Optical implementation of the weighted sliced orthogonal nonlinear generalized correlation for nonuniform illumination conditions

Pascuala García-Martínez, Manuel Tejera, Carlos Ferreira, Daniel Lefebvre, and Henri H. Arsenault

Optical pattern recognition under variations of illumination is an important issue. The sliced orthogonal nonlinear generalized (SONG) correlation has been proposed as an optical pattern recognition tool to discriminate with high efficiency between objects. But, at the same time, the SONG correlation is very sensitive to gray-scale image variations. In a previous work, we expanded the definition of the SONG correlation to the Weighted SONG (WSONG) correlation to modify the discrimination capability in a controlled way. Here, we propose to use the WSONG when pattern recognition is obtained by means of optical correlation under nonuniform illumination. The calculation of the WSONG correlation requires the summation of many linear correlations between binary images. To implement it optically, we use a time sequential joint transform correlator. © 2002 Optical Society of America

OCIS codes: 070.5010, 100.5010, 070.4550, 100.4550.

1. Introduction

Pattern recognition consists of the detection and identification of a known pattern or target in an unknown input scene that may or may not contain the target, and the determination of the spatial location of any such target. Traditional digital pattern recognition techniques require massive computation and can be relatively slow. Optical techniques can provide inherent parallelism, ultrahigh processing speed, noninterfering communication and massive interconnection capability, and sometimes offers a significantly better alternative to the digital pattern recognition approach.[1,2]

Most optical pattern recognition techniques involve either the use of a matched filter based correlator or of a joint transform correlator (JTC).[3,4] A matched filter based correlator uses Fourier domain complex filter synthesis, whereas a JTC utilizes spatial domain filter synthesis. It is difficult to implement the matched filter based correlator for real-time applications because a complex filter must be synthesized, and the filter alignment along the optical axis is critical. On the other hand, a JTC is inherently suitable for real-time matching and tracking operations because no complex filter is needed.[5] The JTC has gone through many improvements and has shown its usefulness for pattern recognition and target detection.[6] Some of those improvements are based on increasing the discrimination capability by applying various mathematical functions, such as thresholding, coding, or nonlinear functions in the power spectrum domain.[7–10] Moreover, some nonlinear correlation based on the addition of binary slices of the power spectra of the input scene and of the reference object have been implemented optically using a JTC.[11,12] Such nonlinear correlations provide higher discrimination capabilities than other JTC techniques.

Those correlations are based on calculating the linear correlations between binary versions of the reference object and of the target. If the binarization is a threshold decomposition, we will obtain the morphological correlation (MC),[11,13] whereas if we use a slicing process we implement the sliced orthogonal

M. Tejera is with both Dpt. D'Òptica, Universitat de València, 46100-Burjassot, València, Spain and Lab. Óptica Coherente, Instituto Superior Politécnico "José Antonio Echeverría", c/ 114 s/n. Marianao, La Habana, Cuba. P. Garcia-Martínez and C. Ferreira are with the Dpt. D'Òptica, Universitat de València, 46100-Burjassot, València, Spain. D. Lefebvre and H. Arsenault are with the COPL, Université Laval, Ste-Foy, Quebec G1K 7P4, Canada. E-mail address for P. Garcia-Martínez is pascuala.garcia@uv.es.

Received 8 May 2002; revised manuscript received 5 August 2002.

0003-6935/02/326867-08$15.00/0

nonlinear generalized (SONG) correlation.[12] The threshold decomposition is understood as a binarization process of a N gray level image into $N - 1$ slices that takes value 1 if the gray-level image is higher or equal to a certain threshold, whereas the slicing process gives value 1 if the gray level of the image is equal to a certain value. These correlations have advantages and disadvantages because of their high discrimination capability and their low tolerance to some common distortion sources, such as additive noise and variable illumination. There are previous works based on proposing different image binarization techniques for correlation. In Ref. 14 the authors show different post-processing operations to reduce the computational burden of correlation-based methods.

Changes of illumination modify the gray-level distribution of the images, especially if the images are captured live from the outside world. Alam and Karim[15] have studied the correlation performance of classical, binary, and fringe adjusted JTCs under varying illumination of the input scene. Other pattern recognition methods using the JTC have studied the influence of the illumination, and techniques invariant to such changes have been proposed.[16,17] However, the proposed illumination model is a uniform illumination over the input scene, so the object alteration consists of multiplying its gray-level distribution by a constant factor. In most real-world applications, the illumination conditions are not uniform. We propose an illumination model based on multiplicative and additive nonuniform noise.

As we have pointed out in previous papers, a weak point of the SONG correlation[12,18] is the strong dependence on the specific gray levels of the objects. We recently showed the optimality of the SONG correlation in a maximum-likelihood sense for images degraded with substitutive noise.[19] Moreover, we have also considered an extension of the SONG correlation to a weighted sliced orthogonal nonlinear generalized (WSONG) correlation to obtain detection under small changes of illumination.[20] Now, taking the noise model that we used previously,[20] we introduce a time sequential JTC to implement optically the WSONG correlation. Experimental results are given.

2. Weighted Sliced Orthogonal Nonlinear Generalized Correlation

The SONG correlation and decomposition was described in Refs. 12 and 18. Nevertheless, for completeness, we repeat fundamental expressions in the following: A two-dimensional image with discrete gray levels can be decomposed into a sum of disjoint elementary images $e_i[f(x, y)]$ defined as

$$e_i[f(x, y)] = \begin{cases} 1 & f(x, y) = i \\ 0 & f(x, y) \neq i \end{cases}. \tag{1}$$

The SONG decomposition of $f(x, y)$ is

$$f(x, y) = \sum_{i=0}^{N-1} i e_i[f(x, y)], \tag{2}$$

where N is the number of gray levels.

The orthogonal property of the SONG decomposition implies that

$$\langle e_i[f(x, y)], e_j[f(x, y)]\rangle = \delta_{ij}\|e_i[f(x, y)]\|^2. \tag{3}$$

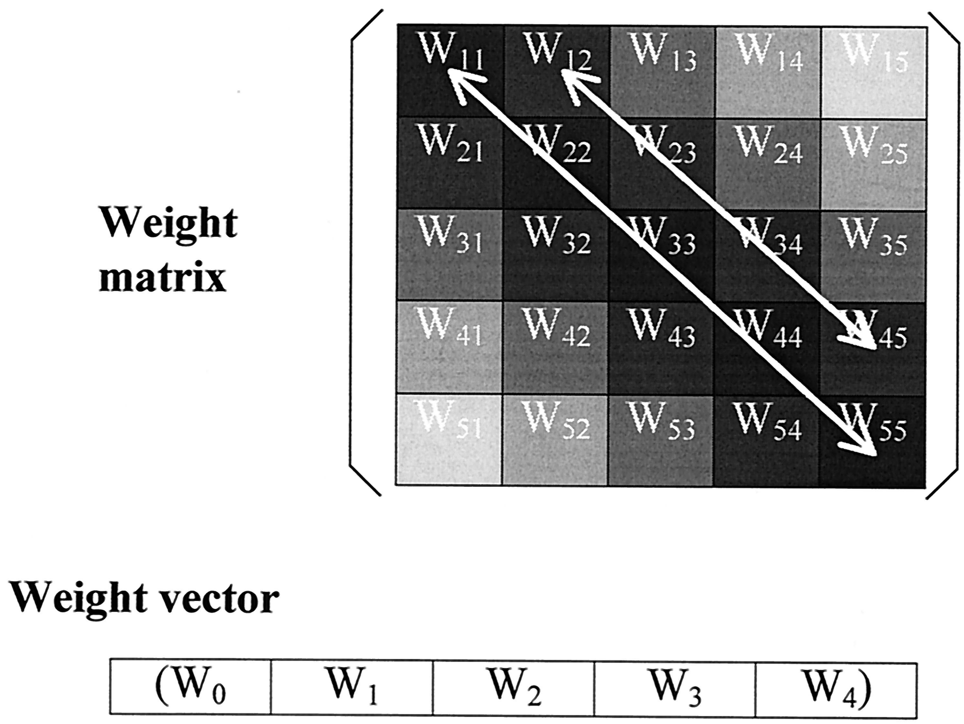


Fig. 1. Diagram of the conversion from the weight matrix into a weight vector.

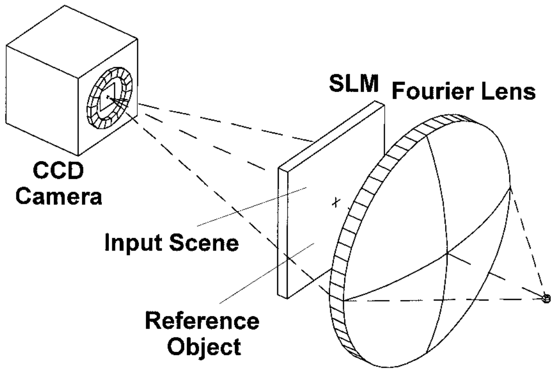


Fig. 2. Classical joint transform correlator.

The Kronecker delta function δ_{ij} has the value one when $i = j$, and the value zero when $i \neq j$, where i and j are integers. So, the squared L_2 norm of $e_i[f(x, y)]$ at the origin (0, 0) is the number of pixels of $f(x, y)$ that have a given gray value i.

The classical linear correlation ($\otimes$) between two functions, $g(x, y)$ and $f(x, y)$, can be expressed in terms of those elementary binary images[18]

$$g(x, y) \otimes f(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} ij e_i[g(x, y)] \otimes e_j[f(x, y)]. \quad (4)$$

The SONG correlation can be expressed in terms of a matrix.[18] In this matrix representation the SONG correlation is defined as the sum of multiple linear correlations between different binary slices of the two functions, $g(x, y)$ and $f(x, y)$, where the matrix elements are weighted by the weight coefficients W_{ij}

$$\Omega_{g,f}(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij} e_i[g(x, y)] \otimes e_j[f(x, y)]. \quad (5)$$

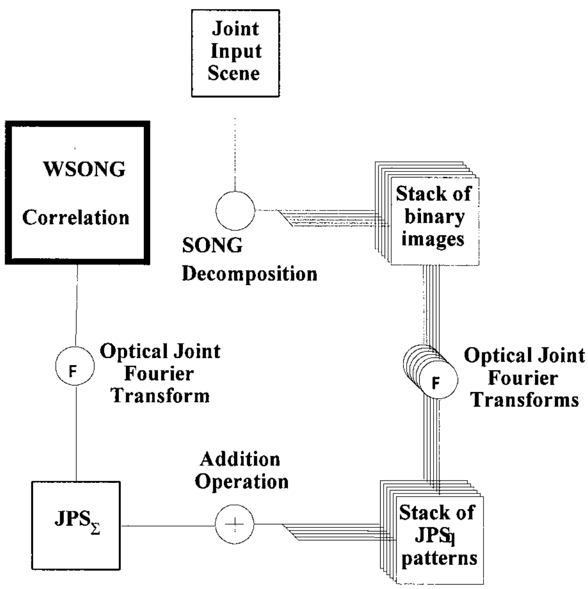


Fig. 3. Block diagram of the opto-electronic WSONG correlation.

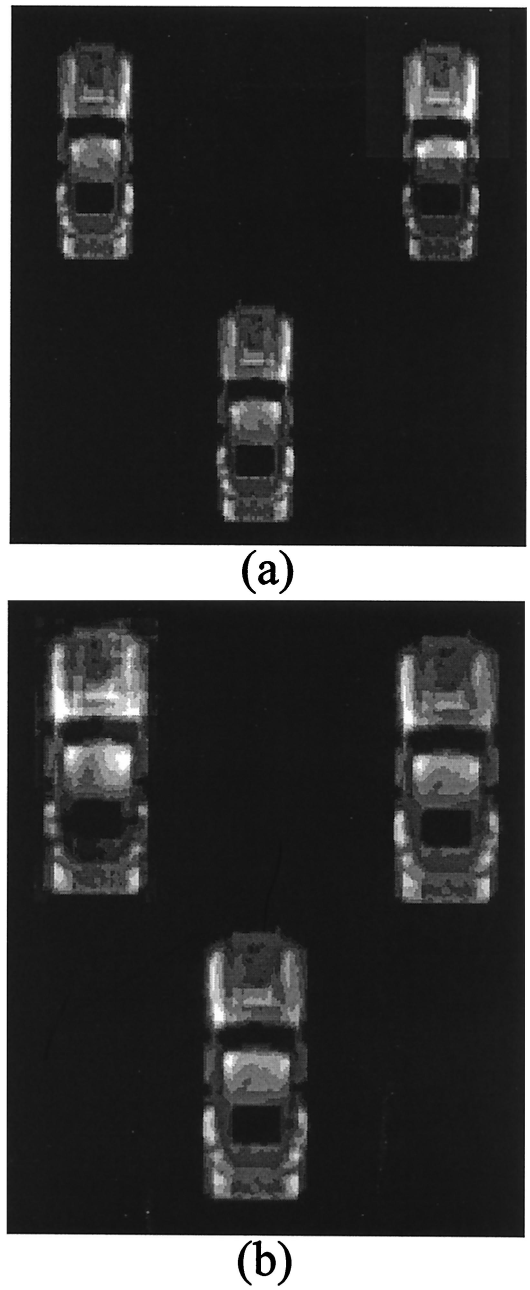


Fig. 4. (a) Joint input scene containing the input scene (top) and the reference object (bottom) for the shadow effect illumination. (b) Joint input scene containing the input scene (top) with the nonuniform degraded image and the reference object (bottom).

The SONG correlation matrix is

$$\Omega_{g,f}(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} R_{gf}{}^{ij}(x, y) = \Theta \begin{bmatrix} R^{0,0} & R^{0,1} & \cdots & R^{0,N-1} \\ R^{1,0} & R^{1,1} & \cdots & R^{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ R^{N-1,0} & R^{N-1,1} & \cdots & R^{N-1,N-1} \end{bmatrix}, \quad (6)$$

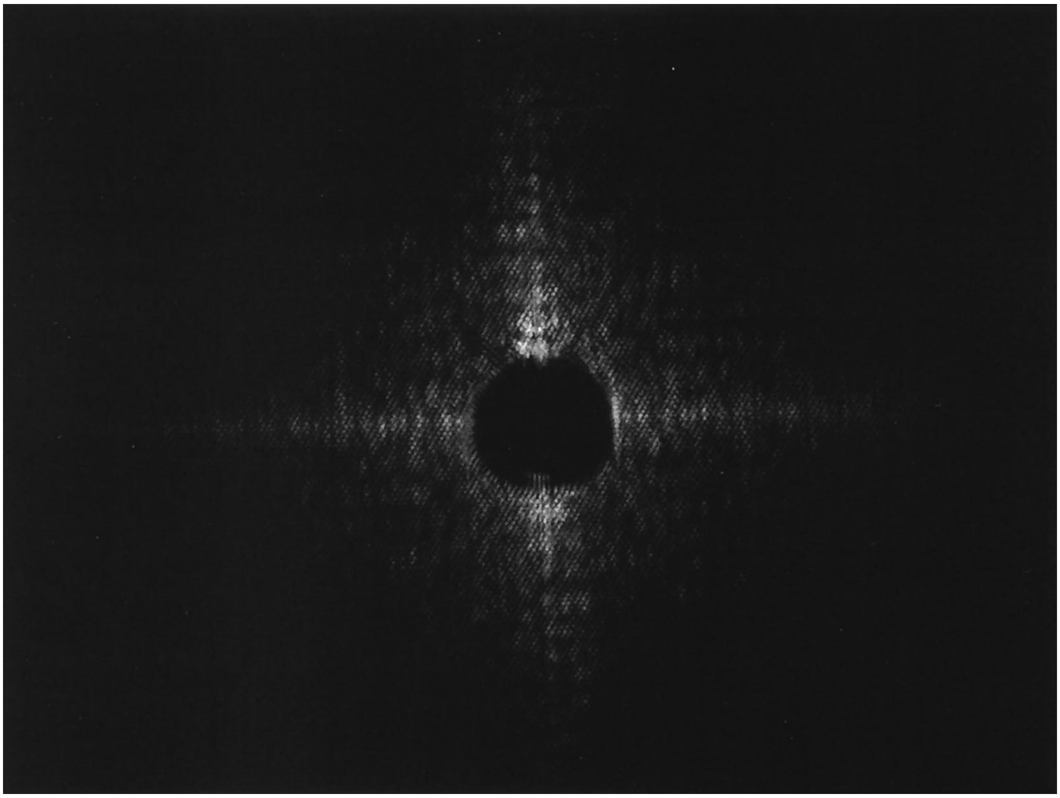


Fig. 5. Joint power spectra (JPS_{Σ}) example for the WSONG correlation.

where $R^{ij}(x, y) = W_{ij}e_i[g(x, y)] \otimes e_j[f(x, y)]$ and the symbol Θ represent the sum of all of the matrix terms.

In early work on the SONG correlation[12,18] the above expression was simplified by setting $W_{ij} = \delta_{ij}$, so giving the same importance to all of the gray levels, and reducing the SONG correlation to the trace of the SONG matrix

$$\Omega_{gf}{}^{P}(x, y) = \sum_{i=0}^{N-1} e_i[g(x, y)] \otimes e_i[f(x, y)]. \quad (7)$$

In this expression, only the gray levels corresponding to image levels having the same value are correlated together after having their pixel values set equal to unity. So the correlation is a sum of correlations between the corresponding binarized gray levels. This is the SONG correlation as used in our previous papers.[12,18]

Effects of the weights in the SONG correlation between gray-level images. From Eq. (7), the main weak point of this correlation is that a change in illumination will change the gray-level distribution and the correlation result. So the limitation of our scheme is the sensitivity to small intensity or gray-level changes that are present in real-world images. To avoid this situation, we consider the importance of weights in the SONG matrix.

We have recently shown that pattern-recognition performance is maintained using a WSONG correlation[20] when the objects are slightly degraded. The idea is to perform not only the correlations that define the principal diagonal but others as well. The nonprincipal diagonals represent the linear correlations between the two slices of the reference and the nonuniform illuminated target from different gray levels. Those diagonals represent the contribution of those pixels with a gray-level displacement from its original value.

Let us assume that the degradation consists of slow

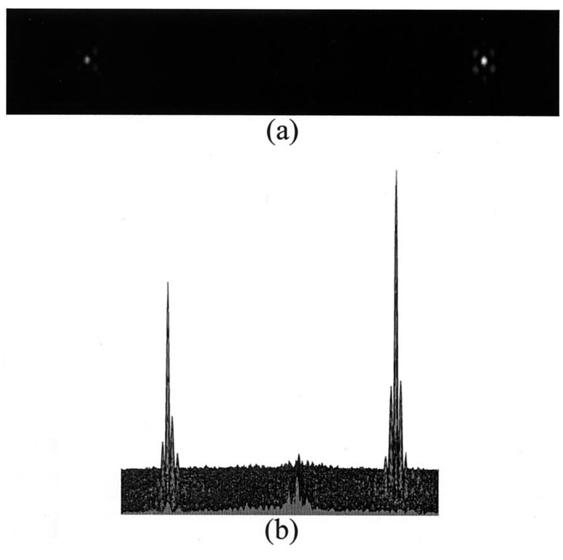
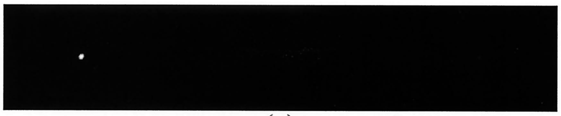
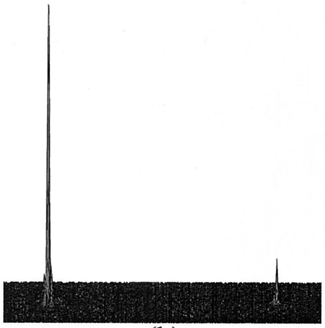


Fig. 6. Experimental output plane containing the optical linear correlation of the scene shown in Fig. 4(a): (a) Output correlation plane covering an area around the correlation peaks, (b) 3-D plots.



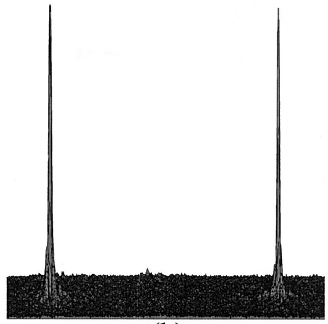
(a)



(b)

Fig. 7. Experimental output plane containing the optical SONG correlation of the scene shown in Fig. 4(a): (a) Output correlation plane covering an area around the correlation peaks, (b) 3-D plots.

(a)



(b)

Fig. 8. Experimental output plane containing the optical WSONG correlation of the scene shown in Fig. 4(a): (a) Output correlation plane covering an area around the correlation peaks, (b) 3-D plots.

changes in the gray-level distribution of the reference object due to small illumination changes. If $g(x, y)$ represents a degraded reference object and $f(x, y)$ the nondegraded object, for recognition or comparison problems the summation of the $R^{ij}(x_o, y_o)$ matrix terms is related to the number of pixels with i-gray level in $g(x_o, y_o)$ that have moved to the j level on image $f(x_o, y_o)$ at the same position.

The values of the weight coefficients for the various diagonals represent the contribution of these value-shifted pixels to the total amount of correct pixels in the correlation. But for recognition purposes, the value of this contribution must be a nonincreasing function of the pixel value shift. For the sake of simplicity we define a vector of weight coefficients $W = (W_0, W_1, \ldots, W_{N-1})$, whose indices are numbers that correspond to the distances of the respective diagonals to the principal diagonal. Figure 1 shows the relationship between the weight-correlation matrix and the weight vector. The weight vector is chosen to have the same values for the diagonals that are equidistant from the principal diagonal.

3. Optical Implementation of the WSONG Correlation

A. Nonuniform Illumination Model

Changes of illumination are a common and natural phenomenon. Images are obtained during the entire day at different moments. So, many shadow effects or specular reflections can appear in the images. Several studies about pattern recognition under illumination changes considered only the multiplication of the input image by a constant factor.[14,16,21] However, when the illumination varies over the image, the internal gray level input-image distribution may change in a nonuniform way. For this reason, other filters have been designed to recognize objects that are heavily distorted by varying illumination.[9,22]

In this paper, we define a gray-scale degradation or illumination-variation model. The illumination variations that we consider are spatially correlated and their range is limited. With this model we try to be as close as possible to real illumination conditions, taking into account sources of illumination variations not considered in other models. The function $g(x, y)$ represents a degraded version of the object $f(x, y)$

$$g(x, y) = \alpha(x, y) f(x, y), \tag{8}$$

with

$$\alpha(x, y) = 1 + A\lambda(x, y)_{lx,ly} \tag{9}$$

So, we are considering multiplicative and correlated noise. The parameter A is the amplitude of the intensity variations, $\lambda(x, y)_{lx,ly} \in (1, -1)$ is a random Gaussian distribution function with correlation lengths l_x and l_y. For simplicity we used $l = l_x = l_y = 5$ in our experiments. To correlate the noise we used a mean low pass filtering.

B. Time Sequential Joint Transform Correlator for Nonlinear Binary Correlations

A classical JTC is shown in Fig. 2, where the reference and the input scene are introduced in the input plane by use of a spatial light modulator (SLM), such as a liquid-crystal television display. From Eq. (5) and Eq. (6), the WSONG correlation is defined as the sum of the amplitudes of the linear correlations between many binary slices of the input scene and of the target, weighted by the weight vector. This amplitude summation is carried out by the JTC by use of using the same system as for the MC and the SONG correlation.[11,12] The optical implementation of those nonlinear correlations are based on a time-sequential JTC. The block diagram is shown in Fig. 3. Each pair of elementary binary joint input slices (one slice from the reference object and one from the input scene) are located next to each other in the input plane. For each pair, the joint power spectrum (JPS) is performed. The JPSs are multiplied by the weight vector and then added. The expression for the final JPS_Σ is

$$\begin{aligned} JPS_\Sigma(u, v) &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} JPS_{ij}(u, v) \\ &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij} |FT(\{e_i[g(x, y)] \\ &\quad + e_j[f(x, y)]\})|^2, \end{aligned} \tag{10}$$

where FT{ } means the Fourier transform. This distribution is input again to the SLM and a second Fourier transformation gives the WSONG correlation.

If the weight vector is $W = (1, 0, 0, 0 \ldots 0)$, we are optically implementing the SONG correlation.[12,18] In this case only $N - 1$ linear correlations are performed optically, that is, only the principal diagonal (D_0) terms are taken into account. This operation is extremely selective because only the pixels whose gray levels are in the same position for the input scene and for the reference object are counted. But if, for instance, we use a weight vector $W = (1, 1, 0, 0 \ldots 0)$, we have to implement D_1 and D_{-1} diagonal correlations in addition to the main D_0 diagonal correlations. This WSONG correlation is more robust to gray-level variations caused by changes of illumination.

We used two different kinds of nonuniform illumination conditions. The joint input scenes are shown in Fig. 4. The objects have 9 gray levels, (i, j = [0, 32, 64, 96, 128, 160, 192, 224, 255]). The input scenes are located in the upper area and the reference object in the lower area. The input scene of Fig. 4(a) contains a reference object that is located on the left-hand side, and a degraded version of the reference object located on the right-hand side. To implement the WSONG correlation taking the contribution of the diagonals D_0, D_1, and D_{-1}, we define, in Fig. 4(a), the degraded object by adding gray-level 32 to the upper part of the reference object and subtracting gray-level 32 from the lower part. So, we simulate shadows in the reference object. It is true that subtracting and adding a particular gray level is somewhat heuristic, however, with this simple specific example we show that the WSONG correlation using the $W = (1, 1, 0, \ldots, 0)$ vector is the same for the two objects.

The other case of nonuniform illumination is shown in Fig. 4(b). The reference object is located on the right-hand side and the degraded object is on the left-hand side. In this case we have used the nonuniform illumination variation defined in Eq. (8). We detect both objects using the WSONG correlation with $W = (1, 1, 0, \ldots, 0)$.

4. Optical Experimental Results

The SLM used for the optical experiments is a 800 × 600 pixel LCTV from SONY. The CCD camera is a Sensys camera by Photometrics with 12 bit resolution. This camera allows 4096 gray levels. This is convenient because we are adding the contribution of many JPSs. Unfortunately, at the end there is a conversion to 8 bits due to the SLM. This only affects when we display the total distribution JPS_Σ on the SLM. Figure 5 shows a JPS_Σ when an opaque spot is inserted at the center of the power spectrum to protect the detector from blooming and from other saturation effects.[23]

We now compare the discrimination capability of the WSONG, of the SONG, and of the common linear correlation for the two kinds of illumination conditions.

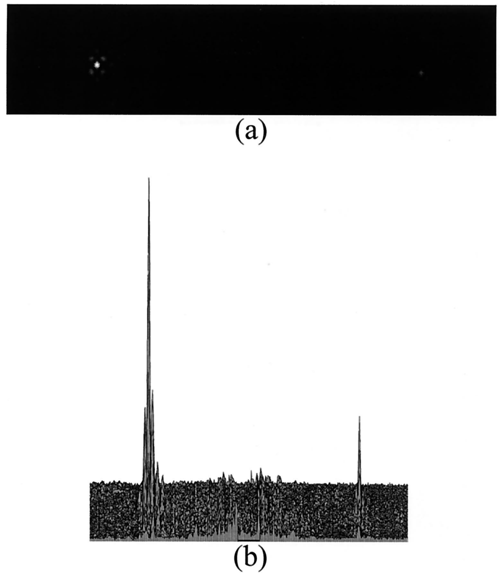


Fig. 9. Experimental output plane containing the optical linear correlation of the scene shown in Fig. 4(b): (a) Output correlation plane covering an area around the correlation peaks, (b) 3-D plots.

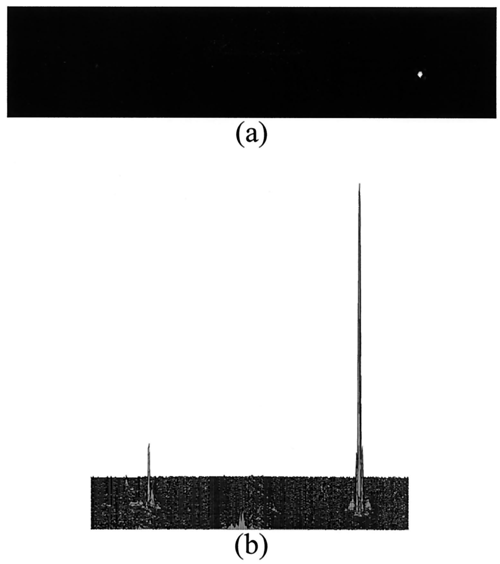
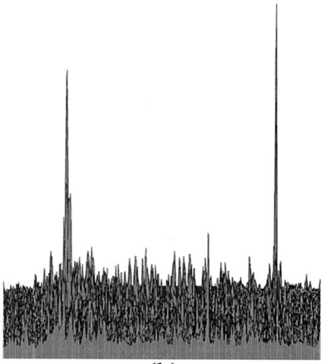


Fig. 10. Experimental output plane containing the optical SONG correlation of the scene shown in Fig. 4(b): (a) Output correlation plane covering an area around the correlation peaks, (b) 3-D plots.

(a)



(b)

Fig. 11. Experimental output plane containing the optical WSONG correlation of the scene shown in Fig. 4(b): (a) Output correlation plane covering an area around the correlation peaks, (b) 3-D plots.

tions. Figure 6 shows the result for the common linear correlation for the input scene of Fig. 4(a). The optical results are plotted only for the regions of interest of the correlation plane. We have also plotted the three-dimensional (3-D) profile around the zone of interest. Figure 7 and Fig. 8, respectively, show the results for the SONG correlation and for the WSONG correlation. The common linear correlation detects the degraded object because it is brighter than the reference object, and the linear correlation is proportional to intensity. The SONG correlation detects the reference object with high discrimination.[12] The WSONG correlation detects the two objects with the same value. This is not surprising because we are adding D_0, D_1, and D_{-1}. Those diagonals contain the gray-level variation between $+32$ and -32.

For the illumination variation or gray-level degradation shown in Fig. 4(b), the results of Fig. 9, 10, and 11 correspond respectively to the common linear correlation, the SONG correlation, and the WSONG correlation. Note that once again neither the linear nor the SONG correlation detects both objects at the same time, whereas the WSONG correlation gives almost the same correlation peak intensity for the two objects with different illumination. Then, we achieved tolerance for the recognition of objects under different illumination conditions.

5. Conclusion

We have optically implemented the WSONG correlation for the detection of nonuniform illuminated objects as a valid alternative of tolerance to the high selectivity of the optical SONG correlation. The idea of using more than one diagonal of the SONG correlation matrix for the total amount of correlation is relevant if we need to detect under conditions presenting a nonuniform illumination variation. To obtain the summation of the amplitude of all the linear correlations, we have used an optical time sequential joint transform correlator. The summation operation is performed in the joint power spectrum level. We have compared the performances of the WSONG correlation with the common SONG correlation introduced in our previous works and the common linear correlation. The possibility to define this nonlinear correlation in terms of the common linear correlation between binary images motivate us to use, in the future, fast spatial light modulators like Ferroelectric SLM, which may work in binary regime. In this paper we use common twisted nematic SLM and a 12-bit camera. However, the results can be also implemented by use of a conventional 8-bit camera. One of the main limitations of the algorithm is to decide how many gray levels can be defined in the image to perform the minimum number of correlations. This further study is under development by us at this moment.

This work has been supported by the Spanish Secretaría de Estado de Política Científica y Tecnológica, Dirección General de Investigación, project BFM2001/3004.

References

1. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996).
2. B. G. Boone, *Signal Processing using Optics: Fundamentals, Devices, Architectures, and Applications* (Oxford University, New York, 1998).
3. A. Vander Lugt, "Signal Detection by Complex Spatial Filtering," IEEE Trans. Inf. Theory, **IT-10,** 139–145 (1964).
4. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," Appl. Opt. **5,** 1248–1249 (1966).
5. P. Purswosumarto and F. T. S. Yu, "Robustness of joint transform correlator versus Vander Lugt correlator," Opt. Eng. **36,** 2775–2780 (1997).
6. *Selected Papers on Optical Pattern Recognition Using Joint Transform Correlation*, M. S. Alam, ed. SPIE Milestone Series, **MS157** (1999).
7. B. Javidi, J. Wang, and Q. Tang, "Nonlinear joint transform correlators," Pattern Recogn. **27,** 523–542 (1999).
8. A. Tanone, C.-M. Uang, F. T. S. Yu, E. C. Tam, and D. A. Gregory, "Effects of thresholding in joint-transform correlation," Appl. Opt. **31,** 4816–4822 (1992).
9. H. Sjöberg, B. Noharet, L. Wosinski, and R. Hey, "Compact optical correlator: preprocessing and filter encoding strategies applied to images with varying illumination," Opt. Eng. **37,** 1316–1324 (1998).
10. E. Perez, K. Chalasinska-Macukow, K. Styczynski, R. Kotynski, and M. S. Millan, "Dual nonlinear correlator based on computer controlled joint transform processor: digital analysis and optical results," J. Mod. Opt. **44,** 1535–1552 (1997).
11. P. Garcia-Martinez, D. Mas, J. Garcia, and C. Ferreira, "Nonlinear morphological correlation: optoelectronic implementation," Appl. Opt. **37,** 2112–2118 (1998).
12. P. Garcia-Martinez, H. H. Arsenault, and S. Roy, "Optical implementation of the sliced orthogonal nonlinear generalized correlation for images degraded by nonoverlapping background noise," Opt. Commun. **173,** 185–193 (2000).
13. P. Maragos, "Morphological correlation and mean absolute error criteria," IEEE Trans. Acoust., Speech Signal Process. **3,** 1568–1571 (1989).

14. W. C. Hasenplaugh and M. A. Neifeld, “Image binarization techniques for correlation-based pattern recognition,” Opt. Eng. **38,** 1907–1917 (1999).
15. M. S. Alam and M. A. Karim, “Joint-transform correlation under varying illumination,” Appl. Opt. **32,** 4351–4356 (1993).
16. S. Jutamulia, G. M. Storti, D. A. Gregory, and J. C. Kirsch, “Illumination-independent high efficient joint-transform correlation,” Appl. Opt. **30,** 4173–4175 (1991).
17. B. Javidi, J. Li, A. H. Fazlollahi, and J. Horner, “Binary nonlinear joint transform correlator performance with different thresholding under unknown illumination conditions,” Appl. Opt. **34,** 886–896 (1995).
18. P. Garcia-Martínez and H. H. Arsenault, “A correlation matrix representation using the sliced orthogonal nonlinear generalized decomposition,” Opt. Commun. **172,** 181–192 (1999).
19. P. Garcia-Martínez, Ph. Réfrégier, H. H. Arsenault, and C. Ferreira, “Maximum likelihood for target location in the presence of substitutive noise,” Appl. Opt. **40,** 3855–3860 (2001).
20. M. Tejera, P. Garcia-Martinez, C. Ferreira, D. Lefebvre, and H. H. Arsenault, “Weighted nonlinear correlation for controlled discrimination capability,” Opt. Commun. **201,** 29–37 (2002).
21. F. M. Dickey and L. A. Romero, “Normalized correlation for pattern recognition,” Opt. Lett. **16,** 1186–1188 (1991).
22. O. Germain and P. Réfrégier, “Optimal snake-based segmentation of a random luminance target on a spatially disjoint background,” Opt. Lett. **22,** 1845–1847 (1996).
23. S. Jutamulia and D. A. Gregory, “Soft blocking of the dc term in Fourier optical systems,” Opt. Eng. **37,** 49–51 (1998).