Interferometric optical Fourier-transform processor for calculation of selected spatial frequencies

Pierre M. Lane and Michael Cada

A novel interferometric optical Fourier-transform processor is presented that calculates the complex-valued Fourier transform of an image at preselected points on the spatial-frequency plane. The Fourier spectrum of an arbitrary input image is interfered with that of a reference image in a common-path interferometer. Both the real and the imaginary parts of the complex-valued spectrum are determined. The source and the reference images are easily matched to guarantee good fringe visibility. At least six interferograms are postprocessed to extract the real and the imaginary parts of the Fourier spectrum at preselected points. The proposed hybrid optical–digital technique is computationally appropriate when the number of desired spatial frequencies is small compared with the number of pixels in the image. When the number of desired points is comparable with the number of image pixels, a conventional or pruned two-dimensional fast Fourier transform is more appropriate. The number of digital operations required by the hybrid optical–digital Fourier processor is proportional to the number of desired spatial frequencies rather than the number of pixels in the image. The points may be regularly distributed over the spatial-frequency plane or concentrated in one or several irregularly shaped regions of interest. The interferometric optical Fourier processor is demonstrated in a moving-object trajectory estimation system. The system successfully estimates the trajectory of multiple objects moving over both stationary and white-noise backgrounds. A comparison of performance was made with all-digital computation. With everything else equal, our hybrid optical–digital calculation was more than 3 orders of magnitude faster.

OCIS code: 070.0070.

1. Introduction

The Fourier transform (FT) is critically important to many signal-processing systems. It is an extremely powerful analytic tool; however, its implementation is very computationally intensive. Even the fastest digital algorithms based on the fast Fourier transform,[1] (FFT) often consume too many CPU clock cycles and prevent real-time processing. It has been known for some time that optical Fourier processors calculate an FT at the speed of light. Unfortunately, only half of the complex-valued spectrum is recoverable with conventional optical detectors: The FT has both magnitude and phase components, and only the former is easily detectable with a conventional optical detector. An interferometer may be employed to detect both magnitude and phase; however, traditional devices are bulky and very sensitive to mechanical shock, vibration, and temperature fluctuations. It has been difficult to realize a stable optical processor that provides both the magnitude and the phase components required by most applications.

Single-chip video processors consisting of multiple parallel-processing digital signal processors can compute the complex-valued FFT of low-resolution images at video frame rates. The TMS320C80 digital signal processor manufactured by Texas Instruments (Dallas, Tex.) can process a complex-valued 256-point FFT in ~5000 instruction cycles. Such a device running at 50 MHz can therefore calculate the FFT of a real-valued 256×256 pixel image in approximately $256 \times 5000/50$ MHz = 25.6 ms, which is equivalent to almost 40 image frames/s.

The interferometric optical Fourier processor proposed here is based on a novel technique that we call joint-transform interference (JTI). It is designed to calculate the complex-valued FT of an arbitrarily large image at a selected number of points on the

When this research was done, both authors were with the Department of Electrical and Computer Engineering, Dalhousie University, P.O. Box 1000, Halifax, Nova Scotia B3H 1H1. P. Lane (E-mail: plane@bccancer.bc.ca) is now with the Department of Cancer Imaging, BC Cancer Research Center, 601 West Tenth Avenue, Vancouver, British Columbia V5Z 1L3.

Received 27 March 2000; revised manuscript received 22 August 2000.

0003-6935/00/356573-14$15.00/0

spatial-frequency plane. When the number of desired points is comparable with the number of pixels in the image, a digital two-dimensional (2-D) FFT is appropriate; however, when a relatively small number of frequency-domain points is desired, the optical technique presented here can offer substantial computational savings.

The number of operations required for the direct inner-product computation of a single frequency-domain point of an $N \times N$ pixel image is $O(N^2)$. The number of operations required for the full 2-D FFT computation is $O(N^2 \log_2 N)$, and the computational complexity of a pruned 2-D FFT algorithm[2] falls between these two limits, depending on the number and distribution of the desired points.

The number of operations required by the optical Fourier processor proposed in this contribution is proportional to the number of desired points rather than the square of the image size. If only a small number of points are required, the optical Fourier processor can be much more efficient than a pruned 2-D FFT, especially when the input images are large. The points may be regularly distributed over the spatial-frequency plane or concentrated in one or several irregularly shaped regions of interest. Such distributions of points can often occur in applications such as discrete image analysis and interpolation[3] and moving object detection and trajectory estimation.[4,5]

This paper is organized as follows. An introduction to coherent detection and the development of interferometric optical FT processors is presented in Section 2. The method of JTI for coherent optical detection in an optical FT processor is described in Section 3. The novel optical–digital technique is applied to a moving object trajectory estimation in Section 4. A summary and concluding remarks are made in Section 5.

2. Interferometric Optical Fourier-transform Processors

Optical spectrum analysis was one of the first recognized applications of optical signal processing. Optical spectrum analyzers are generally divided into two classes on the basis of the variable of integration.[6] The space-integrating class performs a FT with respect to one or more spatial variables, whereas the time-integrating class performs a FT with respect to a single temporal variable. Generally, the integration in the former is provided by a lens that collects a spatial distribution of light. In the latter the integration is usually provided by a photodetector that collects a temporal distribution of light. Hybrid space and time-integrating architectures are also possible. A spectrum analyzer provides the magnitude of a signal of interest as a function of frequency; however, a Fourier analyzer (or heterodyne spectrum analyzer) is required to provide both the amplitude and the phase information.

In 1947 Dennis Gabor recorded on a photographic plate the interference pattern, or interferogram, produced by an object beam and a quasi-monochromatic reference beam. This hologram was probably the first application of coherent optical detection. In 1967 King *et al.*[7] reported one of the first heterodyning time-integrating spectrum analyzers for measuring both amplitude and phase. In their system a reference beam was interfered with a modulated object beam to produce a temporal fringe structure. Two years later, Carleton *et al.*[8] proposed one of the first collinear or common-path heterodyning architectures in which the reference beam was derived from the object beam to make the processor less sensitive to mechanical shock. In a time-integrating interferometric spectrum analyzer proposed by VanderLugt,[9] two acousto-optic modulators and a spatially modulated reference beam were employed to reduce the frequency of the temporal fringe pattern. Turpin[6] reported one of the first cases of coherent detection in a space-integrating spectrum analyzer. In his spectrum analyzer a spatial fringe pattern was detected in the focal plane of a FT lens by use of a high-resolution photodetector array.

Space-integrating 2-D interferometric Fourier processors have not generally been pursued owing to the lack of fast 2-D spatial light modulators (SLM's) with high resolution and high contrast. Those interested can find in an article by Aleksoff and Subotic[10] a review of various architectures for interferometric Fourier-transform processors that are space and time integrating and suitable for processing complex-valued synthetic aperture radar data. Time-integrating one-dimensional (1-D) interferometric Fourier processors are generally based on the more mature technology offered by acousto-optic SLM's. These time-integrating heterodyne spectrum analyzers have been discussed in detail by VanderLugt.[11] The increased dynamic range was the primary motivation behind the research of VanderLugt (since the interference term is proportional to signal amplitude rather than its intensity), and complex-valued optical Fourier processing motivated the research of Aleksoff and Subotic. The latter is the motivation here. In this paper we deal with the space-integrating type of interferometric FT processor for the purpose of computation.

3. Joint-Transform Interference

The JTI technique proposed here is a multipoint extension of the point-diffraction method of coherent detection.[12–15] It is based on interference between the Fourier spectra of two images—the FT of a reference image and the FT of a source image. This is in contrast to the point-diffraction method in which the FT of a single pixel is interfered with the FT of a source image. The difference is that in JTI the reference is an image rather than a single pixel. The amplitude of the reference image can be increased to match its power with that of the source. In addition, the spatial extent of the reference image can be increased such that the source and the reference powers are matched. When the source and the reference image powers are matched, the fringe visibility of the interferogram is optimized, the most effective use is

made of the limited dynamic range of the CCD detector, and the signal-to-noise ratio is maximized.

A. Phase-Shifting Interferometry

The real (or imaginary) part of an optical wave front can be determined by interfering it with a known reference; however, a phase-shifting technique is necessary to determine both at the same time. Phase-shifting interferometry (PSI) is often employed to characterize the quality of an optical element (a lens or mirror, for example) by measurement of the phase distortion added by the element to an undistorted source wave front.[16] A time-varying phase shift is generally introduced between the source and the reference wave fronts in an interferometer to produce a time-varying interferogram. The interferogram time sequence is then postprocessed on a point-by-point basis to determine the phase and magnitude of the distorted wave front. The time-varying phase shift is usually introduced by the mechanical translation of one of the mirrors in a two-path interferometer by use of a piezoelectric transducer. The PSI technique has also been demonstrated in a common-path point-diffraction interferometer by use of a liquid-crystal display to introduce the phase shift.[17–19]

The JTI method introduced here employs a phase-stepping technique to determine the real and the imaginary parts of a source wave front. The phase shift is introduced by translation of the source and the reference images in the front focal plane of a FT lens. The mutual source–reference displacement in the spatial domain corresponds to a phase shift in the frequency domain. In conventional PSI, arbitrary source wave fronts are interfered with phase-shifted reference wave fronts and the phase shift is introduced directly. With the JTI method the FT of an arbitrary source image is interfered with the phase-shifted FT of a known reference image, and the source–reference phase shift is introduced by mutual displacement of the source and the reference images. A quadrature reference is necessary to extract the complex amplitude of the source image FT. The real part is extracted by interfering the source image FT with the in-phase component, and the imaginary part is extracted from the interferogram generated with the quadrature component.

B. Spatial Modulation

The interferograms generated by the JTI technique are indeed very similar to the Gabor holograms, specifically FT Gabor holograms. The holograms first recorded by Gabor were in-line holograms as opposed to the more practical sideband holograms recorded later by Leith and Upatneiks. Gabor recorded holograms of semitransparent objects so that the quiescent light transmitted by the object could be employed as a common-path reference beam. Light incident on the object was diffracted by it to produce the object beam and, at the same time, transmitted through it to provide the reference beam. The real and virtual images reconstructed from a Gabor hologram have the same optical axis, and consequently the images overlap in space.

This practical disadvantage was corrected by the Leith–Upatneiks hologram, which is recorded and reconstructed with appropriate beams inclined with respect to the object beam. The inclination between the two beams causes an intensity grating or spatial carrier to be recorded on the photographic film. The spatial carrier is modulated by the amplitude and phase information encoded in the object beam. When the Leith–Upatneiks hologram is reconstructed, the spatial-carrier grating diffracts the real and the virtual images along the different axes, and the two images are spatially separable.

The interferograms produced by the JTI technique are quite similar to Gabor holograms; however, the source and the reference beams incident on the detector are typically inclined at a very small angle (the angle subtended by one or two pixels on an SLM). This source–reference inclination, introduced by design to provide a linear phase shift in the frequency domain, results in the sinusoidal modulation of the interferogram. The spatial period of the modulation is much less than that of the sinusoidal grating (spatial carrier) observed in the Leith–Upatneiks hologram. (The period of modulation introduced by a single-pixel source–reference displacement is equal to the width of the zero-order spectrum.) The effect of the modulation can be removed by demodulation of the JTI interferograms.

C. Principle of Joint-Transform Interference

The idea behind JTI is simple. The interference pattern produced by a source image and a reference image, displayed on the same SLM and mutually separated by a small distance, is detected by a square-law detector in the back focal plane of a lens. The FT of the source and the reference images produces an interferogram that contains terms proportional to the real and the imaginary parts of the source image FT. The real and the imaginary parts are isolated by addition and subtraction of interferograms generated with different source–reference mutual displacements.

Consider the joint FT of the source image $s(x, y)$ and the reference image $r(x, y)$ illustrated in Fig. 1. The source is written to the SLM at position $\mathbf{p}_s = (x_s, y_s)$, the reference is written at position $\mathbf{p}_r = (x_r, y_r)$, and the mutual displacement between the two is given by $\mathbf{p}_s - \mathbf{p}_r$. A composite source–reference image and its joint FT can be expressed as an FT pair:

$$s(\mathbf{p} - \mathbf{p}_s) + r(\mathbf{p} - \mathbf{p}_r) \Leftrightarrow \exp(i\mathbf{k} \cdot \mathbf{p}_s)S(\mathbf{k}) + \exp(i\mathbf{k} \cdot \mathbf{p}_r)R(\mathbf{k}), \quad (1)$$

where $\mathbf{p} = (x, y)$ and $\mathbf{k} = (k_x, k_y)$ represent the spatial and the spatial angular frequency dependence, respectively. The joint power spectrum detected by a

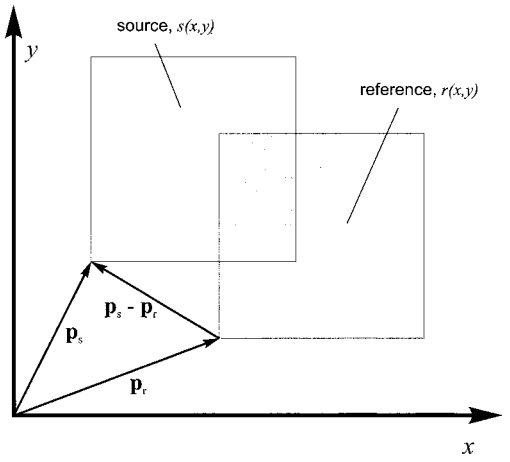


Fig. 1. Construction of a joint image from a source and a reference image: (a) component images, $t(m, n)$, $u(m, n)$, $v(m, n)$, and $w(m, n)$; (b) joint image, $f(m, n)$.

square-law detector in the back focal plane of the lens is given by

$$\begin{aligned} I_\phi(\mathbf{k}) &= I_s(\mathbf{k}) + I_r(\mathbf{k}) + 2\Re[\exp(i\phi)S(\mathbf{k})R^*(\mathbf{k})] \\ &= I_s(\mathbf{k}) + I_r(\mathbf{k}) + 2\{\cos\phi\,\Re[S(\mathbf{k})R^*(\mathbf{k})] \\ &\quad - \sin\phi\,\Im[S(\mathbf{k})R^*(\mathbf{k})]\}, \end{aligned} \tag{2}$$

where $I_s(\mathbf{k}) = |S(\mathbf{k})|^2$ and $I_r(\mathbf{k}) = |R(\mathbf{k})|^2$ represent the power (or intensity) spectra of the source and reference, respectively, $\phi = \mathbf{k}\cdot(\mathbf{p}_s - \mathbf{p}_r)$ is the linear phase difference introduced by the displacement, and operators $\Re(\cdot)$ and $\Im(\cdot)$ represent the real and the imaginary parts, respectively, of their complex-valued operands. The ϕ subscript on $I_\phi(\mathbf{k})$ denotes the dependence on mutual displacement between the source and the reference images. The joint power spectrum is clearly the sum of three distinct terms: (1) the power spectrum of the source image; (2) the power spectrum of the reference image; and (3) an interference term that contains the real and the imaginary parts of the source and the reference images. The real and the imaginary parts of the complex spectrum SR^* are cosine and sinusoid modulated, respectively. The spatial frequency of the modulation is proportional to the displacement of the source and the reference images.

It is possible to recover the real and the imaginary parts of SR^* from two interference spectra $I_\phi(\mathbf{k})$ and $I_{-\phi}(\mathbf{k})$, which have been designed to have opposite phases. The real part is extracted by addition of the two interference spectra, subtraction of the individual power spectra, and division by the cosine modulation term,

$$\Re[S(\mathbf{k})R^*(\mathbf{k})] = \frac{I_\phi(\mathbf{k}) + I_{-\phi}(\mathbf{k}) - 2I_s(\mathbf{k}) - 2I_r(\mathbf{k})}{4\cos\phi}. \tag{3}$$

Similarly, the imaginary part is extracted by subtraction of the two interference spectra and division by the sine modulation term,

$$\Im[S(\mathbf{k})R^*(\mathbf{k})] = \frac{I_\phi(\mathbf{k}) - I_{-\phi}(\mathbf{k})}{4\sin\phi}. \tag{4}$$

Clearly these real and imaginary expressions are indeterminate when their denominators are equal to zero. This apparent deficiency is overcome when Eqs. (3) and (4) are used to determine the Fourier spectra for two different mutual displacements. Both modulated spectra are weighted and then combined into a single demodulated spectrum, as shown in Subsection 3.F.

If the reference image is designed such that $R^*(\mathbf{k}) = 1$, the real and the imaginary parts of the source image are described exactly by the right-hand sides of Eqs. (3) and (4). When $R^*(\mathbf{k})$ is a function of spatial frequency, it must be moved to the right-hand side of the two equations to isolate the real and the imaginary parts of the source spectrum. If the frequency-domain reference is pure real (even reference image) or pure imaginary (odd reference image), the reference term $R^*(\mathbf{k})$ can be moved directly from the left-hand sides of Eqs. (3) and (4) to the respective denominator of the right-hand side. If the frequency-domain reference is complex, the equations are more involved.[20]

D. Joint-Transform Interference on a Binary Spatial Light Modulator

The JTI principle just described is directly applicable to a gray-scale SLM. The method must be extended, however, if a binary SLM is employed. The joint image is composed of a source image and a mutually displaced reference image. The source and the reference images can be directly added and written to a gray-scale SLM if their sum does not exceed the dynamic range of the SLM. [An SLM with $(b + 1)$ bits of quantization is required if both the source and the reference images are quantized with b bits of dynamic range.] The sum of a binary source and a binary reference will exceed the dynamic range of a binary SLM. To realize the JTI method on a binary SLM, the source and the reference images must be interleaved rather than summed. The disadvantage of this simple solution is the reduction in the effective space–bandwidth product of the SLM.

On a binary SLM with a square pixel structure, the source and the reference images may be interleaved horizontally, vertically, or diagonally. Consider, for example, a horizontal row-interleaved (or interlaced) configuration analogous to the even and the odd fields of a video frame of the national television systems committee. When the source and the reference images are displayed as the even and the odd fields, respectively, the source resides in the even lines of the joint image and the reference resides in the odd lines. This composite frame could be referred to as an even-interlaced joint image. Similarly, when the source and the reference images are displayed as the

odd and the even fields, respectively, the resulting composite frame could be referred to as an odd-interlaced joint image. The real part of the FT of the source image can be extracted from the sum of the joint power spectra of the even- and the odd-interlaced images. Similarly, the imaginary part of the FT of the source image can be extracted from the difference in the joint power spectra of the even- and the odd-interlaced images.

Expressions for the JTI spectra, produced by horizontal and vertical displacement of the source and the reference images, are derived and presented in this subsection. In general, a $2M \times 2N$ joint image written to a binary SLM is composed of four $M \times N$ component images, each of which is a source, reference, or zero image. A single source image, a single reference image, and two zero images may be interleaved vertically, horizontally, or diagonally. Multiple source and reference images may also be interfered by construction of the appropriate joint image such that it contains exactly four components. The notation employed in this subsection is quite general, and the expressions may be applied to determine the joint spectrum of any desired source–reference displacement. The JTI spectra for horizontal and vertical source–reference image displacements, and the algorithms necessary to extract the real and the imaginary parts of the source, are presented in this subsection. Expressions for the JTI spectra of other possible source–reference joint images and their algorithms are in Appendix A.

To accommodate all possible single-pixel mutual displacements between the source and the reference images, the $2M \times 2N$ pixels of a binary SLM are grouped into an $M \times N$ array of 2×2 pixel cells. The set of corresponding pixels from each cell is assigned its own $M \times N$ component image that can be a source, a reference, or a zero image. (For example, the top right pixel of each cell would display one of the four component images.) The pixel mapping from component images $t(m, n)$, $u(m, n)$, $v(m, n)$, $w(m, n)$ to joint image $f(m, n)$ for $M = N = 2$ is illustrated in Fig. 2.

The source and the reference images are represented here as infinite 2-D complex sequences, $f = \{f(m, n)\}$, where $m = 0, 1, 2, \ldots, M - 1$ and $n = 0, 1, 2, \ldots, N - 1$, and all other elements in the sequence are set to zero. The optical FT of a source or a reference image is then given by its discrete-space Fourier transform (DSFT), which is continuous and 2π periodic, multiplied by the FT of the pixel aperture function of the SLM. The DSFT is the spatial analog of the discrete-time Fourier transform encountered in discrete-time signal processing.[21] The FT of the pixel aperture function of an SLM with square pixels is described by a 2-D sinc function. If the linear fill factor of the SLM is given by a, the initial zeros of the sinc function appear at $k = \pm(2\pi)/a$.

A joint source–reference image is constructed mathematically by upsampling of each component image and then summation of unit-shifted versions of the upsampled images. The component images are

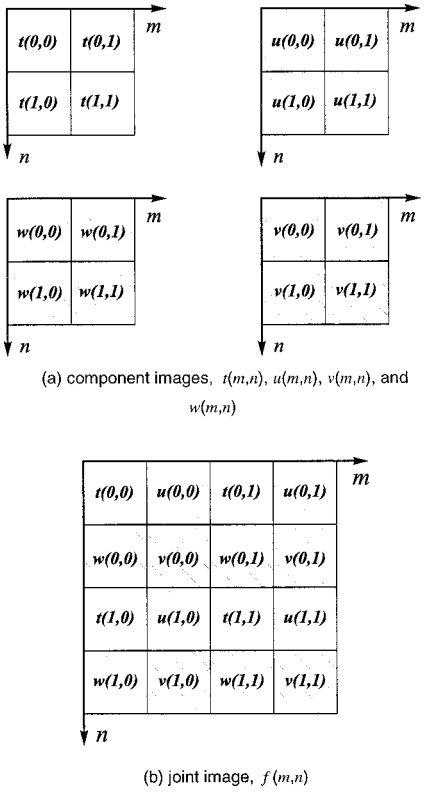


Fig. 2. (a) Component images assembled to construct (b) the joint image.

upsampled by insertion of rows and columns of zeros between each row and column of valid pixel data. If $t(m, n)$ is an $M \times N$ image sequence, its twice upsampled representation is a $2M \times 2N$ image sequence described as

$$\tilde{t}(m, n) = \begin{cases} t\left(\frac{m}{2}, \frac{n}{2}\right) & n, m \text{ even} \\ 0 & \text{otherwise} \end{cases}, \tag{5}$$

where $m = 0, 1, 2, \ldots, 2M - 1$ and $n = 0, 1, 2, \ldots, 2N - 1$. The $2M \times 2N$ composite image sequence $f(m, n)$ illustrated in Fig. 2(b) is then described as

$$f_{wv}^{tu}(m, n) = \tilde{t}(m, n) + \tilde{u}(m - 1, n) + \tilde{v}(m - 1, n - 1) + \tilde{w}(m, n - 1), \tag{6}$$

where the subscript on the joint image represents the spatial position of the four component images. (The subscript notation is in a raster–image coordinate system where the x and the m axes increase from left to right and the y and n axes increase from top to bottom.) If the DSFT's of the original images are given by $T(\mathbf{k})$, $U(\mathbf{k})$, $V(\mathbf{k})$, and $W(\mathbf{k})$, where $\mathbf{k} = (k_x, k_y)$ represents the spatial angular frequency dependence, the DSFT's of the upsampled images are given by $T(2\mathbf{k})$, $U(2\mathbf{k})$, $V(2\mathbf{k})$, and $W(2\mathbf{k})$, respectively.[22] The upsampling operation stretches the component images in the spatial domain by a factor of 2; the spatial-domain expansion corresponds to a

frequency-domain contraction by a factor of 2. The DSFT of the joint image is then given by

$$F_{wv}^{tu}(\mathbf{k}) = T(2\mathbf{k}) + \exp(ik_x)U(2\mathbf{k}) + \exp[i(k_x + k_y)]V(2\mathbf{k}) + \exp(ik_y)W(2\mathbf{k}), \quad (7)$$

and its power (or intensity) spectrum is given by

$$I_{wv}^{tu}(\mathbf{k}) = |F_{wv}^{tu}(\mathbf{k})|^2. \quad (8)$$

The general intensity spectrum described by Eq. (8) consists of 16 complex-valued terms that sum to a real-valued intensity. Expressions for various JTI spectra can be derived by setting each of the four general component spectra in Eq. (7) to source spectrum $S(2\mathbf{k})$, reference spectrum $R(2\mathbf{k})$, or zero and then reducing Eq. (8) to an interference equation form. For example, the expression for the JTI spectrum produced by a nonshifted source, $T(2\mathbf{k}) = S(2\mathbf{k})$, and a right-shifted reference, $U(2\mathbf{k}) = R(2\mathbf{k})$, where the other two images are set to zero, $W(2\mathbf{k}) = V(2\mathbf{k}) = 0$, would be expressed by

$$\begin{aligned} I_{00}^{sr}(\mathbf{k}) &= |S(2\mathbf{k}) + \exp(ik_x)R(2\mathbf{k})|^2 \\ &= I_s + I_r + 2[\cos(k_x)\Re(SR^*) \\ &\quad + \sin(k_x)\Im(SR^*)], \end{aligned} \quad (9)$$

where I_s and I_r represent the power spectra of the source and the reference images, respectively. The frequency-space dependence $2\mathbf{k} = (2k_x, 2k_y)$ has been omitted from both the power spectra and the real and the imaginary parts of the complex function SR^* to improve the clarity of the interference equations.

The JTI spectrum in Eq. (9) represents the interference pattern produced by a joint image composed of a nonshifted source image and a right-shifted reference image. The expression for the interference pattern is identical to the standard interference equation. It has three components:

1. The power spectrum of the source image.
2. The power spectrum of the reference image.
3. The interference term given by two times the sum of a cosine modulated $\Re(SR^*)$ and sine modulated $\Im(SR^*)$.

The modulating functions that premultiply the real and the imaginary parts of the complex function SR^* are due to the mutual source–reference displacement in the spatial domain.

All JTI spectra have a simple three-component form characteristic of the interference equation. In the example in Eq. (9) the sinusoidal modulation terms are simple trigonometric functions. However, for more complicated JTI spectra, the sinusoidal modulation functions become more complicated, but the general form of the interference equation is still evident. The real and the imaginary parts of SR^* are always modulated by terms composed of cosine and sine functions, respectively.

E. Horizontal and Vertical Joint-Transform Interference

In this subsection we consider the simple case of horizontal and vertical mutual displacements of the source and the reference images. Other more complicated source–reference displacements are considered in Appendix A. For horizontal and vertical JTI, the source and the reference images are interfered with respect to a line that can be either horizontal or vertical. If the source is displaced to the left or the right of the reference, the interference is called vertical JTI; if the source is displaced above or below the reference, it is called horizontal JTI. The two different orientations are named after their interference pattern. (The interference pattern of a left-shifted point source and a right-shifted point reference is a series of vertical bars.) The expressions for the eight different interference patterns are as follows:

$$\begin{aligned} I_{00}^{sr} &= I_{sr}^{00} = I_s + I_r + 2[\cos(k_x)\Re(SR^*) + \sin(k_x)\Im(SR^*)], \\ I_{00}^{rs} &= I_{rs}^{00} = I_s + I_r + 2[\cos(k_x)\Re(SR^*) - \sin(k_x)\Im(SR^*)], \\ I_{r0}^{s0} &= I_{0r}^{0s} = I_s + I_r + 2[\cos(k_y)\Re(SR^*) + \sin(k_y)\Im(SR^*)], \\ I_{s0}^{r0} &= I_{0s}^{0r} = I_s + I_r + 2[\cos(k_y)\Re(SR^*) - \sin(k_y)\Im(SR^*)]. \end{aligned} \quad (10)$$

The sign of the sinusoidal modulation function in the interference term depends on the relative position of the source and the reference images. This change in sign between JTI spectra is used to extract the real and the imaginary parts of the complex function SR^*. The sine-modulated imaginary part can be determined with one of the following algorithms:

$$\begin{aligned} 4\sin(k_x)\Im(SR^*) &= I_{00}^{sr} - I_{00}^{rs}, \\ 4\sin(k_y)\Im(SR^*) &= I_{r0}^{s0} - I_{s0}^{r0}, \end{aligned} \quad (11)$$

and the cosine-modulated real part can be determined with one of the following:

$$\begin{aligned} 4\cos(k_x)\Re(SR^*) &= I_{00}^{sr} + I_{00}^{rs} - 2I_s - 2I_r, \\ 4\cos(k_y)\Re(SR^*) &= I_{r0}^{s0} + I_{s0}^{r0} - 2I_s - 2I_r. \end{aligned} \quad (12)$$

The source image power spectrum I_s can be determined from the interference pattern of a composite image in which one of the four component images is equal to the source image and all others are equal to zero,

$$I_s = I_{s0}^{00} = I_{0s}^{00} = I_{00}^{0s} = I_{00}^{s0}, \quad (13)$$

and similarly for the power spectrum of the reference image.

F. Demodulation and Spectral Weighting Functions

The real and the imaginary parts of the source spectrum, described by Eqs. (11) and (12), are modified by sinusoidal and cosinusoidal modulation functions, respectively. In general, for any arbitrary source–reference displacement, the modulation functions are composed of trigonometric functions whose period of

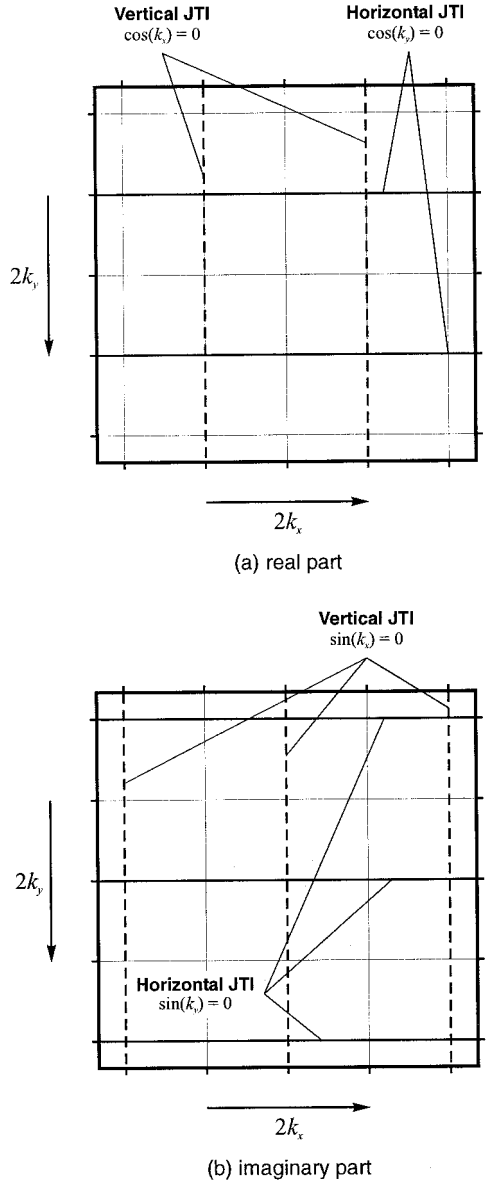


Fig. 3. Modulation-function zeros for horizontal and vertical JTI: (a) real part of spectrum; (b) imaginary part of spectrum.

modulation is much larger than that of the spectrum it is modulating. (In the present case of horizontal and vertical JTI, the 4π period of modulation is twice that of a single order of spatial-frequency space.) The spatial frequencies at which the modulation function equals zero are referred to as the modulation zeros of that function. The modulation-function zeros for horizontal and vertical JTI are illustrated in Fig. 3. The solid lines represent the zeros for the horizontal JTI, and the dashed lines represent the zeros for the vertical JTI. The cosine modulation functions are maximum at the origin and fall to zero at the boundaries of the zero-order spectrum. The sine modulation functions are maximum at the boundaries of the zero-order spectrum and fall to zero along the frequency space axes. The cosine functions have unit magnitude where the sine functions are zero and vice versa.

Clearly the source image spectrum will be difficult to calculate where the modulation functions have a magnitude close to zero. To calculate the imaginary part of the source image spectrum with acceptable accuracy near the coordinate axes, a weighted sum of the horizontal and the vertical modulated spectra is calculated on a point-by-point basis to yield an imaginary spectrum with a single indeterminate point at the origin. The point at the origin is indeterminate because both the horizontal and the vertical modulation functions are zero here. (It can be shown, however, that this point must be zero if the input image is real valued.) Similarly, the horizontal and vertical modulated spectra are combined in a weighted sum to yield a real-valued spectrum with an indeterminate point in each corner. In each case, the weighting functions are chosen to demodulate the horizontal and the vertical modulated spectra and, at the same time, weight their sum, depending on the position of the modulation zeros.

The complex-valued FT of the source image is calculated with the horizontal and the vertical JTI algorithms described by Eqs. (11) and (12). Modulated spectra (the cosine-modulated real part and the sine-modulated imaginary part) are calculated with both the horizontal and the vertical JTI algorithms. The two modulated spectra are added in a weighted sum, on a point-by-point basis, to avoid the indeterminate points associated with the zeros of the modulation functions and, at the same time, to demodulate the spectrum. The complex-valued FT of the source image, on a region of interest defined by $\eta \subseteq \{(k_x, k_y)\}$, is therefore calculated as

$$S(2k_x, 2k_y) = w_1\alpha_{\text{vert}} + w_2\alpha_{\text{horz}} + i(w_3\beta_{\text{vert}} + w_4\beta_{\text{horz}}), \tag{14}$$

where reference spectrum $R^*(\mathbf{k})$ is constant on the region of interest η. The trigonometric weighting functions are defined as

$$\begin{aligned} w_1 &= \frac{\cos(k_x)}{4[\cos^2(k_x) + \cos^2(k_y)]}, \\ w_2 &= \frac{\cos(k_y)}{4[\cos^2(k_x) + \cos^2(k_y)]}, \\ w_3 &= \frac{\sin(k_x)}{4[\sin^2(k_x) + \sin^2(k_y)]}, \\ w_4 &= \frac{\sin(k_y)}{4[\sin^2(k_x) + \sin^2(k_y)]}, \end{aligned} \tag{15}$$

and the modulated spectra calculated through both the vertical and the horizontal JTI algorithms are given by

$$\begin{aligned} \alpha_{\text{vert}} &= I_{00}^{sr} + I_{00}^{rs} - 2I_s - 2I_r, & \beta_{\text{vert}} &= I_{00}^{sr} - I_{00}^{rs}, \\ \alpha_{\text{horz}} &= I_{r0}^{s0} + I_{s0}^{r0} - 2I_s - 2I_r, & \beta_{\text{horz}} &= I_{r0}^{s0} - I_{s0}^{r0}. \end{aligned} \tag{16}$$

The weighting functions were selected by examination of the expressions for the real and the imaginary parts of the modulated spectra [left-hand side of Eqs. (11) and (12)]. The numerator of each function was chosen to be proportional to its corresponding modulated spectra; the denominator was chosen to demodulate the sum of weighted spectra correctly. This particular choice of weighting functions may not be optimal, since other variations are clearly possible. Calculation of the trigonometric weighting functions does not add to the computational complexity of the JTI algorithm because they can be precalculated and read from a static table.

The selected weighting functions demodulate the modulated spectra and, at the same time, weight the spectra according to the magnitude of its absolute value. In this way, zero weight ($w_1 = w_2 = 0$) is assigned to α_{vert} and α_{horz} at the points at which their respective modulation functions go to zero, and only a small weight is assigned to neighboring points. As shown in Fig. 3(a), the real part of the horizontal (vertical) modulated spectrum equals zero on the horizontal (vertical) lines that bound the zero-order spectrum. Similarly, zero weight ($w_3 = w_4 = 0$) is assigned to β_{vert} and β_{horz} at the points at which the modulation functions go to zero. Figure 3(b) illustrates how the imaginary part of the horizontal (vertical) modulated spectrum equals zero on the horizontal (vertical) coordinate axes.

G. Computational Efficiency

When a large number of spatial frequencies are required, the technique of JTI followed by point-by-point interferogram processing is computationally inefficient compared with a 2-D FFT. However, when a small number of points are desired, the optical technique can offer substantial computational savings. The number of operations necessary to realize the 2-D FFT of an $N \times N$ image is $O(N^2 \log_2 N)$. If only a small number of frequency-domain points are required, one can compute them directly or by employing a recursively pruned radix- (2×2) 2-D FFT algorithm.[23] The number of operations required for the direct computation of a single point is $O(N^2)$, whereas that of the pruned 2-D FFT depends on the number and distribution of the points desired. The computational complexity of the pruned algorithm is $O(N^2)$ if a minimum number of points are required and $O(N^2 \log_2 N)$ for all the points (neglecting the frequency-space distribution of the desired points and pruning overheads). The number of operations required by the JTI and point-by-point interferogram processing approach is $O(T)$, where T is the number of spatial-frequency-domain points desired. The computational effort is proportional to the number of desired points rather than the square of the image size. If only a small number of points are required, the JTI technique can be much more efficient than a pruned 2-D FFT, especially when the input images are very large.

4. Application to Moving-Object Trajectory Estimation

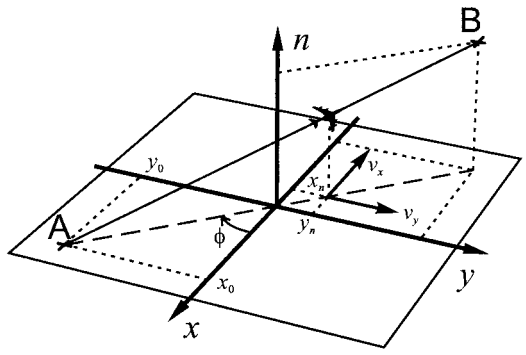
The computational ability of the optical Fourier processor proposed here is ideally suited to the mixed-domain method of moving-object trajectory estimation. The mixed-domain algorithm involves the Fourier transformation of a large number of high-resolution image frames. The number of spatial frequencies at which the FT must be calculated is very small compared with the size of the image frames (generally less than 16 points are required[24]). Consequently, the optical implementation proposed here is much more computationally attractive than either a conventional or a pruned FFT.

The mixed-domain method of moving-object trajectory estimation is particularly well suited to the detection and the trajectory estimation of small, barely discernible, moving objects of unknown position and velocity. An optical implementation is of particular computational interest because typically more than ten high-resolution image frames must be integrated to detect and track targets in a noisy environment. Applications include astronomy (detecting and tracking meteors, satellites, or other small airborne objects moving against a night-sky background), remote sensing (detecting and tracking targets in satellite, radar, sonar, and forward-looking infrared images), meteorology (tracking cloud and storm systems), and biomedical applications such as the study of cell motion.

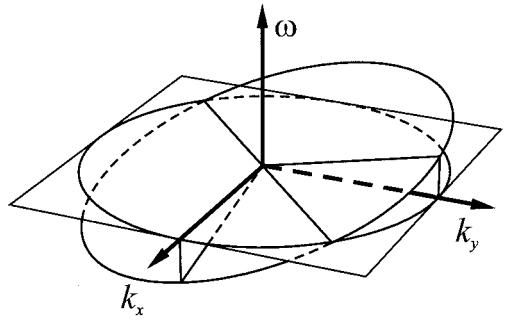
A. Mixed-Domain Trajectory Estimation

The mixed-domain approach to trajectory estimation has several advantages over a purely spatiotemporal or frequency-domain approach. The method, based on the theory of mixed-domain signal processing,[24,25] was first proposed by Knudsen and Bruton. They successfully demonstrated the trajectory estimation of linear[4,26] and nonlinear targets[5] by using their mixed-domain technique. An optical implementation of their algorithm, suitable for an image sequence containing a single moving point object, has been demonstrated by the current authors.[15,27] The optical implementation employed a single-phase interferometric optical Fourier processor and a point-diffraction interferometer to determine the real part of the optical Fourier transform. The inability to measure both parts of the complex-valued spectrum resulted in a direction ambiguity in the estimated trajectories. In addition, it was difficult to match the intensity of the reference image with that of the source to maximize the visibility of the interferograms, and consequently only very simple image sequences (i.e., a single point object on a black background with no additive noise) could be processed. The new interferometric optical Fourier processor presented here addresses these fundamental deficiencies.

The frequency-domain approach to trajectory estimation is illustrated in Fig. 4. The details of the algorithm are not presented here, and the interested reader is referred to the references cited above. An



(a)



(b)

Fig. 4. Frequency-domain trajectory estimation: (a) spatiotemporal domain; (b) 3-D frequency domain.

ideal moving point object describes a line in space–time and a plane in three-dimensional (3-D) frequency space. Multiple objects describe multiple lines in space–time and multiple planes in 3-D frequency space. An object's plane is parameterized by its velocity components. The velocity of an object is determined by estimation of the slope of its frequency–space plane. An optical Fourier processor and the JTI method of coherent detection are employed in the trajectory estimation system to determine optically the complex-valued 2-D spatial FT of each frame in an image sequence. The sequence of complex-valued spectra are processed in the time domain by use of a high-resolution spectral estimation technique. In this mixed-domain approach the space–time to 3-D frequency-space transformation is decomposed into two smaller and computationally simpler operations: The two spatial dimensions of the 3-D data set are optically transformed directly into the frequency domain; the temporal dimension is processed by use of a time-domain linear-prediction method. The temporal frequencies estimated at a preselected set of spatial frequencies constitute a set of frequency triples that describe multiple planes in 3-D frequency space. The slope of each plane is determined from the set of frequency triples to estimate the velocity components of each moving object.

B. Source Image Sequence

Two source image sequences were constructed by addition of different background sequences to an object-only sequence containing three moving objects. The 65-frame object-only sequence was composed of 128 × 128 pixel images and described the motion of three 3 × 3 pixel objects. The three objects, identified as objects M, J, and T, moved on simple linear trajectories with unit pixel displacement per frame. The velocity, speed, and direction and the initial and final positions of the objects are shown in Table 1.

An image consisting of five simple computer-generated objects was added to each frame of the object-only sequence to realize a new sequence with a stationary background. The 17th frame from this new sequence is illustrated in Fig. 5(a). The three objects can be identified as the 3 × 3 pixel squares; their trajectories are illustrated by the narrow lines indicating each object's direction.

A synthetic white-noise image was added to each frame of the object-only sequence to realize a new sequence with an additive white-noise background. The white-noise sequence was constructed by the thresholding of 65 different realizations of a 128 × 128-pixel independent and identically distributed Gaussian random process with zero mean and unit variance. The threshold was chosen to provide a predetermined number of noise pixels for each frame. The object signal-to-noise ratio (SNR) was defined as the object signal power divided by the noise power and was equal to the ratio of the number of object pixels to the number of noise pixels in each frame (both object and noise pixels must have the same intensity on a binary SLM). The 17th frame from the image sequence with SNR = 0.0 dB (27 noise pixels) is shown in Fig. 5(b).

C. Reference Image

The complex-valued FT of a source image is determined from the interferogram produced when it is interfered with a known reference image. The intensity of the source and the reference spectra must

Table 1. Actual Trajectories for Objects in the Object-Only Image Sequence

	Velocity, Speed, and Direction				Initial Position		Final Position	
Object	v_x (ppf)	v_y (ppf)	v (ppf)	ϕ (deg)	m_1	n_1	m_2	n_2
M	−1.00	0.00	1.00	180.0	32	−5	−32	−5
J	1.00	1.00	1.41	45.0	−37	−27	27	37
T	2.00	−1.00	2.24	−26.6	−64	37	64	−27

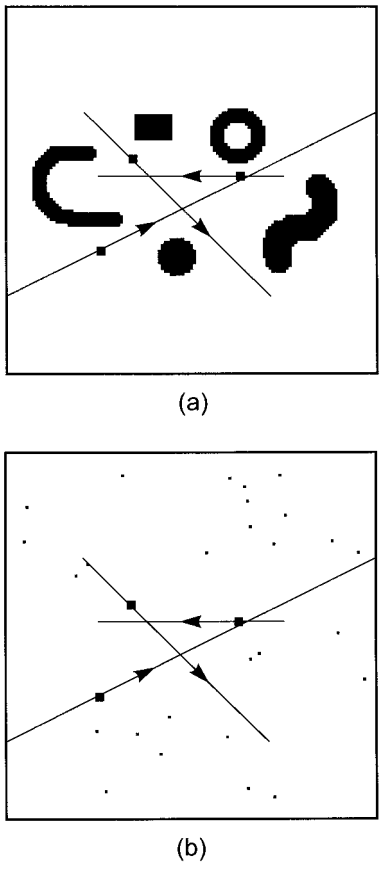


Fig. 5. Frame 17 of the source image sequence: (a) object-only sequence with simple stationary background; (b) object-only sequence with additive noise.

be matched to produce an interferogram with good visibility and to provide complex-valued source spectra with a good SNR. It can be shown[20] that the ideal frequency-domain reference has a constant intensity distribution equal to the maximum intensity of the source spectrum and a uniform phase distribution. In theory, a δ-function reference image provides this ideal frequency-domain reference; however, it is not realizable on a conventional SLM. A single pixel approximates the spatial extent of a δ-function, but its amplitude is finite and is limited by the contrast ratio and the passive transmission or reflection characteristics of the pixel.

The algorithm of the present trajectory estimation application does not require the FT at every point on the spatial-frequency plane. The algorithm necessitates only the FT calculation at a small number of specific points on the spatial-frequency plane. The points are generally chosen as uniformly distributed spatial frequencies on a circle of radius k_r^{circle} (the circle is referred to here as the frequency-domain observation circle). The spatial-domain reference image is engineered such that its FT has the required ideal properties (uniform phase and intensity matched to that of the source) on an annular region of interest η, which includes the frequency-domain observation circle. The frequency-domain reference

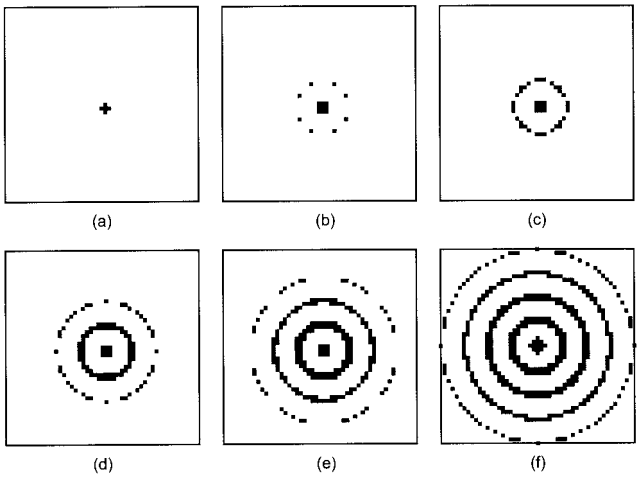


Fig. 6. Bessel reference images: (a) $L = 5$ pixels, (b) $L = 17$ pixels, (c) $L = 41$ pixels, (d) $L = 117$ pixels, (e) $L = 229$ pixels, (f) $L = 449$ pixels.

(and its corresponding spatial-domain reference image) should clearly have a circular symmetry. The zero-order Hankel transform of a zero-order Bessel function is a δ-function ring of radius k_r^{circle}. The Hankel transform pair is given by

$$k_r^{\text{circle}} J_0(k_r^{\text{circle}} r) \Leftrightarrow 2\pi\delta(k_r - k_r^{\text{circle}}), \qquad (17)$$

where $r = (m^2 + n^2)^{1/2}$ and $k_r = (k_x^{\,2} + k_y^{\,2})^{1/2}$ are radial coordinates in the spatial and frequency domains, respectively. A reference image constructed from a zero-order Bessel function satisfies the criteria for uniform phase and constant amplitude on the frequency-domain observation circle. A reference with any radial design frequency k_r^{circle} can be realized by construction of the appropriate spatial-domain Bessel image.

The intensity of the reference spectra can be matched to that of the source by modification of the amplitude of the Bessel image on a gray-scale SLM or by the changing of the threshold employed to binarize the reference on a binary SLM. Six thresholded Bessel reference images are illustrated in Fig. 6. The gray-scale Bessel images were thresholded to produce binary images consisting of exactly 5, 17, 41, 117, 229, and 449 on-pixels. The design frequency was set equal to that of the frequency-domain observation circle [$k_r^{\text{circle}} = (\pi/3)$ rad].

D. Experimental

A block diagram of the experimental setup is illustrated in Fig. 7. A beam expander equipped with a spatial filter (objective lens, L1; pinhole aperture, PH; and collimating lens, L2) was employed to expand the He–Ne laser beam to a diameter of ~7.74 mm. The expanded beam was redirected by mirrors M1 and M2 onto the surface of the digital micromirror device (DMD) at 20° to its normal.

The DMD is a 640 × 480 reflective SLM manufactured by Texas Instruments. The device consists of a superstructure array of movable aluminum mi-

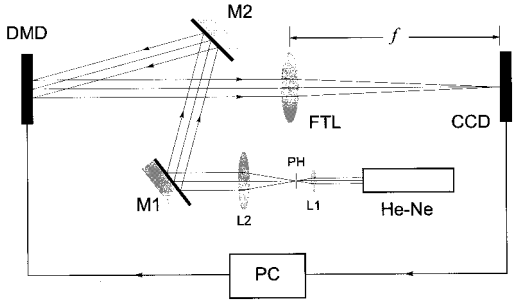


Fig. 7. Schematic diagram of the experimental setup.

cromirrors monolithically integrated over a static random-access memory circuit.[28] A micromirror at each pixel location selectively reflects incident light either into the pupil of a lens or toward an optical stop. The mirrors are 16 μm square and have a center-to-center spacing of 17 μm. The modulated light distribution in the plane of the DMD is Fourier transformed by lens [Fourier-transform lens (FTL)], and the resulting intensity spectrum is measured by a CCD camera located in the back focal plane of the lens.

Frames from the two image sequences described in Subsection 4.B were combined in sequence with an intensity-matched Bessel reference image, and this composite image sequence was written to the DMD. The complex-valued FT of each image frame was calculated with the JTI algorithms described by Eqs. (14)–(16). Examination of Eq. (16) reveals that a total of six intensity spectra are necessary to calculate the full complex-valued FT of a single image frame. The six spectra were measured sequentially at 1/6 video frame rates; however, a parallel implementation would clearly be more attractive in a real-time system.

Details of the mixed-domain trajectory estimation algorithm have been described by Knudsen and Bruton[4,5,24,26] and by the current authors.[15,27] The particular algorithm employed in this research uses a Hough transform technique to identify the three moving objects in spatiotemporal frequency space.[20] The Hough transform is often employed to detect lines and curves in a 2-D image. It is used here to detect planes in a 3-D frequency space. Frequency-domain triples from the optical Fourier processor are Hough transformed into a v_x–v_y velocity space (usually referred to as the parameter or voting space). The velocity-space voting array is searched to find the velocity components of the three moving objects.

The complex-valued FT of each image frame was calculated at a set of $T = 16$ points uniformly spaced around a circle of radius $k_r^{\text{circle}} = (\pi/3)$ rad on the spatial-frequency plane. The sequence of FT's was postprocessed off-line with MATLAB to estimate the trajectory of each object. The trajectory of each object was estimated at 9 different frames over the duration of motion (every 5th frame starting with the 25th frame), and the duration of the observation interval was 25 frames.

The trajectories of the objects in the noisy image sequence were estimated with the hybrid optical–digital trajectory estimation system and then with a completely digital system. The digital system was identical to the hybrid optical–digital system except that the FT of each image frame was calculated with a 2-D FFT algorithm with 64 bits of floating-point accuracy. Each 128 × 128 point FFT required ~1 million floating-point operations with the 2-D FFT algorithm in MATLAB version 5.2 and consumed ~70 ms of CPU time on an Intel 166-MHz Pentium processor.

E. Results and Discussion

The image sequence with a constant background was processed with a Bessel reference image composed of 257 on-pixels. The actual speed and direction, the average speed and direction estimated over the duration of motion, and the estimation errors for the three objects are summarized in Table 2. The objects were successfully tracked, and their speed and direction were correctly estimated at each estimation frame. The average speed of each object was correctly estimated to within 1.2%, and the direction of each object was estimated to within 1.8° of its actual value.

Image sequences with four different levels of white noise were constructed and tested with the system. In Table 3 we summarize the number of frames out of nine in which the velocity components of the three objects were correctly estimated. An object was defined to be correctly estimated if both of its velocity components were estimated to within two times the quantization error introduced by the Hough parameter-space voting matrix.

At a noise level corresponding to an object's SNR of 11.3 dB, the three objects were correctly estimated by the hybrid optical–digital trajectory estimation system at all nine estimation frames. (A total of 27 objects were correctly estimated in each sequence.) The number of correctly estimated objects decreased monotonically with the SNR. A further investigation at other noise levels indicated that the trajecto-

Table 2. Estimated Trajectories for Sequence with Stationary Background

	Speed Estimate			Direction Estimate		
Object	v (ppf)	δ (ppf)	Error (%)	ϕ (deg)	χ (deg)	Abs. Error
M	1.00	1.030 ± 0.004	1.0	180.0	−178.2 ± 0.4	1.8
J	1.41	1.401 ± 0.003	0.5	45.0	45.2 ± 0.1	0.2
T	2.24	2.199 ± 0.016	1.2	−26.6	−27.4 ± 1.3	0.9

Table 3. Number of Frames Correctly Estimated for Sequence with White-Noise Background

		Optical–Digital System				All-Digital System			
SNR (dB)	Noise Pixels	M	J	T	Total	M	J	T	Total
11.3	2	9	9	9	27	9	9	9	27
0.0	27	9	8	7	24	9	8	9	26
−6.0	107	7	7	3	17	8	7	5	20
−10.5	303	1	0	0	1	0	1	0	1

ries were generally estimated correctly for an SNR of greater than −3 to 0 dB (depending on the trajectory) and that the number of correctly estimated objects decreased linearly with a SNR below this threshold. When the noise power was ten times the target power (SNR = −10.5 dB), only 1 out of 27 possible objects was correctly identified.

At the highest SNR level, both the all-digital and optical–digital systems correctly estimated the trajectories of all three moving objects during the nine observation intervals. The two systems performed equally poorly at the lowest SNR's. The pixel aperture of the CCD camera was simulated in the all-digital processor by the calculation of the FT at a particular spatial frequency as an average of the four spectral values positioned at the corners of a square region equivalent in size to a CCD pixel and centered on the desired spatial frequency. When the FT at each spatial frequency was computed as a single sampled point rather than an average of four points, the hybrid optical–digital system usually outperformed the all-digital system. The frequency-domain averaging, realized by the finite CCD pixels in the hybrid optical–digital system and simulated by the four-point average in the all-digital system, appears to smooth out some of the noise contributions transformed into the frequency domain from the noisy image sequence. The frequency-domain smoothing increases the trajectory estimation performance of both systems. The optical–digital calculation of the 16 complex-valued spatial frequencies was more than 3 orders of magnitude faster than the all-digital computation. The optical–digital calculation required a few microseconds of processor time, whereas the 2-D FFT consumed 70 ms.

5. Summary and Conclusions

A novel interferometric technique called joint-transform interference (JTI) has been presented that recovers the complex-amplitude spectrum in an optical Fourier processor. The complex-valued FT of a source image is determined by interfering it with the FT of a reference image in a common-path interferometer. Unlike the other common-path interferometers, both the real and the imaginary parts of the complex-valued spectrum are determined, and, in addition, the source and the reference images are easily matched to guarantee good fringe visibility. The interferometric optical Fourier processor determines the complex-valued FT at only a select number of desired points on the spatial-frequency plane and does not calculate the entire spectrum of an image efficiently. When the number of desired points is small compared with the size of the original image, the optical Fourier processor can be much more efficient than a pruned 2-D FFT, especially when the input images are very large.

The JTI technique was experimentally verified and applied to the problem of the moving-object trajectory estimation. A hybrid optical–digital trajectory estimation system, based on the mixed-domain method of trajectory estimation, was constructed and demonstrated with a JTI-based interferometric optical Fourier processor. The optical processing was performed at 1/6 video frame rates, and the balance of the trajectory estimation algorithm was performed off line with MATLAB. The system was able to estimate the trajectory of 3 × 3 objects when the SNR was greater than −3 dB. The ability of the optical–digital system to track moving objects and accurately estimate their trajectory was equivalent to that of the all-digital system. The FT calculation at the desired frequency-domain points required the measurement of six interferograms in addition to approximately 200 floating-point operations. Calculation of the same points by the FFT required more than 3 orders of magnitude more floating-point operations.

The interferograms required by the JTI algorithms for constructing the complex-valued FT of a single source image are easily computed in parallel. The block diagram in Fig. 8 is a parallel implementation of the weighted horizontal–vertical JTI algorithm described by Eqs. (14)–(16). An appropriate reference image is generated from the source image. Four source–reference composite images are constructed by shifting the source up, down, left, and right and by adding the unit-shifted source images to the reference. Each of the four composite images, and the

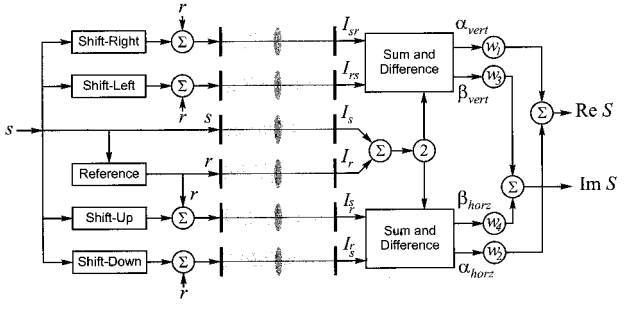


Fig. 8. Parallel interferometric optical Fourier-transform processor.

individual source and reference images, are written to the SLM of a different optical Fourier processor. The six interferograms are combined, according to Eq. (14), to yield the complex-valued source spectrum. The postprocessing is simple pixel-by-pixel addition and subtraction followed by pixel-by-pixel weighting.

Appendix A

1. Diagonal Joint-Transform Interference

A single source image is diagonally interfered with a single reference image. The source and the reference images may be interfered with respect to a line that is diagonal with either a positive or a negative slope. If the source and reference are displaced along a diagonal line with a positive (negative) slope, the interference is called negative (positive) diagonal JTI. (The interference pattern of a point source and point reference, mutually displaced on a positive diagonal line, is a series of negative diagonal bars and vice versa.) The expressions for the four possible interference patterns are

$$\begin{aligned} I_{0r}^{s0} &= I_s + I_r + 2[c_2\Re(SR^*) + s_2\Im(SR^*)], \\ I_{0s}^{r0} &= I_s + I_r + 2[c_2\Re(SR^*) - s_2\Im(SR^*)], \\ I_{s0}^{0r} &= I_s + I_r + 2[c_1\Re(SR^*) + s_1\Im(SR^*)], \\ I_{r0}^{0s} &= I_s + I_r + 2[c_1\Re(SR^*) - s_1\Im(SR^*)], \end{aligned} \tag{A1}$$

where the cosine and the sine modulation functions are given by $c_1 = \cos(k_x - k_y)$, $c_2 = \cos(k_x + k_y)$ and $s_1 = \sin(k_x - k_y)$, $s_2 = \sin(k_x + k_y)$, respectively. As above, the change in sign of the sine modulating function can be utilized to extract the real and the imaginary parts of SR^*.

2. Three-Component Joint-Transform Interference

Two diagonally displaced source images are interfered with a single reference image, or two diagonally displaced reference images are interfered with a single source image. If the source pair and single reference (or single source and reference pair) are displaced along a diagonal line with positive (negative) slope, the interference is called negative (positive) three-component JTI. Equations for the JTI spectra of four representative cases are

$$\begin{aligned} I_{sr}^{0s} &= 2(1 + c_1)I_s + I_r + 2[c_4\Re(SR^*) + s_4\Im(SR^*)], \\ I_{s0}^{rs} &= 2(1 + c_1)I_s + I_r + 2[c_4\Re(SR^*) - s_4\Im(SR^*)], \\ I_{sr}^{r0} &= I_s + 2(1 + c_2)I_r + 2[c_4\Re(SR^*) + s_3\Im(SR^*)], \\ I_{0r}^{rs} &= I_s + 2(1 + c_2)I_r + 2[c_4\Re(SR^*) - s_3\Im(SR^*)]. \end{aligned} \tag{A2}$$

The cosine and sine modulation functions are given by $c_4 = \cos(k_x) + \cos(k_y)$ and by $s_3 = \sin(k_x) - \sin(k_y)$, $s_4 = \sin(k_x) + \sin(k_y)$, respectively. The modulated power spectra of the double-source and double-reference images, $2(1 + c_1)I_s$ and $2(1 + c_2)I_r$, may be determined in practice from the joint images I_{s0}^{0s} and I_{0r}^{r0}, respectively.

3. Symmetric Four-Component Joint-Transform Interference

Symmetric JTI is a special case of a four-component JTI. Two diagonally displaced source images are interfered with two diagonally displaced reference images. There is only one possible orientation (positive and negative configurations do not exist). The interference functions contain only the real part of the complex function owing to the special symmetry of the joint image. Equations for the JTI spectra of both possible configurations are

$$\begin{aligned} I_{rs}^{sr} &= 2(1 + c_2)I_s + 2(1 + c_1)I_r + 4c_4\Re(SR^*), \\ I_{sr}^{rs} &= 2(1 + c_1)I_s + 2(1 + c_2)I_r + 4c_4\Re(SR^*). \end{aligned} \tag{A3}$$

Neither equation contains the imaginary part of the complex-valued function SR^*, and consequently its imaginary part cannot be determined.

4. Four-Component Joint-Transform Interference

Three source images are interfered with a single reference image, or three reference images are interfered with a single source image. If the source triple and single reference (or single source and reference triple) are displaced along a diagonal line with positive (negative) slope, the interference is called negative (positive) four-component JTI. Equations for four representative cases are

$$\begin{aligned} I_{sr}^{ss} &= (3 + 2c_5)I_s + I_r + 2[c_6\Re(SR^*) + s_6\Im(SR^*)], \\ I_{ss}^{rs} &= (3 + 2c_5)I_s + I_r + 2[c_6\Re(SR^*) - s_6\Im(SR^*)], \\ I_{sr}^{rr} &= I_s + (3 + 2c_6)I_r + 2[c_5\Re(SR^*) + s_5\Im(SR^*)], \\ I_{rr}^{rs} &= I_s + (3 + 2c_6)I_r + 2[c_5\Re(SR^*) - s_5\Im(SR^*)], \end{aligned} \tag{A4}$$

where the cosine modulation functions are defined as

$$\begin{aligned} c_5 &= \cos(k_x) + \cos(k_y) + \cos(k_x - k_y), \\ c_6 &= \cos(k_x) + \cos(k_y) + \cos(k_x + k_y), \end{aligned} \tag{A5}$$

and the sine modulation functions are defined as

$$\begin{aligned} s_5 &= \sin(k_x) - \sin(k_y) + \sin(k_x - k_y), \\ s_6 &= \sin(k_x) + \sin(k_y) + \sin(k_x + k_y). \end{aligned} \tag{A6}$$

The modulated power spectra of the triple-source and triple-reference images, $(3 + 2c_5)I_s$ and $(3 + 2c_6)I_r$, may be determined in practice from the joint images I_{ss}^{0s} and I_{0r}^{rr}, respectively.

References

1. J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," Math. Comput. **19,** 297–301 (1965).
2. J. D. Markel, "FFT pruning," IEEE Trans. Acoust. Speech Signal Process. **AU-19,** 305–311 (1971).
3. T. Smit, M. R. Smith, and S. T. Nichols, "Efficient sinc function

interpolation technique for center padded data," IEEE Trans. Acoust. Speech Signal Process. **38,** 1512–1517 (1990).
4. K. S. Knudsen and L. T. Bruton, "Moving object detection and trajectory estimation in the transform/spatiotemporal mixed domain," in *Proceedings of the 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing, San Francisco, March 1992* (Institute of Electrical and Electronic Engineers, New York, 1992), Vol. 3, pp. 505–508.
5. K. S. Knudsen and L. T. Bruton, "Moving object nonlinear trajectory estimation in the transform/spatiotemporal mixed domain," in *Proceedings of the 1992 IEEE Custom Integrated Circuits Conference, San Diego, May 1992* (Institute of Electrical and Electronic Engineers, New York, 1992), Vol. 5, pp. 2481–2484.
6. T. M. Turpin, "Spectrum analysis using optical processing," Proc. IEEE **69,** 79–92 (1981).
7. M. King, W. R. Bennett, L. B. Lambert, and M. Arm, "Real-time electro-optical signal processors with coherent detection," Appl. Opt. **6,** 1367–1375 (1967).
8. H. R. Carleton, W. T. Maloney, and G. Meltz, "Collinear heterodyning in optical processors," Proc. IEEE **57,** 769–775 (1969).
9. A. VanderLugt, "Interferometric spectrum analyzer," Appl. Opt. **33,** 2770–2779 (1981).
10. C. C. Aleksoff and N. S. Subotic, "Compact real-time interferometric Fourier transform processors," in *Optical Information Processing Systems and Architectures II*, B. Javidi, ed., Proc. SPIE **1347,** 427–440 (1990).
11. A. VanderLugt, *Optical Signal Processing* (Wiley, New York, 1992).
12. W. P. Linnik, "A simple interferometer for the investigation of optical systems," C. R. Acad. Sci. URSS **5,** 208–210 (1933) (in Russian).
13. E. C. Tam, S. W. Tannone, F. T. S. Yu, and D. A. Gregory, "Closed-loop binary phase correction of an LCTV using a point diffraction interferometer," IEEE Photon. Technol. Lett. **2,** 143–146 (1990).
14. Y. Zhang, E. Kanterakis, A. Katz, and J.-M. Wang, "Optoelectronic wavelet processors based on Smartt interferometry," Appl. Opt. **33,** 5279–5286 (1994).
15. P. M. Lane and M. Cada, "An optical Fourier processor and point-diffraction interferometer for moving object trajectory estimation," Appl. Opt. **38,** 4306–4315 (1999).
16. D. Malacara, *Optical Shop Testing*, 2nd ed. (Wiley, New York, 1991).
17. H. Kadono, M. Ogusu, and S. Toyooka, "Phase shifting common path interferometer using a liquid-crystal phase modulator," Opt. Commun. **110,** 391–400 (1994).
18. C. R. Mercer and K. Creath, "Liquid-crystal point-diffraction interferometer," Opt. Lett. **19,** 916–918 (1994).
19. C. R. Mercer and K. Creath, "Liquid-crystal point-diffraction interferometer for wave-front measurements," Appl. Opt. **35,** 1633–1642 (1996).
20. P. M. Lane, "The complex-valued optical Fourier transform and its application to moving-object trajectory estimation," Ph.D. dissertation (Dalhousie University, Halifax, N.S., 1999).
21. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, N.J., 1975).
22. G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge Press, Wellesley, Mass., 1996).
23. K. S. Knudsen and L. T. Bruton, "Recursive pruning of the 2-D DFT with 3-D signal processing applications," IEEE Trans. Signal Process. **41,** 1340–1356 (1993).
24. K. S. Knudsen, "Multidimensional mixed domain signal processing," Ph.D. dissertation (University of Calgary, Calgary, Al., 1992).
25. K. S. Knudsen and L. T. Bruton, "Mixed domain filtering of multidimensional signals," IEEE Trans. Circuits Syst. Video Technol. **1,** 260–268 (1991).
26. K. S. Knudsen and L. T. Bruton, "Transform/spatiotemporal mixed domain moving object tracking and enhancement," in *Proceedings of the 1993 European Conference on Circuit Theory and Design, Davos, Switzerland, August 1993* (Elsevier, Amsterdam, 1993), pp. 589–594.
27. P. M. Lane, K. S. Knudsen, and M. Cada, "Moving object trajectory estimation using an optical Fourier processor," in *1998 International Conference on Applications of Photonic Technology*, G. A. Lampropoulos, ed., Proc. SPIE **3491,** 939–943 (1998).
28. L. J. Hornbeck, "Digital light processing and MEMS: timely convergence for a bright future (invited plenary paper)," in *Micromaching and Microfabrication Process Technology*, K. W. Markus, ed., Proc. SPIE **2639,** p. 2 (abstract only), full paper available from Texas Instruments, Dallas, Tex., 1995.