Projection-invariant pattern recognition with a phase-only logarithmic-harmonic-derived filter

A. Moya, D. Mendlovic, J. García, and C. Ferreira

A phase-only filter based on logarithmic harmonics for projection-invariant pattern recognition is presented. This logarithmic-harmonic-derived filter is directly calculated in the Fourier plane. With respect to normal logarithmic-harmonic filters it provides a smaller variation of the correlation intensity with the projection factor of the target. Computer and optical experiments are presented.

Key words: Projection invariance, pattern recognition, phase-only filters.

1. Introduction

Pattern recognition is one of the important problems, as yet unsolved, in robotics and automization of processes. The most general method for recognition of objects is based on computing a set of characteristics of the object. A comparison with a data base may serve to determine if a given object should or should not be detected. This method, although very general, is complicated to implement and is usually performed by numerical calculations. On the other hand, optical methods involve an inherent parallelism of the processing, are easily implemented, and give a fast processing time. A widely used optical method is based on obtaining the correlation between the input scene, which contains or not the object to be detected, and the target, or a function connected with it. The correlation provides a measure of similarity connected with the mean-squared difference between the patterns being correlated. The usual VanderLugt correlator, based on matched filtering,[1] is the main tool for correlation. Usually correlation has as a main drawback the high sensitivity of the system to deformations of the object.

One approach to making pattern recognition more flexible is based on defining invariances. Typically two invariances are considered: scale and rotation. There is a wide variety of algorithms in the literature that are used to cope with these invariances. In particular, the ones based on orthogonal expansions are widely used, being conceptually simple and of easy implementation. For the case of rotation invariance the circular-harmonic expansion[2] is used, whereas the Mellin radial-harmonic expansion[3] serves for scale invariance. Combining the two invariances simultaneously is much more complicated, although some methods have been proposed.[4–6] The cases in which scale or rotation invariance are needed are numerous and obvious. Rotation of the same object on the image plane, or different scales that are due to varying distances to the grabbing systems, are common.

When deformations that involve a scale change are considered, a problem arises. The energies of objects that are differently distorted are also different. A linear system can hardly provide the same correlation output in this case. For this reason a main limitation is imposed: the distortion range must be limited to a certain range. A full-scale or projection [one-dimensional (1D) scaling] invariance is not possible.

For the case of scale invariance, moreover, if an orthogonal set of functions is to be used, an additional constraint is fed into the definition. The correlation is independent of the scale of the target in the sense that the correlation intensity distribution is scaled with the same factor as the input pattern. The magnification factor affects not only the spatial distribution but also its intensity. The output intensity is not strictly invariant but depends quadratically on the scale factor of the object.[3]

A. Moya, J. García, and C. Ferreira are with the Departamento Interuniversitario de Óptica, Universitat de València, Burjasot 46100, Valencia, Spain. D. Mendlovic is with the Faculty of Engineering, Tel-Aviv University, Tel Aviv 69978, Israel.

Received 21 April 1995; revised manuscript received 27 October 1995.
0003-6935/96/203862-06$10.00/0

The variation of the correlation-peak intensity with the scale factor can be avoided by use of a filter derived from the previous idea but elimination of the orthogonality condition.[7] This approach relies on using a phase-only filter. This type of filter does not block any part of the energy of the input; as the correlation is forced to have the same spatial distribution, the peak intensity is kept almost constant.

A third invariance that is often considered is projection invariance. This deformation can arise when images are recorded from remote objects that have different rotations around an axis perpendicular to the camera–object line. The logarithmic-harmonic (LH) functions were proposed for dealing with this deformation.[8,9] The harmonic functions involved are analogous to those of the Mellin radial harmonics, when expressed for one dimension, aside from the intensity of the correlation distribution that depends inversely on the 1D scale factor. Owing to this dependence, only images that have suffered small tilts can be properly detected before the correlation peak decreases to undetectable levels.

In this paper we propose the use of a phase-only LH-derived filter for projection-invariant pattern recognition. The filter is designed directly in the Fourier plane. With this filter the correlation-intensity distributions present a 1D scaling identical to that of the input object, and the peak intensity varies slowly with this factor. The theory for deriving the filter is presented in Section 2. The performance is checked by computer simulations and optical experiments, the results being presented in Sections 3 and 4, respectively. In Section 5 we present the conclusions.

2. Phase-Only Logarithmic-Harmonic-Derived Filter

In order to obtain projection-invariant pattern recognition of a two-dimensional (2D) object, $f(x, y)$, one solution is to decompose the object into a set of logarithmic harmonics[8]

$$f(x, y) = \frac{1}{2L} \sum_{N=-\infty}^{+\infty} f_N(y) |x|^{i2\pi N-(1/2)} \tag{1}$$

with

$$f_N(y) = \int_{-1}^{\exp(-L)} f(x, y)(-x)^{-i2\pi N-(1/2)} \mathrm{d}x + \int_{\exp(-L)}^{1} f(x, y) x^{i2\pi N-(1/2)} \mathrm{d}x. \tag{2}$$

A single harmonic of order N is used to generate a matched filter of the object $f(x, y)$. Thus object versions tilted around the y axis, i.e., having a scale factor along the x axis, can be recognized in a scene.

Taking $g(x, y)$ as a tilted version of $f(x, y)$, $g(x, y) = f(\alpha x, y)$, where α is the 1D scale factor, we obtain the next correlation distribution[8]:

$$C_{fg}(x, y) = \alpha^{-1/2} \exp(i2\pi N \ln \alpha) C_{ff}(x, y). \tag{3}$$

The correlation intensity is inversely proportional to the scale factor α, reducing the true target correlation peaks according to the projection factor α. Thus misdetection owing to large scale factors could appear.

To solve this objection, one would like to obtain a correlation function of the type

$$C(\alpha) = C_0 \exp[i\xi(\alpha)], \tag{4}$$

where C_0 is a constant and $\xi(\alpha)$ is any real function of the scale factor α.

The solution, in a way similar to that employed by Rosen and Shamir[7] to obtain scale-invariant pattern recognition, can be achieved as follows. Denoting $F(u, v)$ as the Fourier transform of the untilted target $f(x, y)$ and $H(u, v)$ as the filter function, we obtain the center correlation value corresponding to any object for which $g(x, y) = f(\alpha x, y)$:

$$C(\alpha) = \frac{1}{\alpha} \int_{-u_1}^{-u_0} \int_{-v_1}^{v_1} F(u/\alpha, v) H^*(u, v) \mathrm{d}u \mathrm{d}v + \frac{1}{\alpha} \int_{u_0}^{u_1} \int_{-v_1}^{v_1} F(u/\alpha, v) H^*(u, v) \mathrm{d}u \mathrm{d}v, \tag{5}$$

where the finite size of the filter is taken into account. Moreover, the filter performance is improved because the low frequencies are removed (Fig. 1).

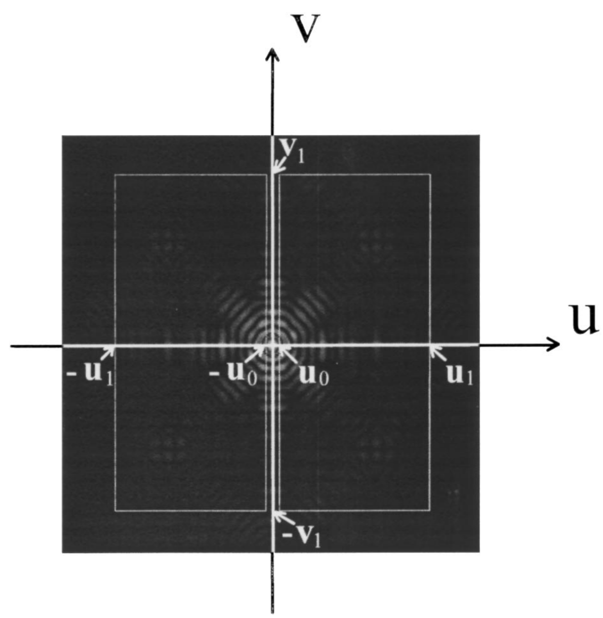


Fig. 1. Fourier plane of the target function. The limits of the filter are indicated by the signatures u_1, $-u_1$, v_1, and $-v_1$. Frequencies between $-u_0$ and u_0 are removed.

Performing a change of the integration variable, $\eta = u/\alpha$, in Eq. (5) results in

$$C(\alpha) = \int_{-u_1/\alpha}^{-u_0/\alpha} \int_{-v_1}^{v_1} F(\eta, v) H^*(\alpha\eta, v) \mathrm{d}\eta \mathrm{d}v + \int_{u_0/\alpha}^{u_1/\alpha} \int_{-v_1}^{v_1} F(\eta, v) H^*(\alpha\eta, v) \mathrm{d}\eta \mathrm{d}v. \quad (6)$$

In this way, the scale variance is included in the filter. The next step is the obtention of a filter function that does not depend on the scale factor. This property can be accomplished provided that the filter is defined as follows:

$$H_q^*(u, v) = \Psi(v) \left(\frac{u}{u_0} \right)^{iq/\varphi}, \quad (7)$$

where $\Psi(v)$ is a function that contains the information of the v frequency in the object function, q is the LH frequency, and φ is a weight constant,

$$\varphi = \frac{1}{2\pi} \ln\left(\frac{u_1}{u_0} \right). \quad (8)$$

It can easily be checked that

$$H_q^*(\alpha u, v) = \exp\left[i \frac{q}{\varphi} \ln(\alpha) \right] H_q^*(u, v). \quad (9)$$

Then, a scale factor in the target results in only a constant phase factor in the center correlation value. Substitution of this filter into correlation equation (6) yields

$$C(\alpha) = \alpha^{iq/\varphi} \int_{-u_1/\alpha}^{-u_0/\alpha} \int_{-v_1}^{v_1} F(\eta, v) \Psi(v) \left(\frac{-\eta}{u_0} \right)^{iq/\varphi} \mathrm{d}\eta \mathrm{d}v + \alpha^{iq/\varphi} \int_{u_0/\alpha}^{u_1/\alpha} \int_{-v_1}^{v_1} F(\eta, v) \Psi(v) \left(\frac{\eta}{u_0} \right)^{iq/\varphi} \mathrm{d}\eta \mathrm{d}v. \quad (10)$$

$\Psi(v)$ may have any arbitrary definition without altering the tilt invariance of the filter. To maximize the correlation output, we select it to compensate the phase in both integrals in Eq. (10) in a way similar to the regular matched filter. The simplest solution of $\Psi(v)$ has two different terms, one for each integral:

$$\Psi_1(v) = \exp\left\{ -\arg\left[\int_{-u_1}^{-u_0} F(\eta, v) \left(\frac{-\eta}{u_0} \right)^{iq/\varphi} \mathrm{d}\eta \right] \right\} \quad (11)$$

for the first term and

$$\Psi_2(v) = \exp\left\{ -\arg\left[\int_{u_0}^{u_1} F(\eta, v) \left(\frac{\eta}{u_0} \right)^{iq/\varphi} \mathrm{d}\eta \right] \right\} \quad (12)$$

for the second one.

Consequently the filter function is defined, in general, as

$$H_q^*(u, v) = \begin{cases} \Psi_1(v) \left(\dfrac{-u}{u_0} \right)^{iq/\varphi} & \text{if } -u_1 \leq u \leq -u_0 \\ \Psi_2(v) \left(\dfrac{u}{u_0} \right)^{iq/\varphi} & \text{if } u_0 \leq u \leq u_1 \end{cases}, \quad (13)$$

and the final correlation expression is identical in form to Eq. (4) except for the integration interval, which changes with the scale factor α. In particular, if $\alpha = 1$, the center correlation value is

$$C_q(1) = \int_{-v_1}^{v_1} \left| \int_{-u_1}^{-u_0} F(\eta, v) \left(\frac{-\eta}{u_0} \right)^{iq/\varphi} \mathrm{d}\eta \right| \mathrm{d}v + \int_{-v_1}^{v_1} \left| \int_{u_0}^{u_1} F(\eta, v) \left(\frac{\eta}{u_0} \right)^{iq/\varphi} \mathrm{d}\eta \right| \mathrm{d}v. \quad (14)$$

The phase of the integration through the v frequency is canceled, and a real and positive correlation results. The LH frequency q must be chosen properly to give the maximal intensity correlation value. In the ideal case it would be $C_q(\alpha) = C_q(1)$. In practice the correlation intensity depends on α through the integration limits, as can be seen in Eq. (10), but this dependence has a small influence on the final correlation intensity.

3. Computer Simulations

To check the capability of the above-defined filter for projection-invariant pattern recognition, both computer and optical experiments were carried out. The input function used in the experiments contains three different models of traffic signals, in a scene of 256×256 pixels, with four gray levels (Fig. 2). Each signal has a tilted version with $\alpha = 2$ as the 1D scale factor. The STOP signal was used to prepare the filter.

Fig. 2. Input image containing three different models of traffic signals. The signal used as the target function is depicted at the upper-left corner of the image. Every signal appears twice in the scene with two projection factors ($\alpha = 1$ and $\alpha = 2$).

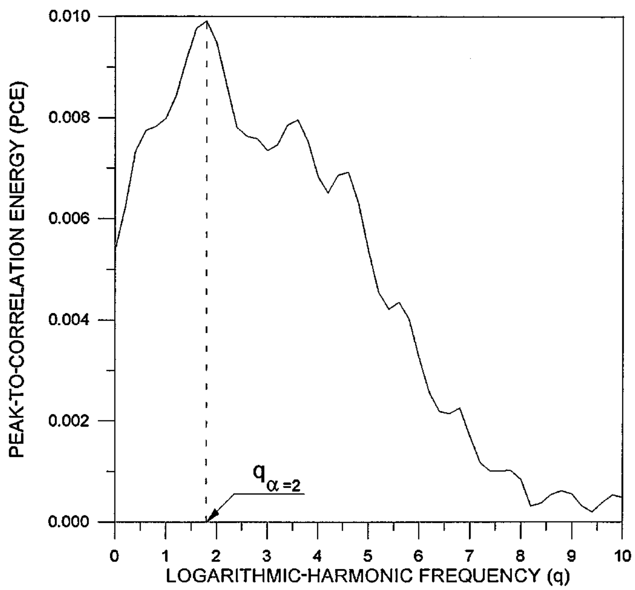


Fig. 3. Plot of the LH frequency versus the PCE. As the input image, the tilted version of the STOP signal for $\alpha = 2$ is used.

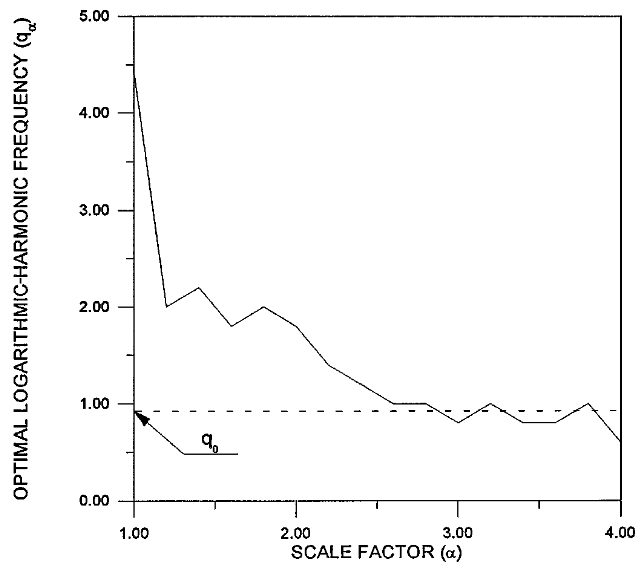


Fig. 4. Plot of the LH frequency q_α that yields the maximum PCE in Fig. 3 versus the 1D scale factor α for different tilted versions of the STOP signal. The q_α values tend to the q_0 value when α increases.

First, we have to establish a criterion to find the optimal LH frequency q. In the above-mentioned Refs. 7 and 8, this parameter is chosen to maximize the energy content of the filter. But this criterion is sensitive to variations of the scale factor α because the energy distribution for each value of the frequency of any object in the scene changes with α. To override this scale dependence, we use the peak-to-correlation energy[10,11] (PCE) distribution instead of the energy distribution of the filter. In principle we want to choose a q value that provides the highest possible PCE. Figure 3 shows a plot of the PCE as a function of q for the tilted STOP with $\alpha = 2$. Thus, the optimal value would be $q_{\alpha=2} = 1.8$. When different tilted versions of the STOP are used the corresponding graph, PCE versus q in Fig. 3 changes. This behavior is explained because the scale factor α affects the correlation intensity through the integration interval [see Eq. (10)], aside from the information loss caused by resampling.

To avoid this scale dependence, we made a plot of the q_α value that yields the maximum PCE in Fig. 3 as a function of the scale factor. The corresponding result shown in Fig. 4 shows that the q_α value stabilizes for increasing values of α. This fact serves to establish the criterion that we propose to select the q value used to prepare the filter as that to which the graph tends. For the STOP signal the graph tends to the value $q_0 = 0.9$. In the following, for simplicity, to record the filter, we use the integer value for q closest to the optimal value, i.e., $q = 1$.

With the filter matched to the untilted STOP signal with $q = 1$, the correlation with the scene in Fig. 2 is computed. The intensity correlation output is shown in Fig. 5(a). With the proposed filter the peak intensities for both versions of the true target differ about an 8% and a 44% threshold, which is enough to reject the peaks corresponding to the other objects in the scene.

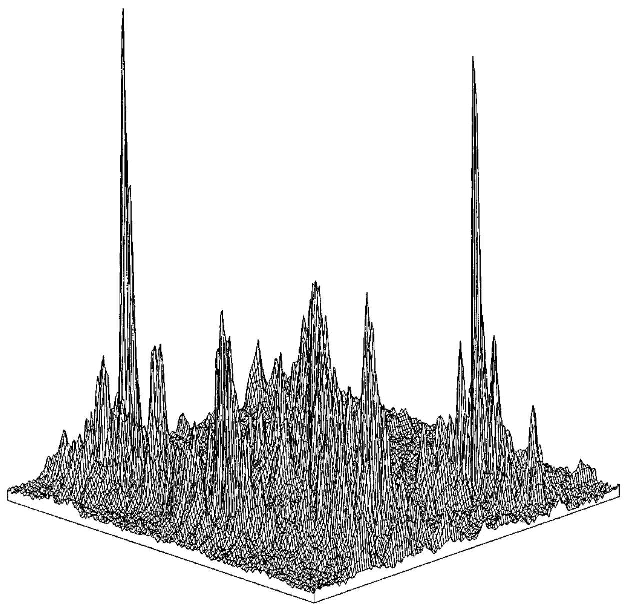
For comparison, in Fig. 5(b) we show the correlation obtained with a LH-component filter, expanded around the center (128, 128) with expansion order $N = 1$ and $L = 2$. In this case the peak intensities for both STOP versions differ in ~50%, in agreement with a projection factor of $\alpha = 2$. Other objects produce higher peaks than that of the scaled true target, impeding the detection.

In Fig. 6 a plot of the intensity correlation peak versus the scale factor for both types of filters was made. For the phase-only LH-derived filter the intensity correlation peaks for the STOP signal never decreases to <90% for projection factors of as much as $\alpha = 4$ [Fig. 6(a)]. In comparison, for LH decomposition [Fig. 6(b)] the intensity values decrease quickly with α, as predicted by Eq. (3) (the correlation peak decreases to <80% of the nontilted correlation value for $\alpha \approx 1.5$). For values of α more than 2.5 the results give a false appearance caused by resampling problems.

In order to check the performance of the filter with other types of objects, we used an input scene that contained four airplane contours, two of them being tilted versions of the other with parameter $\alpha = 2$ (Fig. 7). The targets are the two airplanes in the upper part of the image. With the filter matched to the untilted airplane with $q = 2$, the correlation intensity peaks for both targets differ ~26%, and a very low threshold of 7% is enough to reject the peaks corresponding to the other airplanes in the scene.

4. Optical Results

The above results for the objects depicted in Fig. 2 were experimentally tested by optical experiments.



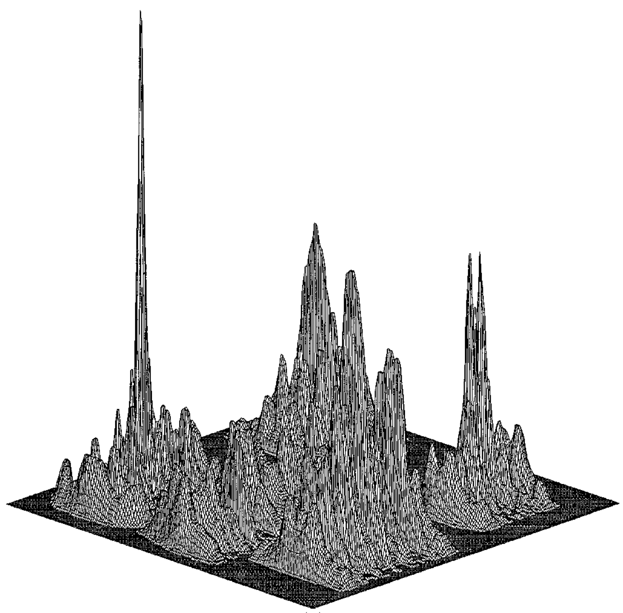


Fig. 5. Computer simulation of the correlation-intensity plane for the input image in Fig. 2 with (a) the phase-only LH-derived filter for the target image in Fig. 2 with $q = 1$ and (b) the LH filter for the same target with $L = 2$ and $N = 1$ as the expansion order and (128, 128) as the expansion center.

The setup consists of a simple correlator with a variable scale. The object is input into the system by means of a photographic transparency. The filter is implemented through a computer-generated hologram by the well-known detour phase method.[12] The computer-generated hologram is written by a 300-dpi laser printer and photoreduced to ~10 mm. No loss occurs in the amplitude information because the filter is a phase-only filter. The output is grabbed with a monochrome CCD camera and stored in a computer.

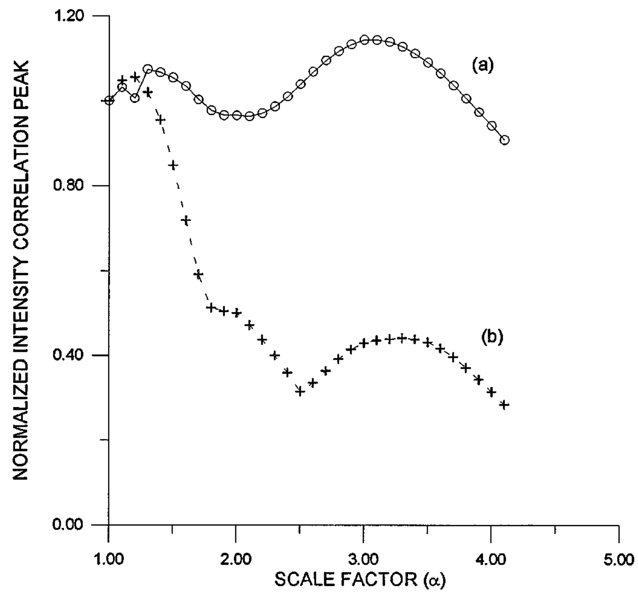


Fig. 6. Normalized intensity-correlation-peak value versus scale factor for the filter matched to the STOP signal (a) with the phase-only LH-derived filter for $q = 1$ and (b) with the LH filter with $L = 2$ and $N = 1$ as the expansion order and (128, 128) as the expansion center.

Figure 8 shows the correlation obtained for the image in Fig. 2 when the phase-only LH-derived filter is used. The experimental results are analogous to those obtained in numerical simulations [Fig. 5(a)]. The correlation peaks corresponding to the true targets present a similar intensity, and a relatively low threshold (52%) is sufficient to reject the other objects. The overall noise is not significantly increased with respect to that obtained in computer simulations.

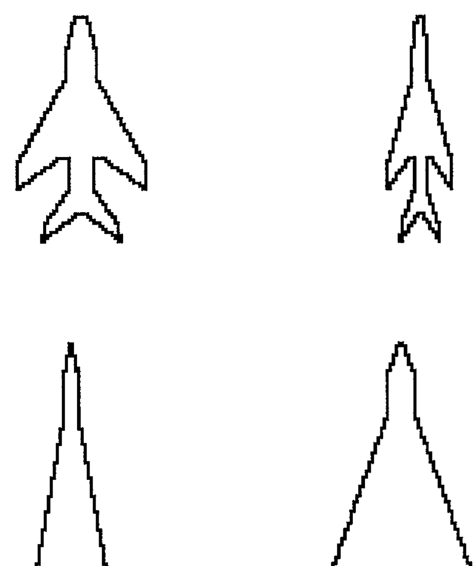


Fig. 7. Input image containing two different types of airplane contours. The reference is located at the upper-left corner of the image. Every signal appears twice in the scene with two projection factors ($\alpha = 1$ and $\alpha = 2$).

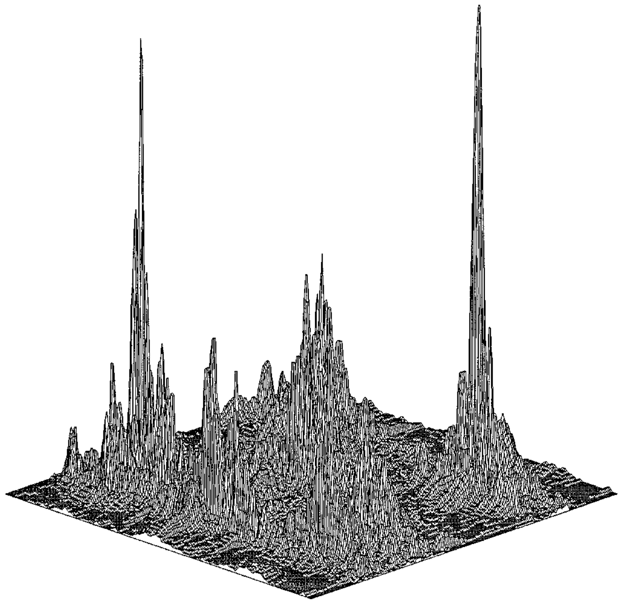


Fig. 8. Correlation intensity distribution obtained in optical experiments. The input image is shown in Fig. 2, and the filter is the phase-only LH-derived filter matched to the target object in Fig. 2 for LH frequency $q = 1$.

5. Conclusions

A new phase-only LH-derived filter has been proposed for projection-invariant pattern recognition. Computer simulations and optical experiments have been performed to verify its theoretical behavior.

First, a procedure to determine the optimal LH frequency to make the filter was outlined. In this way, the harmonic-frequency value selected is optimum for detection of any tilted version of the target in the input plane.

In agreement with the theory, the intensities of the correlation peaks of the true targets change slowly with the 1D scale factor. Conversely, when LH filters are used, the output peaks reduce inversely with the projection scale factor. As a consequence, an important improvement in the discrimination capability of the filter has been obtained because there are only small differences between detecting any tilted object and detecting the untilted one.

This work was carried out with the financial support of Comisíon Interministerial de Ciencia y Technología project TAP93-0667-C03-03. The stay of D. Mendlovic at the Universitat de València was supported by the Israeli-Spanish cultural agreement.

References

1. A. VanderLugt, “Signal detection by complex spatial filtering,” IEEE Trans. Inf. Theory **IT-10,** 139–145 (1964).
2. Y. N. Hsu and H. H. Arsenault, “Optical pattern recognition using circular harmonic expansion,” Appl. Opt. **21,** 4016–4019 (1982).
3. D. Mendlovic, E. Marom, and N. Konforti, “Shift and scale invariant pattern recognition using Mellin radial harmonics,” Opt. Commun. **67,** 172–176 (1988).
4. D. Casasent and D. P. Psaltis, “Position, rotation and scale invariant optical correlation,” Appl. Opt. **15,** 1795–1799 (1976).
5. K. Mersereau and G. M. Morris, “Scale, rotation and shift invariant image recognition,” Appl. Opt. **25,** 2338–2342 (1986).
6. S. Chang, H. H. Arsenault, and D. Liu, “Invariant optical-pattern recognition based on a contour bank,” Appl. Opt. **33,** 3076–3085 (1994).
7. J. Rosen and J. Shamir, “Scale invariant pattern recognition with logarithmic radial harmonic filters,” Appl. Opt. **28,** 240–244 (1989).
8. D. Mendlovic, N. Konforti, and E. Marom, “Shift and projection invariant pattern recognition using logarithmic harmonics,” Appl. Opt. **29,** 4784–4789 (1990).
9. D. Mendlovic, Z. Zalevsky, J. García, and C. Ferreira, “Logarithmic harmonics proper expansion center and order for efficient projection invariant pattern recognition,” Opt. Commun. **107,** 292–299 (1994).
10. B. V. K. Vijaya Kumar, W. Shi, and C. Hendrix, “Phase-only filters with maximally sharp correlation peaks,” Opt. Lett. **15,** 807–809 (1990).
11. B. V. K. Vijaya Kumar and L. Hassebrook, “Performance measures for correlation filters,” Appl. Opt. **29,** 2997–3006 (1990).
12. A. W. Lohmann and D. P. Paris, “Binary Fraunhofer holograms generated by computer,” Appl. Opt. **6,** 1739–1748 (1967).