Three-dimensional optical Fourier transform and correlation

Joseph Rosen

Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel

Received March 10, 1997

Optical implementation of a three-dimensional (3-D) Fourier transform is proposed and demonstrated. A spatial 3-D object, as seen from the paraxial zone, is transformed to the 3-D spatial frequency space. Based on the new procedure, a 3-D joint transform correlator is described that is capable of recognizing targets in the 3-D space. © 1997 Optical Society of America

An optical Fourier transform (FT) obtained by a spherical lens is an efficient tool for optical image processing.[1] The lens transforms a two-dimensional (2-D) image from some transverse plane to the spatial-frequency plane. Therefore, in optical image processing, people usually deal with 2-D objects or at most 2-D projections of three-dimensional (3-D) objects. However, there are applications for which 3-D objects should be processed in the entire space and their 2-D projections do not contain enough information. Because of the essential importance of the FT in image processing of any dimension, a method of implementing a 3-D FT optically is proposed. In this scheme, 3-D objects are scanned from the paraxial point of view and Fourier transformed to the 3-D spatial frequency space. Using a 3-D FT, one can do spatial filtering and a correlation on the entire 3-D input distribution. As an example, a spatial correlation for pattern recognition in 3-D space is demonstrated.

The proposed scheme is shown in Fig. 1. A 3-D input function $o(x_s, y_s, z_s)$ is located in the coordinate system (x_s, y_s, z_s). P_1 is the transverse plane $z_s = 0$. A CCD observes plane P_1 from a distance L. It is assumed that the field of view on P_1 is wider than the transverse dimension of the input function and that the depth of focus of the camera is longer than the longitudinal dimension of the input function. All the observed objects are imaged onto the CCD and displayed on a spatial light modulator (SLM). This SLM is located in the front focal plane of a spherical lens with a focal length f and is illuminated by a plane wave. Thus a 2-D FT of the SLM's image is obtained on the lens's back focal plane P_3, and so far this system is not different from many other well-known hybrid configurations.[2] The new element here is the third spatial frequency variable obtained by displaying a few projections of the plane P_1, each with a different transverse displacement of the CCD.

To express the relation between the 3-D input function and the distributions on plane P_3, let us first look at a single point (x_s', y_s', z_s') from the entire input object. The observed point is displayed on the SLM at point (x_i, y_i) and transformed by the lens into a linear phase function $\exp[i2\pi(x_i u + y_i v)/\lambda f]$, where u and v are the coordinates of back focal plane P_3 and λ is the wavelength of the plane wave. Assuming that plane P_1 is displayed on the SLM with a magnification factor M, we can calculate the location of the observed point on the SLM as a function of its location in object space and the amount of CCD displacement (D_x, D_y). Using simple geometrical considerations, we can locate the point on the SLM:

$$x_i = M(D_x + x_s')/(1 - z_s'/L),$$
$$y_i = M(D_y + y_s')/(1 - z_s'/L). \tag{1}$$

We obtain the overall field distribution on the rear focal plane that results from a 3-D object, $o(x_s, y_s, z_s)$, for a given displacement (D_x, D_y) by integrating over the linear phases contributed by all the object points, as follows:

$$O(u, v, D_x, D_y) = \int\int\int o(x_s, y_s, z_s)\exp[i2\pi(x_i u + y_i v)/\lambda f]\mathrm{d}x_s\mathrm{d}y_s\mathrm{d}z_s. \tag{2}$$

At this point we assume that $L \gg z_{s,\max}$, $Lf\lambda \gg Mu_{\max}z_{s,\max}x_{s,\max}$, and $Lf\lambda \gg Mv_{\max}z_{s,\max}y_{s,\max}$, where $(y_{s,\max}, x_{s,\max}, z_{s,\max})$ are the maximum values of the volume that contains the input function and $(u_{\max}, v_{\max})$ are the maximum values of plane P_3. For

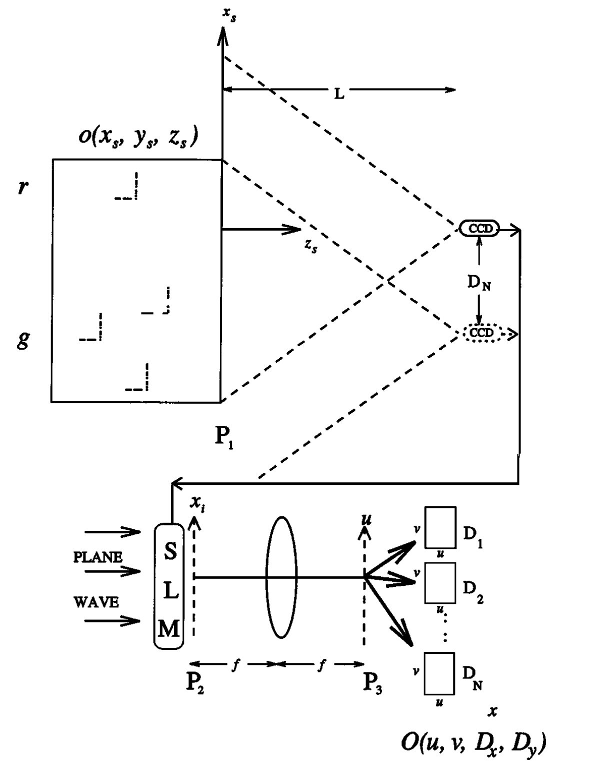

Fig. 1. Schematic of the optical system for the 3-D FT.

0146-9592/97/130964-03$10.00/0 © 1997 Optical Society of America

a given input the validity of these assumptions depends on the system parameters, and therefore we can always design the system to satisfy these assumptions. Following the assumptions, we approximate the location of each object's point on the SLM as

$$x_i \cong M(D_x + x_s + z_s D_x/L)\,,$$
$$y_i \cong M(D_y + y_s + z_s D_y/L)\,. \tag{3}$$

Substituting approximations (3) into Eq. (2) yields

$$O(u, v, D_x, D_y) = \exp[(i2\pi M/\lambda f)(D_x u + D_y v)] \times \int\int\int o(x_s, y_s, z_s)\exp\{(i2\pi M/\lambda f)[x_s u + y_s v + (z_s/L)(D_x u + D_y v)]\}\mathrm{d}x_s\mathrm{d}y_s\mathrm{d}z_s\,. \tag{4}$$

Equation (4) is a 3-D FT, multiplied by a linear phase function, which transforms an object function $o(x_s, y_s, z_s)$ into a 3-D spatial frequency function $O(\omega_x, \omega_y, \omega_z)$, where $\omega_x = Mu/\lambda f$, $\omega_y = Mv/\lambda f$, and $\omega_z = M(D_x u + D_y v)/\lambda f L$.

Following a conventional Fourier analysis, we know that the maximum camera displacement depends on the longitudinal size δz_s of the smallest input element. Assuming that $D_{x,\max} = D_{y,\max} = D$ and $u_{\max} = v_{\max} = B$, the condition $DB \geq \lambda f L/\sqrt{2} M \delta z_s$ should be satisfied in order not to lose the information on the smallest longitudinal element. We conclude that the bandwidth of the system in the third dimension ω_z depends directly on both the transverse bandwidth and the maximal CCD displacement. Another limitation on the camera's displacement is given by the condition $D \leq W - \Delta x_s$, where W is the width of the field of view on P_1 and Δx_s denotes the transverse size of the input object. Shifting the CCD beyond this limitation causes the object to disappear from the field of view.

Although it is convenient to analyze the system in terms of continuous signals, our detected 3-D signal is discrete in its all dimensions. This is so because each 2-D image is recorded separately as a collection of discrete pixels inside the computer. Therefore the limitations on the sampling interval along the CCD's translation should be considered. Let us assume that the maximal sampling intervals $(d_{x,\max}, d_{y,\max})$ satisfy the equation $d_{x,\max} = d_{y,\max} = d$, the maximal transverse sampling intervals $(\delta u_{\max}, \delta v_{\max})$ satisfy the equation $\delta u_{\max} = \delta v_{\max} = b$, and Δz_s denotes the longitudinal size of the input object. The criterion $\mathrm{d}b \leq \lambda f L/2\sqrt{2} M \Delta z_s$ should be satisfied for a signal to be reconstructed completely, along the z_s direction, from its samples in the spectral domain.

A 3-D FT can be useful for spatial filtering and spatial spectroscopy in 3-D space. For spatial filtering one needs first to transform the coordinates of $O(u, v, D_x, D_y)$ from the physical detection space (u, v, D_x, D_y) to the spectral space $(\omega_x, \omega_y, \omega_z)$. Then the transformed function $\tilde{O}(\omega_x, \omega_y, \omega_z)$ is multiplied by some filter function. Finally, the output result is obtained from the last product by an inverse 3-D FT. When the phase distribution of the object's 3-D FT is essential, all these stages should be executed optically, without any photodetection until the final stage of detecting the output signals. Such an optical system could be implemented by a few levels of holograms for coordinate transforms and interconnections. However, this 3-D spatial filtering would become complicated, sensitive to noise, and impractical. To overcome this difficulty without losing the phase information, we adopt the concept of the joint transform correlator[3] (JTC) into our new 3-D correlator. Although the intensity of the spatial spectrum is recorded by a photodetector, the JTC yields a real correlation between two arbitrary functions without losing the phase information.

The 3-D input space of the JTC contains a reference object $r(x_s, y_s, z_s)$ at some point, say, the origin, and a few tested objects, denoted together by the function $g(x_s, y_s, z_s)$ and located around some other point, say, point (a, b, c). Therefore the JTC input function is given by

$$o(x_s, y_s, z_s) = r(x_s, y_s, z_s) + g(x_s + a, y_s + b, z_s + c)\,. \tag{5}$$

Substituting Eq. (5) into Eq. (4) and computing the squared magnitude of the complex amplitude $O(u, v, D_x, D_y)$ yield the intensity distribution on plane P_3:

$$\begin{aligned} I_3(u, v, D_x, D_y) = {} & |R(u, v, D_x, D_y) + G(u, v, D_x, D_y) \\ & \times \exp\{(i2\pi M/\lambda f)[au + bv + (c/L)(D_x u + D_y v)]\}|^2 \\ = {} & |R(u, v, D_x, D_y)|^2 + |G(u, v, D_x, D_y)|^2 + G(u, v, D_x, \\ & D_y)R^*(u, v, D_x, D_y)\exp\{(i2\pi M/\lambda f)[au + bv + (c/L) \\ & \times (D_x u + D_y v)]\} + G^*(u, v, D_x, D_y)R(u, v, D_x, D_y) \\ & \times \exp\{-(i2\pi M/\lambda f)[au + bv + (c/L)(D_x u + D_y v)]\}\,, \end{aligned} \tag{6}$$

where R and G are 3-D FT's, defined by Eq. (4), of r and g, respectively. The intensity distribution I_3 is recorded by another CCD into the computer, in which the coordinate transform of $I_3(u, v, D_x, D_y)$ into $\tilde{I}_3(u, v, D_x u + D_y v)$ is done relatively easily. To simplify the system it is also recommended that the final 3-D FT of $\tilde{I}_3(u, v, D_x u + D_y v)$ be performed numerically inside the computer, although in concept it is possible to display the data on a few SLM's and to do an optical FT on each SLM by use of a coherent illumination. In any case, after an inverse 3-D FT of $\tilde{I}_3(u, v, D_x u + D_y v)$ the output result is

$$\begin{aligned} c(x_0, y_0, z_0) = {} & \int \tilde{I}_3(\omega_x, \omega_y, \omega_z)\exp[-i2\pi(x_0\omega_x + y_0\omega_y + z_0\omega_z)]\mathrm{d}\omega_x\mathrm{d}\omega_y\mathrm{d}\omega_z \\ = {} & r \star r + g \star g + (r \star g) * \delta(x_0 - a, y_0 - b, z_0 - c) \\ & + (g \star r) * \delta(x_0 + a, y_0 + b, z_0 + c)\,, \end{aligned} \tag{7}$$

where $\star$ and $*$ stand for the correlation and the convolution, respectively. Similarly to an ordinary 2-D JTC, the two last terms of Eq. (7) are the cross correlations between the reference and the tested objects. The third and the fourth terms are centered around

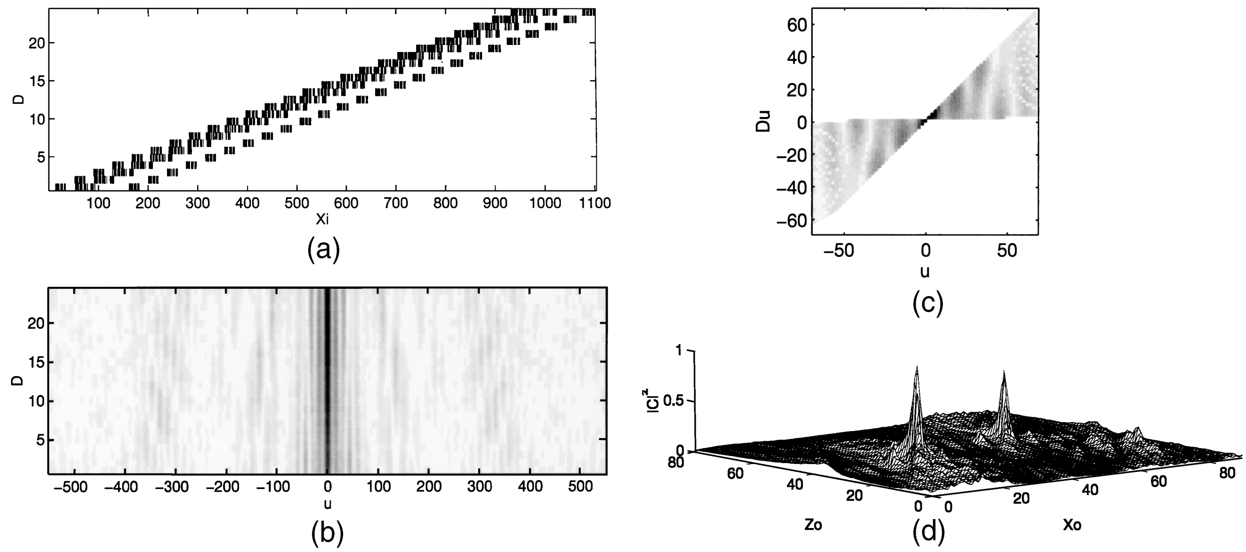

Fig. 2. Computer simulation of the system shown in Fig. 1. (a) Projections of the input image observed from various cameras' translations D. (b) Intensity distribution on plane P_3. (c) The same intensity distribution after the coordinates have been transformed. (d) Cross-correlation results between the reference and the tested objects obtained by a 2-D FT of the pattern in (c).

points (a, b, c) and $(-a, -b, -c)$, respectively. The first two terms of the autocorrelation in Eq. (7) are centered around the origin. Therefore, if one of the distances (a, b, c) is longer than the size of the tested function g, the cross correlation is spatially separated from the autocorrelation terms and becomes detectable.

When there is no information along one transverse coordinate, say y_s, this coordinate can be ignored. Thus one can perform an optical FT of the 2-D function $o(x_s, z_s)$ in the same way. This means that the function $o(x_s, z_s)$ is transformed into the coordinate system (u, D) composed of multiple u axes, each of which contains a one-dimensional FT (done by a cylindrical lens) of the image on P_1, as seen from a different CCD translation D. Then the intensity $I_3(u, D)$ is recorded in the computer and transformed there into the spectral coordinates (u, uD). One performs the final 2-D FT, in an equivalent complication, numerically or optically by illuminating a SLM containing the spectrum $\tilde{I}_3(u, uD)$ in front of a spherical lens. To simplify the presentation of our example, we choose to ignore one transverse coordinate, understanding that the extension of the system to three dimensions is straightforward.

As an example we simulated the optical system shown in Fig. 1. In this example the input plane contains four objects in the shape of corners, distributed in the (x_s, z_s) plane. Each line of each corner has the values (0,1). The uppermost object was used as a reference in the JTC scheme. The other three tested objects are distributed in the (x_s, z_s) plane far from the reference. Two of them are identical to the reference; the third is different. We can also see from Fig. 1 the location of the CCD in the initial state ($D = 0$) and in the final state ($D = L/2$). Between them the CCD sampled the observed plane 24 times. The collected data of 24 projections that should be displayed on the SLM are shown in Fig. 2(a). Each horizontal line at some point D is the one-dimensional pattern seen when the CCD is translated a distance D from the initial state. Each horizontal line was Fourier transformed by a simulated cylindrical lens. The resulting intensity distribution of the one-dimensional FT's obtained on plane P_3 is depicted in Fig. 2(b). The function $I_3(u, D)$ was transformed into $\tilde{I}_3(u, uD)$, shown in Fig. 2(c). Only two quadrants in this plot are occupied with data because the CCD in this example moved from the center only to the left. The horizontal lines $D =$ constant in Fig. 2(b) are transformed into diagonal lines that cross through the origin in Fig. 2(c). Finally, a 2-D FT of $\tilde{I}_3$ in Fig. 2(c) yielded the required cross correlation between the reference and the tested objects. Figure 2(d) is a 3-D plot of the output plane around the region of the first diffraction order [third term in Eq. (7)]. The output correlation is given in the coordinates (x_0, z_0), which are equivalent to the input coordinates (x_s, z_s). The two strong correlation peaks indicate the locations of the two corners, which are identical to the reference.

In conclusion, a method for an optical 3-D FT and correlation has been developed. This FT, in coherent optics, is analogous to the recently discovered relation in partial-coherence optics.[4] The relation between the two-point coherence function and the 3-D intensity distribution of an incoherent source is analogous to the present relation [Eq. (4)] between the complex amplitude distributions $O(u, v, D_x u + D_y v)$ and the 3-D brightness function $o(x_s, y_s, z_s)$.

References

1. J. W. Goodman, *Introduction to Fourier Optics*, 1st ed. (McGraw-Hill, New York, 1968), Chap. 7, p. 141.
2. F. T. S. Yu and S. Jutamulia, *Optical Signal Processing, Computing, and Neural Networks*, 1st ed. (Wiley, New York, 1992), Chap. 6, p. 203.
3. C. S. Weaver and J. W. Goodman, Appl. Opt. **5,** 1248 (1966).
4. J. Rosen and A. Yariv, Opt. Lett. **21,** 1011, 1803, (1996).