Optical nonlinear correlation based on nonuniform subband decomposition

P Garcia-Martinez†, C Ferreira† and D Mendlovic‡

† Departamento d’Optica, Universitat de València, C/Dr Moliner, 50, 46100 Burjassot (València), Spain
‡ Department of Physical Electronics, Faculty of Engineering, Tel Aviv University, 69978 Tel Aviv, Israel

Received 19 July 1999, in final form 6 September 1999

Abstract. We present a nonlinear correlation to improve the selectivity for optical pattern recognition. The approach is based on morphological correlation which involves a threshold decomposition concept. Hereby, we propose a subband decomposition in the Fourier domain to perform the threshold decomposition operation. We consider two frequency bands that give rise to two separate channels. We apply the morphological correlation to each channel using a localized threshold decomposition. Then, we define a two-channel morphological correlation. The final detection decision is made as a combination of both correlation outputs. The two-channel morphological correlation yields improved discrimination capability and noise robustness compared with common linear and morphological correlation. Computer simulations are presented as well as optical experiments performed using a joint transform correlator.

Keywords: Morphological correlation, optical pattern recognition, threshold decomposition, subband decomposition, joint transform correlator

1. Introduction

Correlation is one of the most popular approaches used for signal and image processing [1]. Owing to the achievements of electronic computers, digital techniques play an important role. However, electronics presents physical limits which result from the interaction between electrons and the necessity of permanent conducting lines. Optics offers an alternative solution for interconnections. Moreover, the realization of light-speed operations and the inherent parallelism of optical processors are additional advantages of optics. Correlation is a well known example of an operation that can be ideally performed using simple optical systems. Moreover, this operation is widely used for optical pattern recognition applications. Due to advances in these applications, a high selectivity of the recognition methods is required. Matched filtering is already widely used for template matching and pattern recognition [2]. Moreover, the common matched filter is optimum in the mean square error (MSE) criterion, meaning that minimizing the MSE leads to maximizing the common linear correlation (LC) [3]. The MSE is a function of the norm and the acceptance of the LC is due to the mathematical tractability of the square error metric and the easy optical implementation. Despite these advantages, matching metric norms have been widely used due to their relevance in practical recognition tasks. One of these metrics is the mean absolute error (MAE).

It has been shown that minimizing the MAE criterion is equivalent to maximizing a nonlinear correlation, namely the morphological correlation (MC) [4]. This MC presents high discrimination capability for pattern recognition as compared with the common LC. Also, due to its nonlinearity property, MC offers a more robust detection of low-intensity images in the presence of high-intensity patterns which are to be rejected. Moreover, MC has been implemented optically by means of a joint transform correlator (JTC) [5].

The MC is expressed in terms of a threshold decomposition [6]. It is viewed as the sum over all amplitudes of the LC between thresholded versions of an input scene and a reference pattern to be found in the input scene, at every grey level [4]. The common threshold decomposition is obtained using binary slices that are uniformly distributed according to the grey-level information of the images.

In this paper we propose a nonuniform threshold decomposition based on the widely known subband decomposition (SD) [6] to produce bands of localized frequency information. Those bands will be differently threshold decomposed. Because each threshold decomposition implies a MC, we will consider a multichannel MC. The final combination of the MC outputs provides a high improvement in the discrimination capability when compared with conventional MC. Optical and simulated results are compared for the two-channel MC, the common MC and the conventional LC.
In section 2, we review the definition of the MC based on the threshold decomposition concept. Motivated by the idea of the SD, section 3 presents the local Fourier domain slicing. The novel nonlinear correlation technique for improving the discrimination capability of the MC is described in section 4. Simulation and experimental results are provided in sections 5 and 6, respectively. Finally, in section 7 the conclusions are drawn.

2. Nonlinear MC

Maragos introduced the MC as a robust detection process to minimize the MAE [4]. The MC provides better performances and higher discrimination capability for pattern recognition tasks in comparison with the LC.

The MAE between two functions is

$$\text{MAE}_{gf}(x, y) = \sum_{\varsigma=0}^{Q} |g(\varsigma + x, \eta + y) - f(\varsigma, \eta)|$$

where \(f\) is a pattern to be detected in \(g\). For the sake of clarity, we consider that both \(f\) and \(g\) are defined in a discrete domain. The nonlinear MC \((\mu)\) that minimizes the MAE error, is defined as

$$\mu_{gf}(x, y) = \sum_{\varsigma=0}^{Q} \min[g(\varsigma + x, \eta + y), f(\varsigma, \eta)]$$

which corresponds to the definition of the MC given by Maragos. MC is also performed using the threshold decomposition concept [6]. It has been shown that the MC between the thresholded versions of \(g\) and \(f\) at every grey-level value \(q\) [4], is equal to the LC \((\gamma)\) between them, i.e.,

$$\mu_{gf}(x, y) = \sum_{q=1}^{Q} \mu_{S_{\gamma}f_{\gamma}}(x, y) = \sum_{q=1}^{Q} \gamma_{S_{\gamma}f_{\gamma}}(x, y)$$

where \(Q\) is the number of grey levels in the images. This is not surprising in view of the fact that, for binary signals, the MC and the LC coincide because the minimum of two binary values is equal to their product.

The threshold decomposition concept has been widely used for nonlinear morphological and rank-order filtering [6, 7]. For equation (3), we consider the threshold decomposition as a uniform summation of \(Q\) grey-level binary slices. Nevertheless, this decomposition can be defined in a different way. We hereby decompose the image based on its spectral information, as we explore in the following section.

3. Threshold decomposition based on the SD concept

The basic idea of SD is to separate the Fourier spectrum of an image into nonoverlapping bands and transform each band separately to obtain the final pass-band image [8]. Image compression and object or feature recognition require the study of the codification of a signal in subbands [9, 10]. Each image is subsampled according to its frequency information and then codified separately using methods based on the statistics of the image. Then, the basic definition of SD is to use filter banks to produce localized spectral signals.

One procedure to decompose an image is to use high-pass and low-pass filters. This idea has been widely applied in optical pattern recognition. There are many filters that give more relevance to the high- or low-frequency information of an image. For instance, phase-only filters [11] enhance the high-frequency information. Because the details of the image and the borders are typical information of high frequencies, this filter will have high discrimination capability for pattern recognition. Nevertheless, this filter is very sensitive to distortions.

According to the threshold decomposition [6], the function \(f_{q}(x, y)\) of an input image \(f(x, y)\) is obtained through decomposition by a threshold value \(th(q)\) as

$$f_{q}(x, y) = \begin{cases} 1 & f(x, y) \geq th(q) \\ 0 & f(x, y) < th(q). \end{cases}$$

If the threshold values are equally distributed, a constant distance, \(S\), separates any consecutive levels. That is if \(\delta q = qS\) then

$$\hat{f}(x, y) = \sum_{q} S_{f_{th(q)}}(x, y) \cong f(x, y).$$

In the last equation, the error involved representing the quantization error approaches zero as the number of contributing slices approaches infinity, whereby the threshold step size, \(S\), tends to zero. If the interval value \(S\) coincides or is smaller than the separation between the original quantized-image grey levels, the decomposition is complete, then the image is exactly reconstructed \((\hat{f} = f)\). In contrast, the edges or the fine details of the image usually appear on the high-frequency components. So, a low-pass filtering associates the object with ‘classes’ (‘inter-classes’) whereas high-pass filtering provides high-discrimination ‘intra-classes’.

Motivated by these spectral filtering properties, one can make the supposition that it is possible to give more relevance to the high or low spectral components when one is applying the threshold decomposition, in contrast to the MC.

We decompose the function \(f(x, y)\) as

$$f(x, y) = f^H(x, y) + f^L(x, y)$$

where \(f^H(x, y)\) represents a real function with the high frequencies of \(f(x, y)\) and \(f^L(x, y)\) is a real function with the low-frequency information. Taking into account equation (5), and the separation of the image into two bands, both channels will carry the threshold decomposition in the respective regions. The lower-frequency channel uses a threshold decomposition with a given value \(S_{L}\), and the high-frequency channel uses a different set of grey levels given by \(S_{H}\). Taking into account that our motivation is to improve...
the selectivity for the recognition process, we consider an $S_{\mu}$ value greater than $S_{\mu}$. In this way, we will use more slices for the threshold decomposition in the high-frequency domain channel. This choice will emphasize the details of an image and therefore the discrimination capability. Because the low-frequency information is not as important as the high-frequency information for the selectivity, we used fewer slices for the low-frequency channel. Note that we only used just two widely used channels—the low- and the high-frequency channels—because there are many differences between them. We are also studying the possibility of using more channels and of optimizing the method in the mathematical sense.

This idea is connected with image compression because one of the most common techniques in that subject is related with the subband coding, which consists of creating various spatial frequency bands of the original full-band signal [12]. In these methods one gives more significance, i.e. more bits of information, for different frequency bands.

We define $N = \frac{Q}{2}$ as the number of slices involved in the threshold decomposition.

The consideration of different $S$ values for different bands of the image implies different $N$ values. Normally $Q = 256$ for a 256 grey-level image, but we consider images that have 16 grey levels. To perform the conventional MC we have used $S = 16$, so 16 slices are involved in the process $(N = 16)$. For the high-frequency channel $N_{fL} = \frac{Q}{2}$, and so we are considering eight slices for the MC for the high-frequency channel, and for the low-frequency channel we use $N_{fL} = \frac{Q}{4}$, i.e. four slices. In this way, we reduce the number of operations involved in the case of the two-channel MC to 12 slices, in comparison with the 16 slices needed for the MC.

4. Two-channel MC

In this paper we propose a two-channel MC to improve selectivity. As we have shown in the previous section, we propose a two-band decomposition based on the spectral information of the images. Because different threshold decompositions are applied to both bands we define a two-channel MC.

The MC in each channel will be expressed as

$$\mu_{gf}^{H,L} = \sum_{q} S_{th(q)}^{H,L} \otimes f_{th(q)}^{H,L}$$  (7)

where $\otimes$ denotes LC, and both $f$ and $g$ are decomposed as

$$\hat{f}_{H,L}^{H,L}(x, y) = \sum_{q} S_{th(q)}^{H,L} f_{th(q)}^{H,L}(x, y).$$  (8)

Note that equations (7) and (8) contain the notation for both the low- and high-frequency channels.

Once the correlation outputs $\mu_{gf}^{H,L}$ and $\mu_{gf}^{L}$ are obtained, the final step is to combine both information. We apply the channel combination used for the multichannel pattern recognition [13].

Multichannel pattern recognition is based on applying the correlation process into the channels separately. It is widely used in colour pattern recognition, where the spectral information of each colour is treated as a different channel [13]. In our approach we consider two channels and to avoid false alarms we require positive recognition for both of them. This is eliminated by a combination of the two correlation outputs. There are several ways to combine the correlation outputs. One is based on recognizing the object when the correlation peak is obtained in each channel at the same position of the output plane. This criterion may be represented by means of the logical AND operator applied to the correlation peaks for all the channels. In this paper, we have used other procedures to combine the information of both channels based on a simple multiplication of the correlation channels pixel by pixel. If a pixel has high intensity in each channel then the final image will present a high intensity value in that pixel. This pointwise operation is performed at a high rate, reducing the computational time. Moreover, this method reduces both the false alarms and the sidelobes in the output planes.

The two-channel approach consists of performing the MC for each channel and then combining the results. We extract the low-frequency components of an input image and a reference object. To obtain the lower-frequency components we have applied an average filtering which smooth the images enhancing the lower frequencies. For the high-frequency channel, we simply subtract the information of the object ($f(x, y)$) and the low-frequency image ($f^L(x, y)$), then a detailed image is obtained ($f^H(x, y)$).

We chose the number of slices per channel more or less arbitrarily. However, because we do not have any a priori information about the object and because the choice of the slices is object dependent, we chose a fixed number of slices. Our intention is to use fewer slices for the new correlation in comparison with the normal MC. Because our objects have 16 grey levels, we use 16 slices for the MC and 12 for the two-channel MC. We consider multiples of two, which is why we used eight slices in the high channel and four for the low channel.

5. Computer simulation results

One might expect that because the two-channel MC is very selective for object detection and discrimination, it might have poor noise robustness. In this section we consider images that are degraded by correlated Gaussian non-overlapping noise and overlapping Gaussian noise. We measure the discrimination capability (DC) that is connected with the noise robustness. The DC that we consider is the percentage ratio between the highest cross correlation and the autocorrelation, the cross correlation being the highest peak in the correlation plane that does not correspond to the objects of interest. A low DC value means that the detection is very selective, and that a good discrimination capability is obtained. In contrast, higher DC values mean that the method is not selective, and that false signals are detected instead.

The input scene for the non-overlapping noise case is shown in figure 1. This image consists of different objects. Some of them are very similar to the reference object that is marked with an arrow. In figures 2(a) and (b) the LC and the MC tridimensional plot outputs are shown. We used 16 slices ($S = 16$, in equation (8)) for the MC. The LC yields a false detection whereas the MC needs a DC of 70% to detect the reference object.
Figure 1. The input scene with spatially disjoint correlated noise. The reference object is indicated by an arrow. There are four false targets in the figure.

Figure 2. (a) Tridimensional plot for the LC of figure 1. (b) Tridimensional plot for the MC of figure 1. (c) Final tridimensional two-channel MC and correlation peak profiles when the combination operation is a product for scene shown in figure 1.

Regarding the two-channel approach, we used fewer slices than the MC case. We used eight slices \( S_{L} = 32 \), in equation (8)) for the high-frequency channel and four slices \( S_{L} = 64 \), in equation (8)) for the lower-frequency domain. Figure 2(c) shows the two-channel MC. Note that we detect the image with a DC of 20%.

In a second experiment, we deal with objects that are corrupted with white Gaussian noise. We used zero mean and a standard deviation of \( \sigma = 1 \). Figure 3 shows the input scene, and figures 4(a) and (b) show the LC and MC tridimensional outputs. Note that neither LC nor MC achieve good results. For the LC a false detection is obtained, whereas the DC for the MC is almost 90%. This can be considered as a false detection.

In a second experiment, we deal with objects that are corrupted with white Gaussian noise. We used zero mean and a standard deviation of \( \sigma = 1 \). Figure 3 shows the input scene, and figures 4(a) and (b) show the LC and MC tridimensional outputs. Note that neither LC nor MC achieve good results. For the LC a false detection is obtained, whereas the DC for the MC is almost 90%. This can be considered as a false detection.

In contrast, the two-channel approach shown in figure 4(c) gives sharper peaks and a DC of 40% ensures detection. Therefore, as well as improving the high discrimination we decrease the computational complexity since a reduced number of slices are enough to obtain the correlation outputs.

We would like to point out that even with high-discriminant filters (for instance, phase-only filters, inverse filters or Wiener filters) we are not able to correctly detect the reference pattern. Then, using a nonlinear correlation is an adequate option for pattern recognition.

6. Optoelectronic implementation

The two-channel MC operation requires many calculations. Thus, optics may be attractive for efficient implementation. The approach used for the optoelectronic implementation of the two-channel MC is the same as used for the optoelectronic implementation of the MC [5]. This system is a JTC [14], which exhibits the cross correlation of two functions via a parallel optical computation. For each slice, the joint power spectrum is performed. The pair of slices (one from the reference object and one from the input scene) are placed alongside one another in the input plane. The summation of the joint power spectrum of these pairs is stored by computer and finally fed back to the input plane for a second Fourier transform to produce the MC.
Optical nonlinear correlation based on nonuniform subband decomposition

The optoelectronic system was built to test the performances of the two-channel MC and compare them with the classical LC and common MC performances. The joint input image that we introduce in the JTC setup consists of a reference object and another similar object for testing both autocorrelation and cross-correlation performances. In figure 5 we show the joint input scene for the case of noise-free objects with high similarity. In figure 6 the region of interest and a tridimensional plot of the output correlation plane is shown. It presents the auto- and cross-LC; in this case a threshold of 73% is needed to reject the nondesired object. Figure 7 also shows the optical MC output. In this case a detection threshold of 65% is required. We see that MC presents better selectivity in discrimination capability than the common LC. We used 16 time-sequential joint-input slices to finally obtain the summation of all the joint power spectra. Those spectral distributions are finally Fourier transformed to obtain the MC [5].

As we pointed out in section 3, to perform the two-channel MC we need a high-pass and low-pass image. Due to the optical implementation, one might obtain those images taking profit of the optical means. In figure 8, the final multiplication of both the MC for the high- and low-frequency channels is presented. A threshold of 46% is needed to isolate the reference object. Thus, better discrimination capability is obtained when we compare this novel optical nonlinear correlation with optical common MCs and LCs for objects with high similarity.
7. Conclusion

In conclusion, we have introduced an optical two-channel MC approach that improves the DC for pattern recognition when it is compared with the conventional MC.

The MC operation involves a threshold decomposition. We present a different slicing decomposition depending on the Fourier domain. We have considered the low- and high-frequency information as two separate channels. Because the high-frequency components provide more information about the details of the image, they are more sensitive to be considered for a recognition task. We give more relevance to that information when we apply MC. In contrast, because lower frequencies do not bear an important information in selectivity, we have used less slicing for this channel. The combination of both channels improves the conventional MC discrimination performances as is demonstrated by the simulation and experimental results. We tested recognition in scenes with many objects and different kinds of noise.

Much effort is focused on the optimization of the parameters such as the number of slices or how many channels are involved in the process. These parameters are under development by us, but that is beyond the scope of this paper.

Acknowledgments

This work has been supported by the Spanish DGES, Dirección General de Enseñanza Superior, project PB96-1134-C02-02. The authors thank Professor H H Arsenault for many useful remarks.

References