Input-image homogenization as a method to improve a correlator's discrimination capability

Leonid P. Yaroslavsky

Department of Interdisciplinary Studies, Tel Aviv University, Tel Aviv 69978, Israel

Octavio López-Coronado and Juan Campos

Departamento de Física, Universidad Autónoma de Barcelona, E-08193 Bellaterra, Spain

We propose an input-image preprocessing method consisting of homogenization of the image to improve the discrimination capability of a correlation-based recognition process. This method is an approximation of the optimal filter. It offers the advantage that correlation with the preprocessed images can easily be implemented in an optical correlator working with phase-only spatial light modulators. © 1998 Optical Society of America

OCIS codes: 070.4550, 070.5010, 100.2980.

Detection and localization of objects is a task of extreme importance in many applications. Various approaches can be used to achieve this goal. In all of them the input scene is compared with a reference to yield the degree of similitude. Correlation-based methods can be implemented optically, exploiting the parallelism inherent in optical systems. VanderLugt[1] proposed the use of a filter with an amplitude transmittance proportional to the complex conjugate of the Fourier transform of the object to be recognized. This filter was obtained by optical holography and was called a classical matched filter. It is well known in communication theory that the detection signal-to-noise ratio is maximized when the noise is signal independent, stationary, and additive.[2] Subsequently, many criteria have been defined for evaluation of the performance of the recognition task,[3] and many filters that optimize these criteria have been proposed,[4] most of which can be considered as two filters acting in cascade. The first filter preprocesses the input scene, and the second evaluates the correlation with the preprocessed target.[4] The phase-only filter[5] (POF) is popular because of its high light efficiency and easy implementation in currently available spatial light modulators (SLM's). It has been shown[6] that the POF can be seen as an approximation of the optimal filter[7] in terms of its discrimination capability.

In many real situations the object to be recognized is observed on a complex cluttered background. In these cases most filters (i.e., the classical matched filter and the POF) that were designed without these facts' being taken into account fail to detect the target.[7,8] For this reason, to design robust recognition methods one must take into account the image background.[7,9,10] Also in this case, preprocessing of the input scene before the correlation has been proposed.[11]

We propose a preprocessing method that can be seen as an approximation of the optimal correlation filter to locate objects in spatially inhomogeneous images.[9] The advantage of the method is that the final correlation filter can easily be implemented in a SLM that is working in a phase-mostly regime. Some optical experiments have been carried out to test the proposed method.

From the theory of optimal methods for target location in images[9] it follows that, for spatially inhomogeneous images, the optimal correlation filter should be designed and applied to the input image fragmentwise, with the image fragment size chosen such that the image fragments can be regarded as spatially homogeneous. Here we implement and investigate a correlator input-image preprocessing method that is a simplified version of the above general local adaptation procedure.

Let $\{\beta(x_0, y_0; f_x, f_y)\}$ be the spatial spectrum of an input-image fragment centered at coordinates (x_0, y_0), and let $\alpha^*(f_x, f_y)$ be the complex-conjugate spectrum of the target that is to be located [(f_y, f_y) are spatial frequencies along the (x, y) coordinates]. Then the optimal local adaptive filter for location of the target within the given fragment is defined as[9]

$$H(x_0, y_0; f_x, f_y) = \frac{\alpha^*(f_x, f_y)}{\text{AV}\{|\beta(x_0, y_0; f_x, f_y)|^2\}}, \qquad (1)$$

where AV is an averaging operator that should be applied to the observed image fragment's power spectrum for evaluation of the power spectrum of the fragment's background component (the portion of the fragment that does not contain the target). Fragmentwise processing assumes implementation of the optimal filter of Eq. (1) individually for each input image fragment.

One can simplify the procedure for the filter local adaptation by the following reasoning. Consider the following filter:

$$\tilde{H}(x_0, y_0; f_x, f_y) = \frac{\alpha^*(f_x, f_y)}{\text{Var}(x_0, y_0)}, \qquad (2)$$

where

$$\text{Var}(x_0, y_0) = \iint_{f_x, f_y} |\beta(x_0, y_0; f_x, f_y)|^2 \mathrm{d}f_x \, \mathrm{d}f_y \qquad (3)$$

is the local variance of the input image within the given fragment. This variance can be regarded as a zero-order approximation to the local power spectrum $AV\{|\beta(x_0, y_0; f_x, f_y)|^2\}$, and therefore the filter of Eq. (2) can be regarded as a zero-order approximation to the filter of Eq. (1). Furthermore, one can separate the filter of Eq. (2) into two filters in cascade:

$$\tilde{H}(x_0, y_0; f_x, f_y) = \widetilde{H_{\mathrm{mf}}}(x_0, y_0; f_x, f_y)\widetilde{H_{\mathrm{pp}}}(x_0, y_0; f_x, f_y)$$

$$= \frac{\alpha^*(f_x, f_y)}{[\mathrm{Var}(x_0, y_0)]^{1/2}} \frac{1}{[\mathrm{Var}(x_0, y_0)]^{1/2}} . \quad (4)$$

The first filter in the cascade,

$$\widetilde{H_{\mathrm{pp}}}(x_0, y_0; f_x, f_y) = \frac{1}{[\mathrm{Var}(x_0, y_0)]^{1/2}} , \quad (5)$$

can be regarded as a preprocessing filter that makes the local variance of the input image uniform and normalized to unity. This preprocessing can be called image homogenization because it makes an image homogeneous in terms of its local variances.

The second filter,

$$\widetilde{H_{\mathrm{mf}}}(x_0, y_0; f_x, f_y) = \frac{\alpha^*(f_x, f_y)}{[\mathrm{Var}(x_0, y_0)]^{1/2}} , \quad (6)$$

represents a modified matched filter, that is, the filter matched to the target modified by the input-image local variance.

In principle, image local variance can be computed optically by low-pass filtering of the squared input image intensity, which could be accomplished either by means of an optical correlator or simply by defocusing of the image. With the latter method we have to capture two images. The first image should be on focus and corresponds to the image to be analyzed. The second one should be out of focus and gives us the local variance $\mathrm{Var}(x_0, y_0)$. The final step is to evaluate the preprocessed image by electronically dividing the input scene by the local variance. This is a pixelwise operation and can be quickly evaluated electronically. Finally, this homogenized image is introduced into the input plane of the correlator.

To verify this proposal experimentally we carried out an optical experiment with computer-generated preprocessed images and used POF's instead of the modified matched filter given by Eq. (6).

An optical correlator was used for testing the image preprocessing proposed here. The experimental setup is shown in Fig. 1. Two twisted nematic liquid-crystal-type spatial light modulators (SLM's) extracted from an Epson video projector were used to display the filter and the scene in the optical correlator. We performed several optical experiments. The original and the preprocessed input scenes are represented in Fig. 2. They had 256×256 pixels. These scenes were displayed in the input SLM (SLM1 in Fig. 1) working in the phase-mostly regime. A CCD camera registered the energy distribution on the correlation plane. The object to be recognized in the first experiment, which was a small area of the whole image, is shown in Fig. 3(a). It had a diameter of 74 pixels.

Another target of 36-pixel diameter is also shown in the figure. A POF matched to the target was displayed in the second SLM. Although phase modulation of twisted nematic liquid-crystal modulators is in general cross coupled with amplitude modulation, it is possible to obtain phase-mostly modulation by proper selection of the plane of the incident polarized light and adjustment of the brightness, contrast, and color controls of the Epson video projector. In our case, the best configuration of the potentiometers was brightness 0, contrast 10, and color 3 for the scene SLM and brightness 0, contrast 10, and color 0 for the filter SLM. Figure 4(a) shows the energy distribution on the correlation plane when the original picture without preprocessing and a target of 74-pixel diameter were used. There is no distinguishable peak and, in consequence, the target could not be recognized. When the preprocessing was used, a high correlation peak appeared, as can be seen from Fig. 4(b). Note that in the correlation plane a noisy background appeared that was due to the nonlinearity introduced by the filter SLM,[12] which produced different superimposed diffraction orders. The zero order reproduced the input scene, and it is responsible for the background.
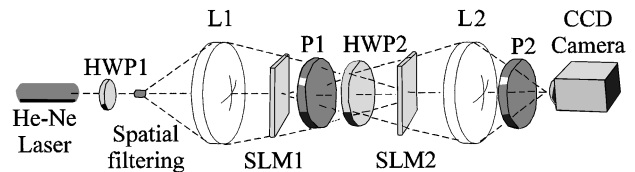

Fig. 1. Real-time optical correlator. SLM1 implements the input scene and SLM2 implements the matched filter. P1, P2, polarizers; HWP1, HWP2, half-wave plates; L1, L2, convergent lenses.
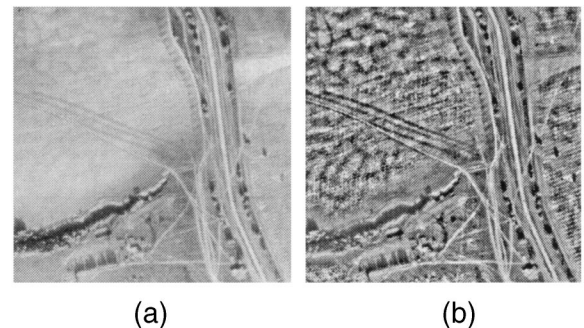

(a)   (b)

Fig. 2. (a) Original and (b) preprocessed scenes for the pattern-recognition experiments.

(a)   (b)
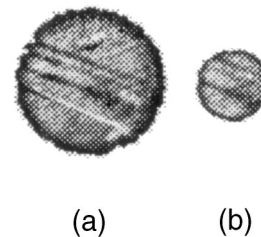
Fig. 3. Two different sizes of preprocessed target extracted from Fig. 2(b): (a) target diameter, 74 pixels; (b) target diameter, 36 pixels.

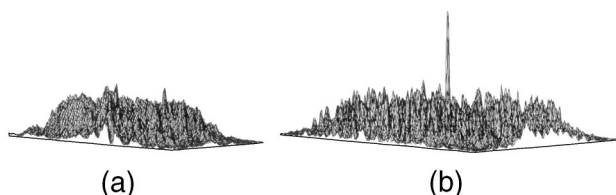(a)   (b)

Fig. 4. Correlation peaks resulting from experimental optical correlation between (a) the original scene and the original target and (b) the preprocessed scene and the preprocessed target. Both preprocessed and original targets have a diameter of 74 pixels.
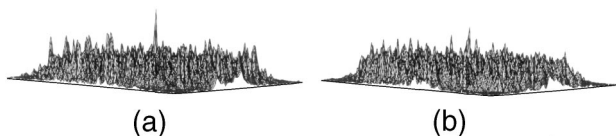

(a)   (b)

Fig. 5. Correlation peaks resulting from experimental optical correlation between preprocessed scenes and preprocessed targets of different sizes: (a) target diameter, 49 pixels; (b) target diameter, 36 pixels.

When the size of the target decreased, the energy that was concentrated in the correlation peak also decreased, and in consequence, the height of this peak diminished, as shown in Figs. 5(a) and 5(b), for which the diameters of the target were 49 and 36 pixels, respectively. If the target was too small, the correlation peak was too small compared with the background and could not be detected. Figure 5(b) shows the limit in this example. From the results one can see that the preprocessing allows us to detect targets that cannot be detected without preprocessing.

In summary, we have proposed a preprocessing of an input image, which we call homogenization, that can be seen as a zero-order approximation of the optimal filter for target location in images. We have implemented the recognition method in an optical correlator that uses two SLM's working in phase mode. In the first, the preprocessed input image is phase encoded and displayed. In the second, a phase-only filter matched to the target is implemented. We performed optical experiments showing the big improvement obtained with the proposed method when it is applied to input scenes with a spatially nonhomogeneous background.

References

1. A. B. VanderLugt, IEEE Trans. Inf. Theory **IT-10,** 139 (1964).
2. J. B. Thomas, *An Introduction to Statistical Communication* (Wiley, New York, 1969), Chap. 5.
3. B. V. K. Vijaya Kumar and L. Hassebrook, Appl. Opt. **29,** 2997 (1990).
4. J. Campos, F. Turon, L. P. Yaroslavsky, and M. J. Yzuel, Int. J. Opt. Comput. **2,** 341 (1991).
5. J. L. Horner and P. D. Gianino, Appl. Opt. **23,** 812 (1984).
6. L. P. Yaroslavsky, Appl. Opt. **31,** 1677 (1992).
7. L. P. Yaroslavsky, *Digital Picture Processing: An Introduction* (Springer-Verlag, Berlin, 1985), p. 276.
8. B. Javidi and J. Wang, Appl. Opt. **31,** 6826 (1992).
9. L. Yaroslavsky, in *Progress in Optics*, E. Wolf, ed. (Elsevier/North-Holland, Amsterdam, 1993), Vol. XXXII, pp. 145–201.
10. B. Javidi, Ph. Réfrégier, and P. Willet, Opt. Lett. **18,** 1660 (1993).
11. F. Guérault, L. Signac, F. Goudail, and Ph. Réfrégier, Opt. Eng. **36,** 2660 (1997).
12. I. Moreno, J. Campos, C. Goreki, and M. J. Yzuel, Jpn. J. Appl. Phys. **34,** 4923 (1995).