatial light modulators operating at over 1 kHz, and high resolution fast CCD cameras facilitates recent development of specialized computer controlled optoelectronic processors which perform various image processing techniques at frame-rates not attainable by pure digital signal processing electronic devices.

Correlation function is a crucial operation in many methods of optical pattern recognition and usually plays the role of the measure of similarity between recognized images. Optics offers a simple and fast method to calculate the correlation between two images in real-time and this approach is expected to replace in the near future the traditional techniques of matching in applications which require the greatest possible speed of searching through large databases [1].

Classical correlation matched filter (CMF) has several drawbacks: correlation signal is broad and difficult to localize, discrimination ability is poor, correlation signal is sensitive to image distortion, orientation, etc. Since the introduction of the CMF much effort has been put into improving the performance of the optical correlation method for pattern recognition [2]. Hybrid approach which combines the flexibility of digital processing with the parallelism and speed of optical methods has opened up the possibility of creating several sophisticated algorithms and
processors adapted to the particular recognition problem. Various modified filters have been developed [3] and the new family of composite filters [4], nonlinear correlators [5]—[7] and trade-off filters [8], [9] has been proposed.

Pure phase correlation (PPC), [10], [11], also called phase-extracting correlation [12] belongs to the methods developed from nonlinear matched filtering [5]. The very name PPC means that we use only phase information of the Fourier transform in the recognition procedure. The algorithm of PPC performs a whitening of the scene spectrum and subsequently the correlation with the impulse response of the phase-only filter. It is well known that phase information of the Fourier transform is more important in image reconstruction than knowledge about amplitude distribution [13]. Following the great success of phase-only filtering resulting from high performance in light efficiency and discrimination capability the introduction of pure phase correlation was the next natural step in modification correlation algorithms.

Studies investigating phase correlation method show clearly that the sharpness of correlation peak, the light efficiency and the discrimination capability are superior in comparison with the results obtained with other methods [7]. During the last few years several papers have been published presenting various algorithms realizing phase correlation: symmetrical nonlinear filtering [5], binary joint transform correlator [6], phase extraction correlator [14], or the joint transform correlator with fringe-adjusted-filter [15]. But it is necessary to emphasize that phase correlation method works well in the noise-free case. If the signal-to-noise ratio of the input scene is low, the results given by phase correlation are less satisfactory [7]. Also in the case of multiobject input scenes some false alarms due to intermodulation effect are observed. Fortunately, in practice there are several methods proposed recently which allow us to avoid this problem [14]—[16].

In this paper, we present a review of recent results concerning the properties, performance, and optoelectronic realization of pure phase correlation, taking into consideration their application to pattern recognition.

2. Pure phase correlation — definition and properties

The concept of PPC can be explained by a comparison with classical correlation method. Let \( f(x) \) and \( g(x) \) be an input scene and the reference target, respectively. \( F(u) \) and \( G(u) \) are corresponding Fourier transforms. Classical correlation method is defined as

\[
f(x) * g(x) = \text{IFT} \{ \text{FT} \{ f(x) \} \cdot \text{MSF} \{ g(x) \} \}^{
\star}
\]

where: \( * \) denotes a correlation operation, FT and IFT denote Fourier and inverse Fourier transforms, respectively, and MSF[\( g(x) \)] is the matched spatial filter of the object \( g(x) \) to be detected. The Fourier transform of the real and positive function is a complex function described by their amplitude and phase distribution. We define phase extraction operation as a nonlinear procedure \( N[.] \) described by
Optical pattern recognition ...

\[
N\{F(u)\} = \begin{cases} 
\exp[i\varphi(u)] & \text{if } |F(u)| \neq 0, \\
0 & \text{if } |F(u)| = 0 
\end{cases}
\]

where \(|F(u)|\) and \(\exp[i\varphi(u)]\) are amplitude and phase distribution of the Fourier transform, respectively, and \(u\) are spatial frequencies in the Fourier plane. Pure phase correlation is equivalent to the symmetrical nonlinear matched filtering [5], see Fig. 1. Both the filter and the input scene Fourier transform pass through the same nonlinearity \(N\). In the output plane of correlator we obtain PPC function of the form

\[
\text{PPC}[f(x), g(x)] = \text{ITF}\{N[\text{FT}(f(x))]N[\text{FT}(g(x))^*]\}.
\]

![Fig. 1. Nonlinear optical correlator equivalent to the PPC method: FT — Fourier transform, N — nonlinear operator, CC — complex conjugate operator.](image)

The PPC operation may also be seen as two successive filters. The first filter is the whitening filter \(W\{\text{FT}[f(x)]\}\) of the scene spectrum which depends on the input scene spectrum. The second filter is the phase-only filter \(\text{POF}[g(x)]\) of the object to be detected. In such an approach the PPC distribution takes the form

\[
\text{PPC}[f(x), g(x)] = \text{IFT}\{\text{FT}(f(x)) W[\text{FT}(f(x))] \text{POF}[g(x)]\}
\]

where \(W[\text{FT}[f(x)]] = |\text{FT}[f(x)]|^{-1}\) is the whitening factor of the scene spectrum depending on the spectrum of the scene, and \(\text{POF}[g(x)] = \text{FT}[g(x)]^* |\text{FT}[g(x)]|^{-1}\) is the transmittance of the phase-only filter.

The PPC has properties typical of high frequency filtering and is very efficient as a method of sharpening the correlation signal [17]. It also optimizes peak-to-correlation energy criterion [3]. In digital image processing the use of phase only components of both the test and the reference data has been firstly suggested by Pearson et al. [18]. They created video-rate correlation processor giving delta-like correlation output signal. A very sharp correlation peak improves the probability of correct target localization and detection.
Note that for segmented input scenes, with the input scene equal to the target, the autocorrelation function obtained by PPC is equivalent to the output signal given by the inverse filtering (IF) [3], or MACE filter [19], which can be considered as a multitarget version in inverse filtering. The three methods mentioned above optimize peak-to-correlation energy criterion and assure very sharp, delta-like correlation. All of them can be seen as two filters operating in cascade. Nevertheless, they differ in the case of a general scene. The PPC and IF perform a whitening of the scene spectrum and subsequently the correlation with phase-only filter is obtained. The difference between the PPC and the inverse filter lies in the definition of the whitening part. In the PPC the whitening is proportional to the spectrum of the total input scene, while in the inverse filter it takes into account only the spectrum of the target. Thus we obtain

$$\text{IF}[f(x), g(x)] = \text{IFT}\{\text{FT}[f(x)] W[\text{FT}(g(x))] \text{POF}[g(x)]\}$$

where $W[\text{FT}(g(x))] = |\text{FT}(g(x))|^{-1}$ is the whitening factor which depends on the target spectrum. Similar situation is in the case of MACE filter [19], which becomes the IF in the case of only one training image.

2.1. Intermodulation effect

The nonlinearity used in the pure phase correlation method improves pattern recognition, but at the same time causes higher-order harmonics in the correlation plane in the case of multiobject scenes and creates intermodulation effect [14], [17], [20], [21]. Higher-order harmonics and the presence of aliasing (in the case of digital pattern recognition or in the case of optoelectronic processor with pixelated devices) may cause false alarms and misses. In most recognition problems the level of false alarms resulting from intermodulation is very low. False alarms are strongly enhanced when the multiobject input scene is composed of identical separated objects positioned at some regular intervals [14], [17], [21]. In order that the problem of intermodulation be better analyzed in the case of PPC we quote here the theoretical description of this problem presented by KOTZER et al. [14] for phase extracting correlation.

Let $f(x)$ be a multiobject input scene composed of different separated objects $f_i(x-a_i)$ positioned in different places $a_i$ in the input plane. The input scene takes the form

$$f(x) = \sum_{i=1}^{s} f_i(x-a_i)$$

and the corresponding Fourier transform distribution is given by the formula

$$F(u) = |F(u)| \exp[i\psi(u)] = \sum_{i=1}^{s} |F_i(u)| \exp\{i[\varphi_i(u) - ua_i]\}.$$
Optical pattern recognition ...

\[ F_\psi(u) = N \{ F(u) \} = \sum_{i=1}^{S} \frac{F_i(u)}{|F(u)|} \exp \{ i[\varphi_i(u) - a_i u] \} \quad \text{if} \quad |F(u)| \neq 0, \]

or \( = 0 \quad \text{if} \quad |F(u)| = 0. \) \hspace{1cm} (8)

If the input scene is composed of \( S \) identical objects, then the phase distribution in the input channel takes the form

\[ F_\psi(u) = \exp \{ i[\varphi(u)] \} = \sum_{i=1}^{S} \exp \{ -i a_i u \} \frac{1}{\sum_{i=1}^{S} \exp \{ -i a_i u \}}. \] \hspace{1cm} (9)

If the reference target is the same as input object, the output correlation function corresponding to the PPC has the distribution \([14]\)

\[ \text{PPC}[f(x), f(x)] = \sum_{i=1}^{S} \delta(x - a_i) \oplus V(x) \] \hspace{1cm} (10)

where

\[ V(x) = \text{IFT} \left( \frac{1}{[S + D(u)]^{1/2}} \right) \] \hspace{1cm} (11)

and

\[ D(u) = \sum_{i=1}^{S} \sum_{k=1}^{S} \exp[iu(x_i - x_k)] \] \hspace{1cm} (12)

are functions of the input scene configuration independent of the shape or size of the objects. Here \( \oplus \) denotes the convolution operator.

The modulation term \( V(x) \) depends only on the input object distribution. If the position of the input objects is random then, following the analysis presented by KOTZER et al. \([14]\), the \( D(u) \) is derived from many uncorrelated terms and the function \( V(u) \) describes a random process. The function \( V(x) \) is expected to be similar to the delta function (see Appendix A in \([14]\)). If the multiobject input scene is composed of many various objects they will not correlate independently of their arrangement. In this case, the function \( V(x) \) also retains delta like shape. In the case of regular arrangement of the input scene containing identical images the function \( V(x) \) takes the form \([14], [20]\)

\[ V(x) = \text{IFT} \left\{ \frac{\sin(\pi ub)}{\sin(\pi uSb)} \right\} \] \hspace{1cm} (13)

where \( b \) is the spatial distance between images. Therefore \( V(x) \) is an array of delta functions with modulated intensity. The correlation is no longer shift-invariant. It gives various correlation peak values depending on the position in the input plane and creates false alarms. Each object affects the correlation response of all the other objects in an essentially unpredictable way.
Fig. 2. Numerical results of intermodulation effect: a — input scene (distance between letters is 64 pixels), b — correlation plane in the case of the 128 x 128 input matrix, c — correlation plane if three letters are embedded in a 256 x 256 matrix.

If the phase distribution in the Fourier plane is sampled and displayed by means of a spatial light modulator, the correlation is periodically repeated in different orders. The peaks corresponding to harmonics of one order are mixed up with the peaks corresponding to other orders and this effect of aliasing may enhance intermodulation effect [20], [22]. TURON et al. [20] presented the digital analysis of how the effect of aliasing together with intermodulation can have strong influence on the recognition results. Let the input scene consists of three identical objects arranged in a regular array (Fig. 2a). The value of false alarm depends on the dimension of the input matrix and orientation of the direction of repetition in comparison with the sampling direction [20]. If the sampling direction coincides with the direction of the object repetition and the matrix dimension is a multiple of the distance between objects, the false alarm is the highest. The fourth peak (Fig. 2b), which is a false alarm, is due to the additive superposition of the harmonics. When the scene is embedded in a bigger matrix, the false alarms decrease significantly (Fig. 2c).

The study of the intermodulation effect leads to the conclusion that the PPC performs well for any nonpatterned inputs, in which the linear phases generated by the displacement of the object do not modulate the amplitude distribution of the Fourier transform of the input scene. Each perturbation of regularity, input noises, presence of non-target objects in the input scene, etc., reduce intermodulation effect. It is for this reason that in the experiment reported by KOTZER et al. [23] correlation signals achieved approximately equal peak values, instead of being unequal in height according to the theory and simulation.
3. PPC with improved discrimination capability

The discrimination capability (DC) of the recognition processor is one of the most important parameters and several methods have recently been proposed to avoid false alarms. In the case of binary or nonlinear joint transform correlator some improvement has been achieved by introducing different types of thresholding methods in the Fourier plane [15], [16], [22], [24]. Similarly, KOTZER et al. [14] proposed a method that uses adaptive thresholds to remove false alarms in phase extracting correlator.

The PPC with improved discrimination capability by blocking technique (BPPC) [25] has recently been proposed by AHOUZI et al. [26]. The region of support for the PPC which improves the discrimination capability is designed following this approach. In nonlinear correlators the output signal depends on the input scene and, in consequence, on the spatial distribution of the objects. Therefore, to optimize the DC in the PPC case by blocking technique some a priori information about the input scene should be taken into account during the definition of the region of support.

In the proposed blocking technique no complete knowledge about the input scene is required. There are three approaches [26]: intermodulation approach, segmented scene approach, and POF approach. Intermodulation approach takes into consideration knowledge about the modulus of the Fourier transform of the input scene, where the spatial distributions of the input objects are recorded. We define intermodulation function depending on the scene arrangement [26]. This function coincides with the modulation function introduced by KOTZER et al. [14]. The modulus of the weighting function introduced by JAVIDI et al. [27] coincides with the inverse of the intermodulation function. The two other methods take into account the objects which may appear in the input scene but not their spatial distribution. Thus, in the case of the segmented scene approach the design of the region of support is based on the knowledge about segmented input scene and the resulting BPPC is applied to the multiobject input scene. The third approach is suggested by the relationship between the PPC and POF. Firstly, the discrimination capability of a POF is optimized by means of the blocking technique [25]. The BPPC function is obtained by the convolution of the POF correlation function and the intermodulation function in spatial domain.

Figure 3 illustrates qualitatively the improvements achieved with three methods of blocking technique. Figure 3a shows the input scene with four different characters. Figure 3b presents the correlation plane obtained with PPC method without blocking technique. Figure 3c presents correlation plane distribution when all a priori information about the input scene has been used to design region of support. Figure 3d shows the result of the intermodulation approach. The results of correlation for the segmented scene approach and POF approach are shown in Figs. 3e and f, respectively.

A comparison of Figs. 3b—f with Fig. 3b shows that in the case of PPC the discrimination capability increases significantly for all the approaches when the
Fig. 3. Input scene to be recognized (a). Recognition by using the PPC without blocking technique (b). Correlation results obtained with various blocking techniques: all a priori information is used (c), intermodulation approach (e), segmented scene approach (d), POF approach (f).

blocking technique is applied. Note that if we want to obtain higher discrimination capability the number of blocked elements is higher and the reduction of light efficiency of the setup is observed [26].

4. Optoelectronic realization of PPC

With the development of addressable SLM's and the advances of the CCD technology, various optoelectronic systems performing nonlinear processing were proposed. ROZEN et al. [12], [14], [23] proposed an optoelectronic phase extraction correlator. Several optoelectronic realizations of PPC are based on the joint transform correlator architecture. WANG and JAVIDI [27] showed that binary joint transform correlator with proper threshold function leads also to the PPC in the first order Fourier terms. CHALASINSKA-MACUKOW and GORECKI [28] presented an optoelectronic quasiphase only correlation architecture based on classical joint transform correlator architecture. ALAM and KARIM [29] introduced fringe adjusted
joint transform correlation to obtain typical output signal for the PPC. The PPC function can also be realized in optoelectronic dual nonlinear correlator based on computer controlled joint transform processor with variable discrimination capability [30]. COHN [31] proposed another approach — an adaptive real-time architecture for PPC, where the phase extraction of Fourier transform of the scene and the target were obtained by a Mach–Zehnder interferometer.

A convergent correlator (L1, L2) is introduced in the arm of a Mach–Zehnder interferometer (BS1, BS2, PZT, BS3). The lens L3 images the plane \((X_2, Y_2)\) in the CCD1 camera where the interferences are digitized. The computed phase is displayed in the SLM to obtain the phase correlation in the CCD2 camera.

**Fig. 4.** Phase-shifting interferometric correlator: M — mirror, BS — beam splitter, L — lenses, PH — pinhole, SLM — spatial light modulator, CCD — CCD camera, PZT — piezo-electric transducer.

AHOUZI *et al.* [32] proposed new experimental setup shown in Fig. 4 to perform the PPC. It is an optoelectronic phase processor based on a conventional convergent correlator. Using a Mach–Zehnder interferometer the phase correlation is extracted in the Fourier domain. After digital analysis the setup has been performed experimentally [33]. The phase-shifting interferometry (PSI) techniques have been used to improve the accuracy of the phase extraction processor by using a four step algorithm. To improve the discrimination capability the blocking technique has been introduced [26]. Figure 5 shows an example of optical results obtained by using
phase shifting optoelectronic correlator. The input scene contains four characters (Fig. 5a). Figure 5b shows the distribution of the binary mask which has been introduced into Fourier plane to improve the discrimination capability by blocking technique. Figures 5c—e show correlation plane obtained experimentally in the case of phase only filtering (for comparison), the PPC method, and the PPC method with blocking technique. The best result of discrimination is obtained by using the PPC method with blocking technique.

5. Intensity and contrast invariance of PPC

The whitening of an input scene and reference target Fourier spectra results in intensity invariance in such a sense that a constant amplitude scaling factor of the input signal does not modify the detection signal [34]. The effect of varying the
illumination of the input scene has been investigated recently. It has been shown in computer simulation and experimental results that the intensity invariance always accompanies the pure phase correlation irrespective of the approach and realization [7], [35], [36].

In the case of segmented input scenes the PPC demonstrates contrast-invariant property [34]. Contrast invariant properties are especially important if only a part of the object information is used for the recognition procedure to obtain, for example, the distortion-invariant method [37], to recognize low contrast objects in the presence of high clutter, or in colour pattern recognition [38]. In the case of multiobject input scenes, the contrast invariance is lost, but constant amplitude scaling factor still does not modify the correlation signal.

6. Application of the PPC to colour pattern recognition

The improvement of the recognition process when colour information is added consists in the fact that for a colour object both the shape and intensity distribution may depend strongly on the illumination wavelength. The multichannel recognition of polychromatic objects [39] is based on the fact that either the shape or the intensity distribution depend on the channel used. Each channel corresponds to an illuminating wavelength. In the visible region three channels (red, green, blue) usually suffice. In this case, three correlations are obtained. The filters are designed from the targets corresponding to each channel. By combining the information provided from the three channels false alarms may be eliminated. At the end of the multichannel correlator we use logical operation AND if we are interested in the colour of the object and the shape. If we are interested only in the shape, we use logical operation OR.

One of the problems in colour pattern recognition is to recognize properly a character over a background regardless of the character-background colour combination. For each channel this corresponds to a different transmission for the character and the background. So, in each channel there is a problem related to contrast. Depending on the colour combination between object and background we can have direct contrast: a bright letter over a dark background, or inverse contrast: a dark letter over a bright background. According to the strategy proposed by MILLAN et al. [39] two POF's filters are needed in each channel. One of them is matched to a direct contrast object, and the other one is matched to an inverse contrast object.

As was mentioned above, the PPC method in the case of mono-object input scene is contrast-invariant. But we have to emphasize that in the case of multiobject input scene the PPC is still invariant in such a sense that the sign of contrast, direct or inverse, does not affect the results of correlation [39]. This is an interesting finding for application in colour pattern recognition, because in the PPC method we need only one filter in each channel.

Tables 1 and 2 [39] show the digital result of recognition (correlation peak values) for the six colour combination by using phase only filtering or the PPC
Table 1. Correlation peaks for the six different colour combinations. Phase only filtering.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Contrast</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>R dir.</td>
<td>0.64</td>
<td>1</td>
<td>0.29</td>
<td>0.49</td>
<td>0.20</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>R inv.</td>
<td>1</td>
<td>0.48</td>
<td>0.53</td>
<td>0.24</td>
<td>0.26</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>G dir.</td>
<td>0.23</td>
<td>0.21</td>
<td>0.61</td>
<td>0.37</td>
<td>0.61</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>G inv.</td>
<td>0.24</td>
<td>0.12</td>
<td>0.20</td>
<td>0.69</td>
<td>1</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2. Correlation peaks for the six different colour combinations. Pure phase correlation.

<table>
<thead>
<tr>
<th>Channel</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>0.83</td>
<td>1</td>
<td>0.28</td>
<td>0.42</td>
<td>0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>G</td>
<td>0.06</td>
<td>0.07</td>
<td>0.76</td>
<td>0.82</td>
<td>1</td>
<td>0.95</td>
</tr>
<tr>
<td>OR</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

method, respectively. To simplify the problem we have only two channels, red and green, and we play with only these two colours. The shape to be recognized is capital letter X and the POF filter is matched to the shape of X.

From Table 1 corresponding to phase only filtering we can see that it is necessary to use two filters, matched to direct and inverse objects, to be sure that all the objects will be recognized properly. In the case of the PPC method one filter is enough and all objects are perfectly recognized. Moreover, the peaks obtained with the PPC method are more detectable than those obtained with POF's.

7. Conclusions

We have reviewed the crucial results of studies on the pure phase correlation method. We have discussed the problem irrespective of optical realization of the PPC, which can be based on different architectures: 4f correlator, classical joint transform correlator and their nonlinear and binary versions. Also interferometric techniques can be engaged in the phase extraction in the Fourier plane.

It has been shown that the PPC gives good correlation performance. The sharpness of the correlation peak, the light efficiency and the discrimination capability are superior with the PPC method in comparison with the results obtained with other methods. When multiobject input scenes include multiple identical targets, some false alarms may appear due to intermodulation effects. Fortunately, using a spatially varying threshold, blocking technique, modified fringe adjusted filter, etc., good performance can be achieved even in this case.

In the case of low signal to noise ratio of the input scene the PPC method does not give equally good results as in the noise-free case, because of its high frequency
characteristics. In the case of high noisy input scenes, it is necessary to find the compromise between discrimination capability and the sharpness of correlation peak or the noise robustness.

We presented an example of application of the PPC method to colour pattern recognition. The list of references to this paper contains more examples of how a very narrow correlation peak and enhanced discrimination capability can improve recognition, location, and classification of objects.

Acknowledgements — This work has financed by the Dirección General de Ensenanza Superior de Ministerio de Educacion y Cultura (Spain) under the project PB96-1134-C02-01 and State Committee for Scientific Research (KBN), Poland, grant No. 8 T11F 028.

References


Received May 22, 2000