

Position, rotation, and scale invariant optical correlation

David Casasent and Demetri Psaltis

A new optical transformation that combines geometrical coordinate transformations with the conventional optical Fourier transform is described. The resultant transformations are invariant to both scale and rotational changes in the input object or function. Extensions of these operations to optical pattern recognition and initial experimental demonstrations are also presented.

Introduction

Considerable attention has recently been focused on extending the flexibility of the operations possible in an optical processor. This work has involved various geometrical transformations,¹ preprocessing of imagery by halftone screens,² nonlinear operations using optical feedback,³ and Mellin transforms,^{4,5} among others. The Mellin transform has been shown to be quite useful in the correlation of imagery that differs in scale, since this transform is scale invariant. In this paper, a new transformation that is invariant to both a scale and a rotational change in the input image is discussed. Experimental demonstrations of the optical implementation of this transformation are presented. Its use in optical pattern recognition is discussed, and the experimental results of an optical correlation that is scale and rotationally invariant are included.

The need for such transformations and operations is quite apparent when one considers Figs. 1 and 2. In Fig. 1, the SNR (signal-to-noise ratio) of the correlation peak is plotted vs the percent scale change between the input and holographic matched filter function. The analogous plot of SNR vs the rotation angle between the input and filter function is shown in Fig. 2. These plots were experimentally obtained on a 35-mm transparency of an aerial image with about 5–10-lines/mm resolution. The rate of decrease of SNR increases with the space bandwidth product of the image. The SNR is seen to decrease from 30 dB to 3 dB with a 2% scale change and from 30 dB to 3 dB with a 3.5° rotation. These represent severe losses and demonstrate the practical problems associated with an optical pattern recognition system.

These scale and rotation problems can be addressed by scanning the image point by point, digitizing the data, and performing the operations to be described

with a digital computer. Such schemes do not have the real-time parallel processing advantages of an optical processor and have extensive data storage requirements. Therefore, only coherent optical implementations of solutions to these problems will be considered both to enhance the practicality of these processors and because of their higher potential processing capacity.

Transformations

Several methods of overcoming these scale and rotational problems have been considered. One approach involves the storage of many multiplexed holographic spatial filters of the object at various scale changes and various rotational angles. While theoretically feasible, this approach suffers from a severe loss in diffraction efficiency proportional to the square of the number of stored filters. In addition, a precise synthesis system is required to fabricate the filter bank as well as a high storage density recording media. The second approach that has been suggested involves positioning the input behind the transform lens. As the input is moved along the optical axis, the transform is scaled. While useful in the laboratory, this method is only appropriate for smaller 20% scale changes. Since this method involves mechanical movement of components, it is not compatible with the real-time capability of an optical processor. Similar mechanical rotation of the input can also be used to compensate for orientation errors. However, limitations similar to those in the mechanical scaling system are still present.

The novel method to be discussed overcomes the practical problems present in the above implementations by utilizing a transformation which is itself invariant to scale and orientational changes in the input. The first step in the synthesis of such a transformation is to form the magnitude of the Fourier transform $|F(\omega_x, \omega_y)|$ of the input function $f(x, y)$. This eliminates the effects of any shifts in the input and centers the resultant light distribution on the optical axis. A rotation of $f(x, y)$ rotates $|F(\omega_x, \omega_y)|$ by the same angle, and a scale change in $f(x, y)$ by a scales $|F(\omega_x, \omega_y)|$ by $1/a$.

The authors are with Carnegie-Mellon University, Department of Electrical Engineering, Pittsburgh, Pennsylvania 15213.

Received 8 January 1976.

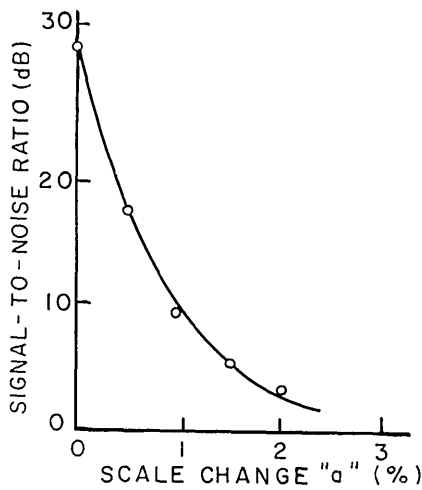


Fig. 1. SNR of the correlation peak vs the percent scale change between the input and holographic spatial filter functions.

The effects of rotation and scale changes in the resultant light distribution can be separated by performing a polar transformation on $|F(\omega_x, \omega_y)|$ from (ω_x, ω_y) coordinates to (r, θ) coordinates. Since $\theta = \tan^{-1}(\omega_y/\omega_x)$ and $r = (\omega_x^2 + \omega_y^2)^{1/2}$, a scale change in $|F|$ by a does not effect the θ coordinate and scales the r coordinate directly to $r' = ar$. A two-dimensional scaling of the input function is thus reduced to a scaling in only one dimension (the r coordinate) in this transformed $F(r, \theta)$ function.

If a one-dimensional Mellin transform^{4,5} in r is now performed on $F(r, \theta)$, a completely scale invariant transformation results. This is due to the scale invariant property of the Mellin transform.^{4,5}

The one-dimensional Mellin transform of $F(r, \theta)$ in r is given by

$$M(\omega_\rho, \theta) = \int_0^\infty F(r, \theta) r^{-j\omega_\rho - 1} dr, \quad (1)$$

where $\rho = \ln r$. The Mellin transform of the scaled function $F(ar, \theta)$ is then

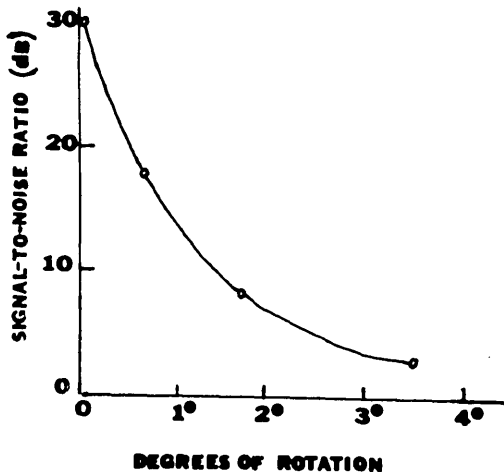


Fig. 2. SNR of the correlation peak vs the rotational angle θ_0 between the input and holographic spatial filter functions.

$$M'(\omega_\rho, \theta) = a^{-j\omega_\rho} M(\omega_\rho, \theta), \quad (2)$$

from which the magnitudes of the two transforms are seen to be identical. The optical implementation of the Mellin transform has been previously discussed,⁴ and it has been shown that

$$M(\omega_\rho, \theta) = \int_{-\infty}^\infty F(\exp \rho, \theta) \exp(-j\omega_\rho \rho) d\rho, \quad (3)$$

where $\rho = \ln r$. From Eq. (3), it can be seen that the realization of the required optical Mellin transform simply requires a logarithmic scaling of the r coordinate followed by a one-dimensional Fourier transform in r . This follows from Eq. (3) since $M(\omega_\rho, \theta)$ is the Fourier transform of $F(\exp \rho, \theta)$.

Let us now consider the effects of a rotation of the input function $f(x, y)$ by an angle θ_0 . The rotation will not affect the r coordinate in the (r, θ) plane. The effects on the θ coordinate in the (r, θ) plane are best seen with reference to Fig. 3 in which the original input $F(\omega_x, \omega_y)$ is partitioned into two sections $F_1(\omega_x, \omega_y)$ and $F_2(\omega_x, \omega_y)$ for convenience, where F_2 is the segment of F that subtends an angle θ_0 as shown in Fig. 3(a). The polar transformation of this function is shown in Fig. 3(b). The $F_1(r, \theta)$ section of $F(r, \theta)$ extends from 0 to $2\pi - \theta$ in the θ direction, while $F_2(r, \theta)$ occupies the $2\pi - \theta$ to 2π portion of the θ axis.

A version of Fig. 3(a) rotated clockwise by θ_0 is shown in Fig. 3(c). For convenience, the sections of this rotated version are denoted by $F_1'(\omega_x, \omega_y)$ and $F_2'(\omega_x, \omega_y)$. The polar coordinate transformation of this F' function is shown in Fig. 3(d). The effect of a rotation by θ_0 is seen to be an upward shift in $F_1(r, \theta)$ by θ_0 and a down-

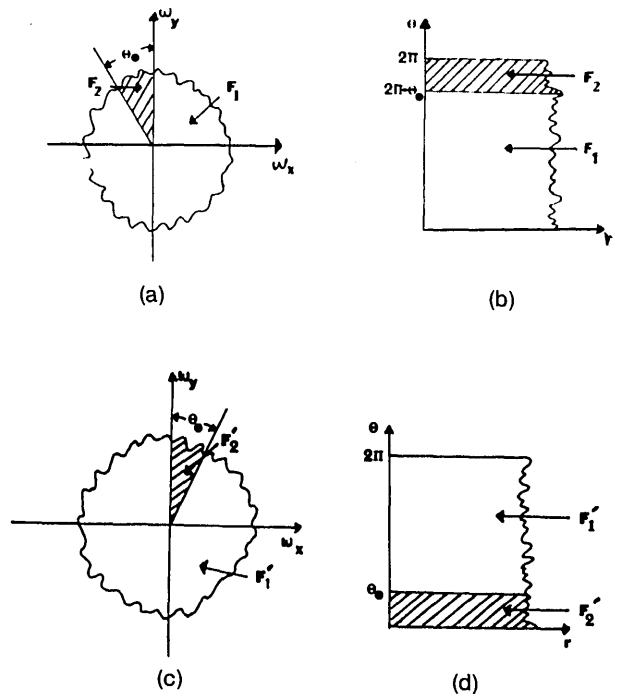


Fig. 3. Effects of a rotation by θ_0 in the input function $F(\omega_x, \omega_y)$ on the transformed function $F(r, \theta)$: (a) input function; (b) polar coordinate transform of (a); (c) rotated input function; (d) polar coordinate transform of (c).

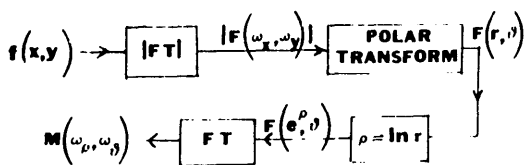


Fig. 4. Block diagram of the positional, rotational, and scale invariant PRSI transformation system. FT denotes Fourier transform.

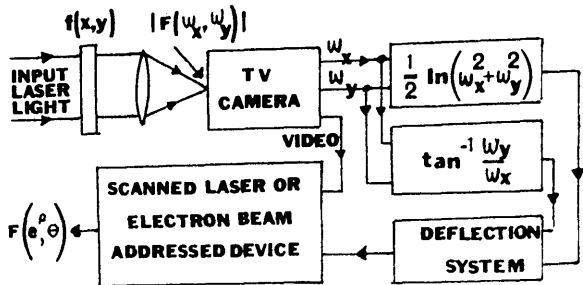


Fig. 5. Block diagram of the real-time implementation of a PRSI transformation using a real-time optically or electron beam addressed spatial light modulator.

ward shift in $F_2(r, \theta)$ by $2\pi - \theta_0$. Thus while this transformation has converted a rotation in the input to a shift in the transformed space, the shift is not the same for all parts of the function.

These shifts in $F(r, \theta)$ space due to a rotation in the input can be converted to phase factors by performing a one-dimensional Fourier transform on $F(r, \theta)$. This operation will prove useful in the eventual correlation of two scaled and rotated images. If the one-dimensional transform of $F(r, \theta)$ in the θ direction is denoted by $G(r, \omega_\theta)$, the one-dimensional transform of the original unrotated input,

$$F(r, \theta) = F_1(r, \theta) + F_2(r, \theta), \quad (4)$$

is given by

$$G(r, \omega_\theta) = G_1(r, \omega_\theta) + G_2(r, \omega_\theta), \quad (5)$$

whereas the transform of the rotated input,

$$F'(r, \theta) = F_1(r, \theta - \theta_0) + F_2(r, \theta + 2\pi - \theta_0), \quad (6)$$

is described by

$$G'(r, \omega_\theta) = G_1(r, \omega_\theta) \exp(-j\omega_\theta \theta_0) + G_2(r, \omega_\theta) \exp[j\omega_\theta (2\pi - \theta_0)]. \quad (7)$$

If these scale, rotational, and positional transformations are combined, the system block diagram shown in Fig. 4 results. The final Fourier transform (FT) is now a two-dimensional transform, in which the Fourier transform in ρ is used to effect scale invariance by the Mellin transform, and the Fourier transform in θ is used to convert the shifts due to θ_0 to phase terms as described by Eqs. (4)–(7). The resultant function is thus a Mellin transform in r and hence is denoted by M in Fig. 4. If the complete transformation of $f(x, y)$ is represented by

$$M(\omega_\rho, \omega_\theta) = M_1(\omega_\rho, \omega_\theta) + M_2(\omega_\rho, \omega_\theta), \quad (8)$$

the transformation of the function $f'(x, y)$, which is scaled by a and rotated by θ_0 , is given by

$$M^*(\omega_\rho, \omega_\theta) = M_1(\omega_\rho, \omega_\theta) \exp[-j(\omega_\rho \ln a + \omega_\theta \theta_0)] + M_2(\omega_\rho, \omega_\theta) \exp[-j[\omega_\rho \ln a - \omega_\theta (2\pi - \theta_0)]]. \quad (9)$$

Correlations

The positional, rotational, and scale invariant (PRSI) correlation is based on the form of Eqs. (8) and (9). If the product M^*M' is formed, we obtain

$$M^*M' = M^*M_1 \exp[-j(\omega_\rho \ln a + \omega_\theta \theta_0)] + M^*M_2 \exp[-j[\omega_\rho \ln a - \omega_\theta (2\pi - \theta_0)]]. \quad (10)$$

The Fourier transform of Eq. (10) then consists of two terms:

(a) the cross-correlation $F_1(\exp \rho, \theta) * F(\exp \rho, \theta)$ located at $\rho' = \ln a$ and $\theta' = \theta_0$;

(b) the cross-correlation $F_2(\exp \rho, \theta) * F(\exp \rho, \theta)$ located at $\rho' = \ln a$ and $\theta' = -2\pi + \theta_0$, where the coordinates of this output Fourier transform plane are (ρ', θ') . If the intensities of these two cross-correlation peaks are summed, the result is the autocorrelation of $F(\exp \rho, \theta)$. Therefore, the cross correlation of two functions that are scaled and rotated versions of one another can be obtained. Most important, the amplitude of this cross correlation will be equal to the amplitude of the autocorrelation function itself.

From the locations of the cross-correlation peaks, the scale factor a and the rotation θ_0 between the two functions can be obtained. By forming the magnitude of the Fourier transform in the initial step in Fig. 4, positional information was lost. To retrieve these data, the input can be scanned⁶ or the scale and rotational data from the PRSI correlation can be used to perform conventional correlation.

Experiments

The real-time implementation of the PRSI transformation can be realized as shown in block diagram form in Fig. 5. The Fourier transform $F(\omega_x, \omega_y)$ of the input image $f(x, y)$ is formed on a TV camera. The geometrical transformations are implemented by electronic processing of the deflection ramps, which are then used to deflect a scanning laser beam or a scanning electron beam. These scanned beams can be used to record the resultant $F(\exp \rho, \theta)$ distribution on a variety

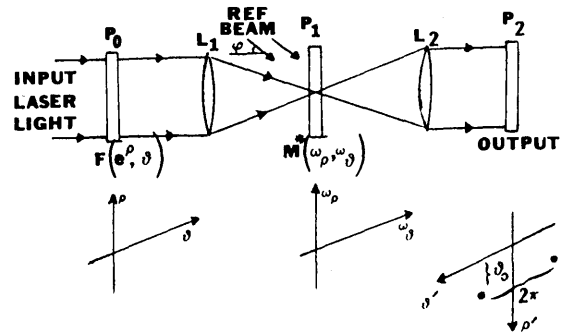


Fig. 6. PRSI correlator schematic diagram.

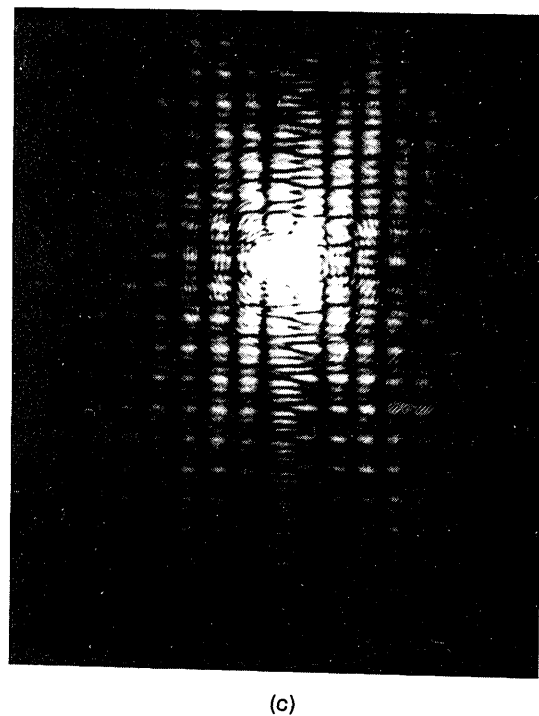
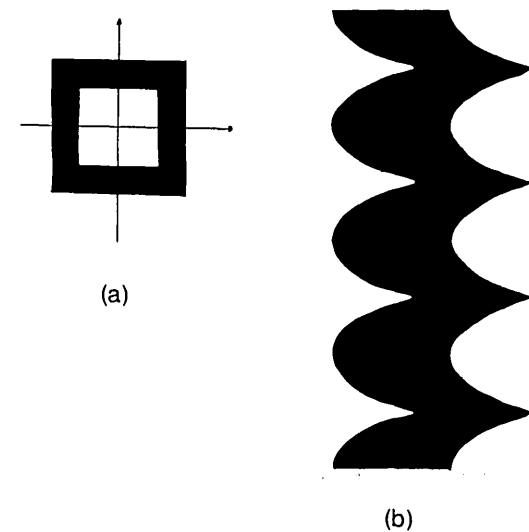


Fig. 7. (a) Input function $F(\omega_x, \omega_y)$; (b) polar transform of (a) with logarithmic scaling in r to form $F(\exp\rho, \theta)$; (c) Fourier transform of (b) or $M(\omega_\rho, \omega_\theta)$.

of real-time reusable spatial light modulators.⁷⁻⁹ Other promising methods using computer generated geometric transformation filters can be utilized. This will be the subject of a future paper. The optical implementation of the Mellin transform in r has been previously discussed.⁴ The Fourier transform in θ is the conventional optical operation. Thus all operations denoted in Fig. 4 can be realized on an optical processor, and $F(\exp\rho, \theta)$ and $F'(\exp\rho, \theta)$ can be produced.

The actual correlator section of this PRSI correlator is shown in Fig. 6. It is analogous to the conventional frequency plane correlator¹⁰ and will thus not be ex-

tensively discussed. The function $M^*(\omega_\rho, \omega_\theta)$ is produced by placing $F(\exp\rho, \theta)$ in the input plane P_0 of Fig. 6 and interfering its transform $M(\omega_\rho, \omega_\theta)$ (formed by L_1) and a plane wave reference beam (at an angle θ) in plane P_1 . By analogy with conventional holographic spatial filter synthesis, one of the four terms recorded in P_1 will be proportional to $M^*(\omega_\rho, \omega_\theta)$ as required. If the reference beam is now blocked and $F'(\exp\rho, \theta)$ is placed in P_0 , the light distribution incident on P_1 will be $M'(\omega_\rho, \omega_\theta)$. One term in the light distribution leaving P_1 is given by $M'M^*$ as described by Eq. (10). The Fourier transform of $M'M^*$ is then formed in P_2 and contains two cross-correlation peaks as shown in Fig. 6 and proven earlier.

To provide initial experimental confirmation of these concepts, a simple square [Fig. 7(a)] was used as the input function $F(\omega_x, \omega_y)$. Its polar transform with logarithmic scaling of r was digitally calculated. The resultant pattern $F(\exp\rho, \theta)$ is shown in Fig. 7(b). A transparency of this pattern was prepared, and its Fourier transform was generated optically. The resultant pattern $M(\omega_\rho, \omega_\theta)$ is shown in Fig. 7(c). A scaled and rotated version of this square with a 45° rotation angle and a 200% scale change was then produced. This $F'(\omega_x, \omega_y)$ function is shown in Fig. 8(a). Its polar transform $F'(\exp\rho, \theta)$ logarithmically scaled in r is shown

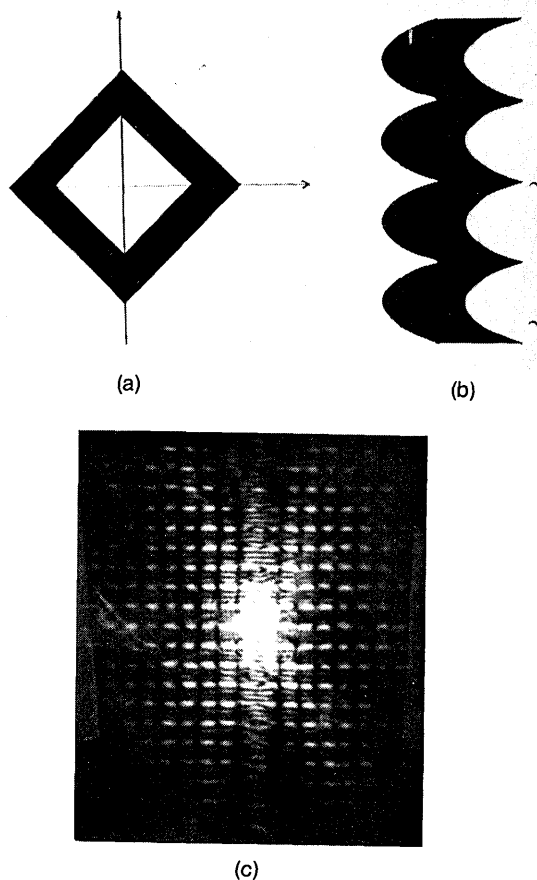


Fig. 8. (a) Rotated and scaled input function $F'(\omega_x, \omega_y)$; (b) polar transform of (a) with logarithmic scaling in r to form $F'(\exp\rho, \theta)$; (c) Fourier transform of (b) or $M'(\omega_\rho, \omega_\theta)$.

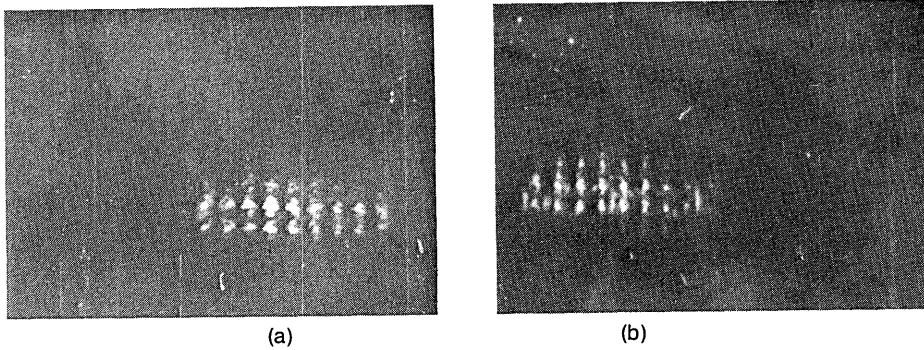


Fig. 9. (a) Autocorrelation of $F(\omega_x, \omega_y)$; (b) cross correlation of $F(\omega_x, \omega_y)$ and $F'(\omega_x, \omega_y)$.

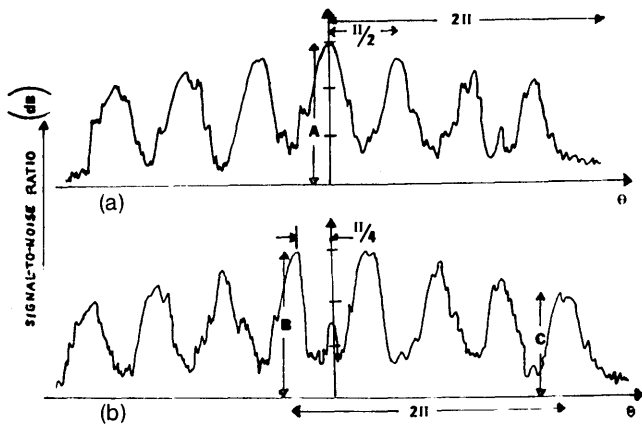


Fig. 10. (a) Cross section of autocorrelation peak in Fig. 9(a); (b) cross section of cross-correlation peaks in Fig. 9(b).

in Fig. 8(b), and its optically produced Mellin transform $M'(\omega_x, \omega_y)$ is shown in Fig. 8(c). The pattern of peaks in Figs. 7(c) and 8(c) is due to the repetitive input pattern used and the limited aperture containing the input.

The correlator shown in Fig. 6 was used to form the Mellin transform hologram of $F(r, \theta)$ at P_1 . With $F(\exp \rho, \theta)$ placed at P_0 and the Mellin hologram at P_1 , the optical pattern in the output plane P_2 is the autocorrelation function and contains a single bright peak as shown in Fig. 9(a). With $F'(\exp \rho, \theta)$ placed at P_0 , the cross-correlation pattern shown in Fig. 9(b) appears at P_2 . As predicted, it contains two cross-correlation peaks of light. The cross sections of these auto correlations and cross-correlations of $F(\exp \rho, \theta)$ and $F'(\exp \rho, \theta)$ are shown in Fig. 10. These were obtained by scanning the light distribution of Figs. 9(a) and 9(b) with a photometric microscope. The periodic structure in the cross-sectional scans in Fig. 10 and in Fig. 9 is due to the repetitive pattern in the input.

The central autocorrelation peak in Fig. 10(a) is shifted in Fig. 10(b) by the $\theta_0 = 90^\circ$ rotation between the two inputs. The sum of the amplitudes of the cross-correlation peak B in Fig. 10(b) (29 dB) and the peak C (shifted by 2π) (21 dB) is found to equal 920 which

equals the amplitude of the autocorrelation peak A (30 dB = 1000) in Fig. 10(a) within 8%. This is well within the accuracy of these experiments. The displacement of the correlation peak from the ρ' axis is found to be proportional to the scale difference $a = 200\%$ between the two inputs.

Thus, input objects that differ in scale as well as rotation can be correlated optically. In addition, the scale change and amount of rotation can be determined from the coordinates of the correlation peak. At present, the sum of the outputs of two detectors spaced by 2π is required for no loss in SNR in the cross correlations.

Summary/Conclusion

A new optical transformation that makes possible the optical correlation of two objects that differ in both scale and rotational alignment has been discussed and initial experimental demonstrations performed. This technique, and similar approaches in which the operations performed by an optical computer are extended, should be of major practical importance in optical pattern recognition and optical data processing.

This research was supported by the Office of Naval Research on contract NR-350-011 and in part by the Air Force Office of Scientific Research, Air Force Systems Command USAF, under grant AFOSR-75-2851. Initial technical discussions with George Huang of Philco-Ford are gratefully recognized.

References

- O. Bryngdahl, *J. Opt. Soc. Am.* **64**, 1092 (1974).
- A. Sawchuck, in *Proc. Elec. Opt. Sys. Des. Conf.* (Anaheim, Calif. 1974).
- S. Lee, *Opt. Eng.* **13**, 196 (1974).
- D. Casasent and D. Psaltis, *Opt. Eng.* (1976) to appear.
- D. Casasent and D. Psaltis, *Opt. Commun.* to appear.
- W. R. Callen and J. E. Weaver, in *Proc. Elec. Opt. Sys. Des. Conf.* (Anaheim, Calif., 1975).
- D. Casasent, *Proc. IEEE* (Nov. 1976).
- D. Casasent, "Materials and Devices for Optical Computing," in *Optical Information Processing*, G. W. Stroke *et al.* Eds. (Plenum, New York, 1976).
- D. Casasent, "Recyclable Input Devices and Spatial Filter Materials," in *Laser Applications*, M. Ross, Ed. (Academic, New York, 1976).
- A. Vander Lugt, *IEEE Trans. Inf. Theory* **IT-10**, 139 (1964).