# The Complex-Valued Optical Fourier Transform and its Application to Moving-Object Trajectory Estimation

by

Pierre M. Lane

A Dissertation Submitted to the
Faculty of Engineering
in Partial Fulfillment of the Requirements
for the degree of

DOCTOR OF PHILOSOPHY

Major Subject: Electrical and Computer Engineering

APPROVED:

_____
Dr. M. Cada, Supervisor

_____
Dr. P. Gregson

_____
Dr. W. Phillips

_____
Dr. W. Robertson

_____
Dr. D. Gingras, External Examiner

DALHOUSIE UNIVERSITY - DALTECH

Halifax, Nova Scotia 1999

DALTECH LIBRARY

**"Authority to Distribute Manuscript Thesis"**

TITLE:

The Complex-Valued Optical Fourier Transform
and its Application to
Moving-Object Trajectory Estimation

The above library may make available or authorize another library to make available individual photo/microfilm copies of this thesis without restrictions.

Full Name of Author: Pierre M. Lane

Signature of Author: _____

Date: June 18, 1999

*To my best friend,*
*Terri.*

# Table of Contents

LIST OF TABLES ..... viii
LIST OF FIGURES ..... ix
LIST OF SYMBOLS AND ABBREVIATIONS ..... xiii
ACKNOWLEDGEMENTS ..... xx
ABSTRACT ..... xxi

## CHAPTER 1

**INTRODUCTION** ..... 1

1.1 Optical Computing ..... 2
1.1.1 Analog Optical Computing ..... 4
1.1.2 Digital Optical Computing ..... 9

1.2 Main Contributions ..... 11

1.3 Dissertation Framework ..... 15
1.3.1 Objectives and Scope ..... 15
1.3.2 Approach and Methodology ..... 16
1.3.3 Organization and Overview ..... 18

## CHAPTER 2

**OPTICAL SIGNAL AND IMAGE PROCESSING** ..... 20

2.1 Introduction ..... 21

2.2 Fourier Optics ..... 23
2.2.1 Propagation of Light in Free Space ..... 25
2.2.2 Optical Fourier Transform ..... 27

2.3 Fourier Transform of a Pixelated SLM as a DSFT ..... 30
2.3.1 Spectrum of an SLM ..... 30
2.3.2 DSFT of a 1-D Aperture ..... 34

2.4 Coherent Detection ..... 35
2.4.1 Interference of Two Waves ..... 37
2.4.2 Mach-Zehnder Interferometer ..... 39
2.4.3 Twyman-Green Interferometer ..... 40
2.4.4 Point-Diffraction Interferometer ..... 41

2.5 PDI-Based Interferometric Optical Fourier Processor ..... 44

2.6 Chapter Summary ..... 47

## CHAPTER 3

**SINGLE-OBJECT TRAJECTORY ESTIMATION** ..... 48

3.1 Introduction ..... 49

3.2 Trajectory Estimation ..... 53
3.2.1 Frequency-Domain Trajectory Estimation ..... 53
3.2.2 Optical Implementation ..... 56
3.2.3 Interpretation of the Mixed Domain ..... 57
3.2.4 Accuracy and Aliasing ..... 60

3.3 Experimental ..... 64

3.4 Results and Discussion ..... 70
3.4.1 Speed Estimation ..... 70
3.4.2 Direction Estimation ..... 73
3.4.3 Estimation Accuracy ..... 74
3.4.4 Multiple-Object Trajectory Estimation ..... 76

3.5 Chapter Summary ..... 77

## CHAPTER 4

**JOINT TRANSFORM INTERFERENCE** ..... 79

4.1 Introduction ..... 80
4.1.1 Phase-Shifting Interferometry ..... 80
4.1.2 Spatial Modulation ..... 81

4.2 Principle of JTI ..... 83

4.3 Joint Transform Interference on a Binary SLM....85
4.3.1 Horizontal and Vertical JTI....90
4.3.2 Diagonal JTI....92
4.3.3 Three-Component JTI....93
4.3.4 Symmetric Four-Component JTI....94
4.3.5 Four-Component JTI....95

4.4 Symmetric and Non-Symmetric Reference images....97

4.5 JTI Example....98

4.6 Modulation-Function Zeros....101

4.7 Signal-to-Noise Ratio....106
4.7.1 Dynamic Range and Region of Interest....107
4.7.2 Source and Reference Spectra....108
4.7.3 Saturation Intensity....110
4.7.4 Optimum Reference Intensity....111
4.7.5 Noise in Detection....113
4.7.6 Imaginary Part SNR....115
4.7.7 Real Part SNR....118
4.7.8 Maximizing the SNR....120

4.8 Chapter Summary....121

## CHAPTER 5

**MULTIPLE-OBJECT TRAJECTORY ESTIMATION**....123

5.1 Introduction....124

5.2 The Digital Micromirror Device (DMD)....126
5.2.1 Pixel Structure and Switching Time....127
5.2.2 The DMD as a Blazed Diffraction Grating....128

5.3 Experimental Verification of JTI....133
5.3.1 Demodulation and Spectral Weighting Functions....134
5.3.2 Computational Efficiency....137
5.3.3 Experimental Setup and Procedure....138
5.3.4 Results and Discussion....140

5.4 Trajectory Estimation with a Single-Pixel Reference....142
5.4.1 Single Point-Object Trajectory Estimation....143
5.4.2 Multiple Point-Object Trajectory Estimation....151

5.5 Trajectory Estimation with a MultiPixel Reference ..... 159
5.5.1 Bessel Function Reference Image ..... 160
5.5.2 Multiple 3×3 Pixel Objects ..... 163
5.5.3 Multiple 3×3 Pixel Objects on a Stationary Background ..... 167
5.5.4 Multiple 3×3 Pixel Objects with Additive Noise ..... 169
5.5.5 Multiple Single-Pixel Objects with Additive Noise ..... 173

5.6 Chapter Summary ..... 176

**CHAPTER 6**

**CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK** ..... 178

6.1 Objectives, Achievements, and Contributions ..... 179

6.2 Suggestions for Further Work ..... 182
6.2.1 Joint Transform Interference ..... 182
6.2.2 Moving-Object Trajectory Estimation ..... 187

6.3 Other Applications ..... 190

**REFERENCES** ..... 192

**APPENDIX A**

TEMPORAL AND SPATIAL FOURIER TRANSFORMS ..... 204

**APPENDIX B**

MODIFIED FBLP ALGORITHM ..... 207

**APPENDIX C**

JTI EQUATIONS AND ALGORITHMS ..... 211

# List of Tables


**Table 3-1** Actual and Estimated Speeds ..... 72

**Table 3-2** Actual and Estimated Directions..... 75

**Table 5-1** Actual Trajectories for Objects A, B, C, D, and I through P ..... 144

**Table 5-2** Actual and Estimated Trajectories for Objects A, B, C, and D..... 149

**Table 5-3** Actual and Estimated Trajectories for Objects I through P ..... 151

**Table 5-4** Actual Trajectories for Image Sequences 1 and 2..... 154

**Table 5-5** Estimated Trajectories for Image Sequences 1 and 2..... 159

**Table 5-6** Actual Trajectories for Image Sequences 3 and 4..... 164

**Table 5-7** Estimated Trajectories for Image Sequences 3 and 4..... 165

**Table 5-8** Estimated Trajectories for Image Sequences 3 and 4 with a Background ..... 168

**Table 5-9** Number of Frames Correctly Estimated in Image Sequence 3 for $3\times3$ pixel objects (Total of 9 Estimation Frames) ..... 171

**Table 5-10** Number of Frames Correctly Estimated in Image Sequence 4 for $3\times3$ pixel objects (Total of 9 Estimation Frames) ..... 172

**Table 5-11** Number of Frames Correctly Estimated in Image Sequence 3 for single-pixel objects (Total of 9 Estimation Frames)..... 173

**Table 5-12** Number of Frames Correctly Estimated in Image Sequence 4 for single-pixel objects (Total of 9 Estimation Frames)..... 174

# List of Figures


**Figure 1-1** Conventional 4-$f$ optical correlator. ..... 6

**Figure 2-1** An arbitrary wave shown as a superposition of plane waves. ..... 24

**Figure 2-2** Plane waves incident on a thin optical element whose amplitude transmittance is a harmonic function with spatial period $\Lambda_x = \nu_x^{-1}$. 26

**Figure 2-3** Principle of the optical Fourier transform. The amplitude distribution $f(x, y)$ in the back focal-plane of the lens (FTL) is Fourier transformed into the front focal-plane $g(x, y)$. ..... 28

**Figure 2-4** Spectrum of a pixelated SLM. The OTF is determined by the pixel aperture and the periodicity is determined by the pixel pitch ($a = b = 0.8$). ..... 33

**Figure 2-5** Discrete space Fourier transform (DSFT) of a rect-sequence. The SLM spatial duty cycle (pixel width-to-pitch ratio) is $a = 0.8$ and the sequence is length $N = 9$. ..... 35

**Figure 2-6** Intensity distribution versus phase for the interference of two plane waves of unequal intensity. ..... 38

**Figure 2-7** Mach-Zehnder interferometer. ..... 39

**Figure 2-8** Twyman-Green interferometer. ..... 41

**Figure 2-9** Point-diffraction interferometer (PDI). ..... 42

**Figure 2-10** Optical Fourier transformer with coherent detection. ..... 45

**Figure 3-1** Spatiotemporal-domain representation of a moving point-object with velocity vector $\mathbf{v} = \left[v_x \; v_y\right]^{\mathsf{T}}$ and angle $\phi$. ..... 54

**Figure 3-2** Frequency-domain representation of a moving point-object when the temporal Fourier transform is applied to the real part of the mixed-domain representation only. ..... 57

**Figure 3-3** Moving object trajectories. ..... 58

**Figure 3-4** Real part of the mixed-domain representation of a moving object. The top and bottom images sequences represent the trajectories of the two objects shown in Figure 3-3. ..... 59

**Figure 3-5** Relationship between frequency-plane radius $k_r^{circle}$ and maximum target speed $v_{max}$. ..... 63

**Figure 3-6** Block diagram of the optical Fourier processor with point-diffraction interferometer (single-pixel reference). ..... 65

**Figure 3-7** Target animation and data acquisition flow chart. ..... 67

**Figure 3-8** Post processing and trajectory estimation flow chart. ..... 69

**Figure 3-9** Speed experiment actual trajectories. The duration of each trajectory is 64 frames and the lengths ranged from 36 pixels (trajectory A-A') to 282 pixels (trajectory D-D'). ..... 70

**Figure 3-10** Speed estimation for targets moving with velocity 0.56, 1.12, 2.24, and 4.47 ppf. ..... 71

**Figure 3-11** Direction estimation for targets moving on paths 0, ±45, ±90, ±135, and 180º to the $x$ axis. ..... 74

**Figure 3-12** Relationship between absolute estimation error and frequency-domain radius $k_r^{circle}$ ..... 76

**Figure 4-1** Construction of a joint image from a source and reference image. 83

**Figure 4-2** Component images shown in part (a) are assembled to construct the joint image shown in (b). ..... 88

**Figure 4-3** JTI example: (a) individual source and reference intensity spectra; (b) joint intensity spectra; (c) spectra sum and difference; (d) modulation and normalization functions; (e) real and imaginary parts of source amplitude spectra. ..... 99

**Figure 4-4** Modulation function zeros for horizontal and vertical JTI: (a) real part of spectrum; and (b) imaginary part of spectrum. ..... 102

**Figure 4-5** Modulation function zeros for diagonal JTI: (a) real part of spectrum; and (b) imaginary part of spectrum. ..... 103

**Figure 4-6** Modulation function zeros for three-component JTI: (a) real part of spectrum; and (b) imaginary part of spectrum. ..... 104

**Figure 4-7** Modulation function zeros for 4-component JTI: (a) real part of spectrum; and (b) imaginary part of spectrum. ..... 105

**Figure 4-8** Interference effect illustrated as a optical bias modulated by a cosinusoidal information signal..... 111

**Figure 4-9** Normalized intensity variation as a function of phase for different source intensities: (a) maximum intensity; (b) $\frac{1}{4}$ of maximum intensity; (c) $\frac{1}{16}$ of maximum intensity; and (d) $\frac{1}{64}$ of maximum intensity. ..... 113

**Figure 4-10** Block diagram illustrating the extraction of the imaginary part of the source spectrum from two source-reference joint spectra with opposite displacement phase. ..... 116

**Figure 4-11** Block diagram illustrating the extraction of the real part of the source spectrum from two source-reference joint spectra with opposite displacement phase and their individual spectra. ..... 118

**Figure 5-1** DMD pixel structure: (a) switching principle of a single mirror; and (b) SEM photomicrograph of a $3\times3$ mirror array (SEM reproduced with permission from Texas Instruments [TI98]). ..... 127

**Figure 5-2** A row of DMD on-state mirrors as a blazed diffraction grating. ... 130

**Figure 5-3** DMD far-field diffraction pattern illustrating the absolute value of its OTF (diffraction envelope) and replicated DSFT spectra ($2\pi$-periodic interference maxima). ..... 132

**Figure 5-4** Experimental setup employed to verify the JTI method of coherent detection and for the trajectory estimation experiments. ..... 138

**Figure 5-5** Experimental method and post-processing procedure employed to verify the weighted three-component JTI algorithm. ..... 139

**Figure 5-6** Spectrum of single-pixel $(m,n)=(3,3)$ source image: (a) real part; (b) imaginary part; (c) real and (d) imaginary parts along $k_x$ axis; and (e) real and (f) imaginary parts along $k_y$ axis..... 141

**Figure 5-7** Spectrum of single-pixel $(m,n)=(10,10)$ source image: (a) real part; (b) imaginary part; (c) real; and (d) imaginary parts along $k_x$ axis; and (e) real and (f) imaginary parts along $k_y$ axis..... 142

**Figure 5-8** JTI and trajectory estimation post-processing procedure. ..... 147

**Figure 5-9** Speed estimates for objects A, B, C, and D. Primed objects move along the same path in the opposite direction. .....148

**Figure 5-10** Direction estimates for objects I through P. .....150

**Figure 5-11** Actual trajectories for: (a) image sequence 1 (objects O, L and C'); and (b) image sequence 2 (objects R, Q, and S). .....155

**Figure 5-12** Modified trajectory estimation algorithm extended to accommodate three moving objects. .....156

**Figure 5-13** Speed and direction estimates for the three objects O, L, and C' of image sequence 1. .....157

**Figure 5-14** Speed and direction estimates for the three objects R, Q, and S of image sequence 2. .....158

**Figure 5-15** Algorithm to generate a binarized Bessel-function reference image. .....162

**Figure 5-16** Bessel reference images: (a) $L$=5 pixels; (b) $L$=17 pixels; (c) $L$=41 pixels; (d) $L$=117 pixels; (e) $L$=229 pixels; and (f) $L$=449 pixels.....163

**Figure 5-17** Average fringe visibility versus the ratio of maximum source intensity to reference intensity for various Bessel reference images. .....164

**Figure 5-18** Synthetic backgrounds and object trajectories: (a) image sequence 3 (objects M, J, and T) with background; and (b) image sequence 4 (objects U, V, and M) with background. The $3\times3$ pixel objects are shown at frame $n$=17. .....167

**Figure 5-19** Image sequence 3 with four levels of noise: (a) SNR = 11.3 dB (2 noise pixels); (b) SNR=0.0 dB (27 noise pixels); (c) SNR=-6.0 dB (107 noise pixels); and (d) SNR=–10.5 dB (303 noise pixels). The objects are shown at frame $n$=17.....170

**Figure 6-1** Parallel interferometric optical Fourier transform processor. .....183

# List of Symbols and Abbreviations

Greek Symbols

| | |
|---|---|
| $\alpha(k_x, k_y)$ | interferogram sum image |
| $\beta(k_x, k_y)$ | interferogram difference image |
| $\gamma_{xy}, \gamma$ | DMD pixel blaze and projection onto the coordinate axes |
| $\Delta, \Delta_x, \Delta_y$ | SLM pixel pitch (center-to-center spacing), [m] |
| $\Delta', \Delta'_x, \Delta'_y$ | SLM pixel dimension, [m] |
| $\Delta I = 2^{-b}$ | intensity quantization width |
| $\Delta v, \Delta v_x, \Delta v_y$ | velocity quantization width, [ppf] |
| $\theta = -\text{atan}\left(\frac{x_n}{y_n}\right)$ | angle between lines of constant phase and $k_x$ axis |
| $\theta_k$ | azimuthal component of spatial angular frequency |
| $\theta_x, \theta_y$ | angle between wavevector $\mathbf{k}$ and the $y$-$z$ and $x$-$z$ planes |
| $\theta_i, \theta_r$ | DMD angle of incidence and reflectance |
| $\lambda$ | wavelength, [m] |
| $\lambda_1, \lambda_2, \ldots, \lambda_K$ | singular values of data matrix $\mathbf{A}$ |
| $\Lambda_x$ | spatial period of a grating in the $x$ direction, [m] |
| $\nu_x, \nu_y$ | spatial frequencies in the $x$ and $y$ directions, [$m^{-1}$] |
| $\sigma^2 = \frac{1}{12}2^{-2b}$ | variance of the quantization-error sequence |
| $\varphi$ | phase of a complex-valued amplitude |
| $\phi$ | phase difference between DMD pixels; pinhole diameter, [m] |
| $\phi, \hat{\phi}$ | actual and estimated directions of a moving object |
| $\phi = \mathbf{k} \cdot (\mathbf{p}_s - \mathbf{p}_r)$ | linear phase introduced by a source-reference displacement |
| $\phi_t = k_x x_0 + k_y y_0$ | phase of $S(k_x, k_y, n)$ at $\mathbf{k}_t = (k_x, k_y)$ |
| $\mathbf{\Sigma}$ | $K \times K$ diagonal matrix of singular values |
| $\omega$ | temporal angular frequency, [rad·$s^{-1}$] |
| $\{\omega_p\}$ | set of $P$ unknown angular frequencies, [rad·$s^{-1}$] |
| $\omega_t = k_x v_x + k_y v_y$ | angular frequency of $S(k_x, k_y, n)$ at $\mathbf{k}_t = (k_x, k_y)$, [rad·$s^{-1}$] |
| $\overline{\omega}(k_x, k_y)$ | average of two diametrically opposed $\omega_t$, [rad·$s^{-1}$] |
| $\hat{\boldsymbol{\omega}}$ | column vector of angular frequency estimates, [rad·$s^{-1}$] |

Roman Symbols

| | |
|---|---|
| $a, b$ | horizontal and vertical spatial duty cycles of an SLM |
| $\{a_p\}$ | set of $P$ unknown complex amplitudes |
| $\mathbf{a}$ | $(K+1)\times 1$ tap-weight vector for the prediction-error filter |
| $\mathbf{A}$ | $2(N-K)\times K$ data matrix for modified FBLP |
| $b$ | number of quantization bits |
| $\mathbf{b}$ | $2(N-K)\times 1$ desired response vector for modified FBLP |
| $c_j(\mathbf{k}),\ j=1, 2, \ldots, 6$ | cosinusoidal modulation images |
| $d$ | distance between an SLM and the FTL, [m] |
| $d=\{n_1, n_2, \ldots\}$ | set of frames at which the velocity is to be estimated |
| $D$ | duration of the estimation interval, [frames] |
| $e(n)$ | white-noise sequence |
| $e_j(n),\ j=1, 2, 3, 4$ | quantization-error sequences |
| $E_j=\{d_j-D+1, \ldots\}$ | set of frames in estimation interval $j$ |
| $f$ | focal length of a lens, [m] |
| $f(x,y)$ | arbitrary function of two spatial variables |
| $f_{\substack{t\ u\\ w\ v}}(x,y)$ | composite image composed of four component images |
| $F(k_x,k_y)$ | Fourier transform of $f(x,y)$ |
| $F_{\substack{t\ u\\ w\ v}}(k_x,k_y)$ | Fourier transform of a composite image |
| $i=\sqrt{-1}$ | imaginary number |
| $I(k_x,k_y)$ | optical intensity spectrum |
| $I(n), \hat{I}(n)$ | exact and quantized intensity sequences |
| $I_{\min}, I_{max}$ | minimum and maximum intensity |
| $I_s^{max}$ | maximum intensity of a source interferogram |
| $I_{bias}=I_s^{\max}+I_r$ | optical-bias intensity |
| $I_r(\mathbf{k})=\lvert R(\mathbf{k})\rvert^2$ | reference intensity (power) spectra |
| $I_s(\mathbf{k})=\lvert S(\mathbf{k})\rvert^2$ | source intensity (power) spectra |
| $I_{\substack{t\ u\\ w\ v}}(\mathbf{k})=\left\lvert F_{\substack{t\ u\\ w\ v}}(\mathbf{k})\right\rvert^2$ | joint transform interferogram; composite intensity spectra |
| $I_\phi(\mathbf{k}), I_{-\phi}(\mathbf{k})$ | interferograms with opposite displacement phase |
| $j$ | current estimation interval; integer |
| $k$ | wavenumber, [$m^{-1}$] |

| | |
|---|---|
| $\mathbf{k} = (k_x, k_y)$ | point on the 2-D spatial frequency plane, [rad] |
| $\mathbf{k}_{\max}$ | spatial frequency where intensity is maximum, [rad] |
| $\mathbf{k} = (k_x, k_y, k_z)$ | wavevector, [$\mathrm{m}^{-1}$] |
| $k_x$, $k_y$ | spatial angular frequencies in the $x$ and $y$ directions, [rad] |
| $(\tilde{k}_x, \tilde{k}_y) = (2k_x, 2k_y)$ | compressed spatial-frequency space, [rad] |
| $k_x'$, $k_y'$ | non-normalized spatial angular frequencies, [$\mathrm{rad \cdot m}^{-1}$] |
| $k_r$ | radial component of spatial angular frequency, [rad] |
| $k_r^{circle}$ | observation circle radial angular frequency, [rad] |
| $K = \frac{3N}{4}$ | optimum prediction order for modified FBLP |
| $l$ | integer; moving-object number |
| $L$ | number of on-pixels in a Bessel reference-image |
| $m$, $n$ | pixel coordinates, [pixel] |
| $(m_1, n_1), (m_2, n_2)$ | initial and final moving-object positions, [pixel] |
| $n$ | frame number, [frame] |
| $M$, $N$ | dimensions of an image, [pixel] |
| $N$ | number of image frames, [frame] |
| $\mathrm{OTF}(k_x, k_y)$ | optical transfer function |
| $p$, $q$ | diffraction orders; also, integers |
| $\mathbf{p} = (x, y)$ | point on the 2-D spatial plane, [m] |
| $\mathbf{p}_s$, $\mathbf{p}_r$ | source and reference position, [m] |
| $P$ | number of moving objects in an image sequence |
| $Q = \{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_T\}$ | set of $T$ frequency domain observation points |
| $r(x, y)$, $r(m, n)$ | continuous and discrete reference images (spatial domain) |
| $r = \sqrt{m^2 + n^2}$ | radial pixel coordinate, [pixel] |
| $r_n = \sqrt{x_n^2 + y_n^2}$ | object displacement from the origin at frame $n$, [m] |
| $\mathbf{r} = (x, y, z)$ | position vector, [m] |
| $R(k_x, k_y)$ | reference spectrum (frequency domain) |
| $s_j(\mathbf{k})$, $j = 1, 2, \ldots, 6$ | sinusoidal modulation images |
| $s(x, y)$, $s(m, n)$ | continuous and discrete source images (spatial domain) |
| $s(x, y, n)$ | spatiotemporal-domain representation of a moving object |
| $S(k_x, k_y)$ | source spectrum (frequency domain) |
| $S_{\max}$ | source complex-amplitude at the point of maximum intensity |
| $S(k_x, k_y, n)$ | mixed-domain representation of a moving object |
| $S(k_x, k_y, \omega)$ | frequency-domain representation of a moving object |

| | |
|---|---|
| $S'(k_x, k_y, \omega)$ | same as above calculated from the real part of $S(k_x, k_y, n)$ |
| $\mathrm{SNR}_{\mathrm{Re}\,S}$, $\mathrm{SNR}_{\mathrm{Im}\,S}$ | real and imaginary parts of the JTI signal-to-noise ratio |
| $t$ | time, [s] |
| $t(m,n)$ | component image, no displacement |
| $T$ | number of frequency-domain observation points |
| $u(m,n)$ | component image, displaced horizontally |
| $u(n)$ | temporal components of data vector $\mathbf{u}$ |
| $\mathbf{u}$, $\mathbf{u}_t$ | data vector, data vector measured at $\mathbf{k}_t = (k_x, k_y)$ |
| $U(\mathbf{r}, t)$ | complex amplitude of an optical wave |
| $v(m,n)$ | component image, displaced horizontally and vertically |
| $v$, $\hat{v}$ | actual and estimated speeds of a moving object, [ppf] |
| $v_{max}$ | maximum object speed, [ppf] |
| $v_x$, $v_y$ | velocity components in the $x$ and $y$ directions, [ppf] |
| $(v_x, v_y)$ | Hough transform parameter space, [ppf] |
| $v_x^{min}$, $v_y^{min}$, $v_x^{max}$, $v_y^{max}$ | dimensions of Hough transform parameter space, [ppf] |
| $\mathbf{v}^{peak} = (v_x^{peak}, v_y^{peak})$ | peak coordinates in Hough parameter space, [ppf] |
| $\mathbf{v} = [v_x \; v_y]^{\mathsf{T}}$ | velocity column-vector, [ppf] |
| $\mathcal{V}$ | fringe visibility (depth of modulation) |
| $w_0$, $w_1$ | waist radius of original and expanded laser beams, [m] |
| $w_1$, $w_2$ | weighting images for real interference spectra |
| $w_3$, $w_4$ | weighting images for imaginary interference spectra |
| $w(m,n)$ | component image, displaced vertically |
| $\hat{\mathbf{w}}$ | estimate of the $K \times 1$ tap-weight filter for modified FBLP |
| $x$, $y$ | coordinates in a plane perpendicular to optic axis, [m] |
| $x_0$, $y_0$ | initial position of a moving object, [m] |
| $x_n$, $y_n$ | position of a moving object at frame $n$, [m] |
| $\mathbf{x}_k$, $\mathbf{y}_k$ | right and left singular vectors of data matrix $\mathbf{A}$ |
| $\mathbf{X}$, $\mathbf{Y}$ | $K \times K$ and $2(N-K) \times K$ unitary matrices for SVD |
| $z$ | complex variable; also coordinate along the optic axis, [m] |
| $z_1, z_2, \ldots, z_K$ | roots of a $K^{\text{th}}$-order polynomial in $z$ |

## Special Functions

| | |
|---|---|
| $\operatorname{atan}(y,x)$ | four-quadrant inverse tangent |
| $\delta(x,y)$ | Dirac-delta or impulse function |
| $\operatorname{diric}_{M,N} = \frac{\sin\left(\frac{M k_x}{2}\right)}{M \sin\left(\frac{k_x}{2}\right)} \frac{\sin\left(\frac{N k_y}{2}\right)}{N \sin\left(\frac{k_y}{2}\right)}$ | Dirichlet function of order $(M,N)$ |
| $J_0(k_r)$ | zero-order Bessel function |
| $\operatorname{rect}(x,y) = \begin{cases} 1 & \lvert x\rvert, \lvert y\rvert < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$ | rectangular function |
| $\operatorname{rect}_{M,N}(m,n) = \begin{cases} 1 & \lvert m\rvert < \frac{M-1}{2}, \lvert n\rvert < \frac{N-1}{2} \\ 0 & \text{otherwise} \end{cases}$ | rectangular sequence of length $(M,N)$ |
| $\operatorname{sinc}(k_x,k_y) = \frac{\sin(\pi k_x)}{\pi k_x} \frac{\sin(\pi k_y)}{\pi k_y}$ | filtering and interpolation function |

## Operators

| | |
|---|---|
| $z^*$ | complex conjugate |
| Re $z$, $\Re(z)$ | real part |
| Im $z$, $\Im(z)$ | imaginary part |
| $\lvert z \rvert$ | magnitude or absolute value |
| $[v_x \; v_y]^{\mathsf{T}}$ | complex-conjugate matrix transpose |
| $\mathscr{F}\{f(x,y)\}$ | continuous-space Fourier transform |
| $\mathscr{F}_{DSFT}\{f(m,n)\}$ | discrete-space Fourier transform (DSFT) |
| $\mathscr{F}\{f(t)\}$ | continuous-time Fourier transform |
| $\mathscr{F}_{DTFT}\{f(n)\}$ | discrete-time Fourier transform (DTFT) |
| $f(x,y) * g(x,y)$ | convolution of $f(x,y)$ and $g(x,y)$ |
| $f(x,y) \otimes g(x,y)$ | cross-correlation of $f(x,y)$ and $g(x,y)$ |
| $E[f(n)]$ | temporal expectation or ensemble average |

Acronyms and Abbreviations

| | |
|---|---|
| 1-D | one-dimensional |
| 2-D | two-dimensional |
| 3-D | three-dimensional |
| ATR | automatic target recognition |
| BS | beam splitter |
| CCD | charge-coupled device |
| CD | compact disc |
| CMOS | complementary metal oxide silicon |
| COG | center of gravity |
| CPU | central processing unit |
| CRT | cathode ray tube |
| DC | direct current |
| DMD | digital micromirror device |
| DOC | digital optical computer |
| FBLP | forward-backward linear prediction |
| DSFT | discrete-space Fourier transform |
| DTFT | discrete-time Fourier transform |
| DVD | digital versatile disc (or digital video disc) |
| FFT | fast Fourier transform |
| FHT | fast Hartley transform |
| FLC | ferroelectric liquid crystal |
| FPN | fixed-pattern noise |
| FT | Fourier transform |
| FTL | Fourier transform lens |
| He-Ne | Helium-Neon |
| HPOC | high-performance optoelectronic computing |
| ID | identification |
| i.i.d. | independent and identically distributed |
| JTC | joint transform correlator |
| JTI | joint transform interference |
| L | lens |
| LCTV | liquid-crystal television |
| LHS | left hand side |
| M | mirror |
| MACE | minimum average correlation energy |
| MEMS | microelectromechanical systems |
| MOSFET | metal-oxide-silicon field effect transistor |
| MPEG | moving picture experts group |
| MQW | multiple quantum well |
| MTF | modulation transfer function |

| | |
|---|---|
| MVSDF | minimum variance synthetic discriminant filter |
| N | normal |
| NTSC | national television systems committee |
| OTF | optical transfer function |
| ppf | pixels per frame |
| PC | personal computer |
| PD | photodetector |
| PDI | point-diffraction interferometer |
| PH | pinhole |
| PSF | point spread function |
| PSI | phase-shifting interferometry |
| RF | radio frequency |
| RHS | right hand side |
| ROI | region of interest |
| ROM | read only memory |
| SAR | synthetic aperture radar |
| SDF | synthetic discriminant function |
| SEM | scanning electron microscope |
| SLM | spatial light modulator |
| SNR | signal-to-noise ratio |
| SRAM | static random access memory |
| TI | Texas Instruments |
| TN-LCD | twisted-nematic liquid crystal display |
| VCSEL | vertical-cavity surface-emitting lasers |
| VLSI | very large scale integration |

# Acknowledgements

It has been a great privilege and a pleasure to work with my supervisor Michael Cada. Thank you for your excellent supervision, guidance, encouragement over the course of my studies, and for suggesting the project. Peter Gregson has contributed a great deal to my understanding of machine vision and image processing – thank you very much Peter, for sharing your unique and fruitful insights. I would also like to thank Bill Robertson and Bill Phillips for the guidance and advice offered during several helpful meetings and discussions.

I would like to thank Steven Knudsen, the inventor of the mixed-domain method of trajectory estimation, for his assistance with the implementation of his algorithms, and his quick email responses to my countless questions.

I owe a great deal to fellow graduate students Mark Starzomski and Ronnie Van Dommelen, who slogged through several painfully-early versions of this manuscript, and provided helpful feedback during the ensuing late night discussions. In addition, I would like to thank Jing-Yi Wang for his advice regarding preparation of the dissertation. My parents, Fenella and Alan Lane, mathematician and engineer, respectively, deserve a great deal of credit for reviewing an early draft of the dissertation. Thank you to my sister, Penelope Gumley, for your continued support.

I owe a substantial debt of gratitude to Robert Gray, who machined several mounting plates, under severe time constraints, for both of the optical Fourier processors. The technical staff, in particular John Lee, Christopher Hill, Mark LeBlanc, and Joe Isenor, contributed to the design, layout, and fabrication of several interface boards, and their assistance is very much appreciated.

Financial support from NSERC (Natural Sciences and Engineering Research Council of Canada), OPCOM (Optical Processing and Computing Consortium of Canada), and Dalhousie University (DalTech) is gratefully acknowledged. Additional support provided by Michael Cada and Steven Knudsen is also acknowledged.

# Abstract

This dissertation investigates interferometric optical Fourier processors for the purpose of computation. Two optical Fourier processors with coherent detection are proposed and demonstrated. Both processors contain a conventional optical Fourier processor and employ a common-path type of interferometer to indirectly measure the complex-valued optical distribution produced in the back focal-plane of the Fourier transform lens. The first employs a point-diffraction interferometer while the second uses a novel interferometric technique. Both interferometric optical Fourier processors are demonstrated in a moving-object trajectory-estimation application.

A novel interferometric technique, referred to as *joint-transform interference* (JTI), is proposed to recover the complex-valued Fourier transform of an image at selected points on the spatial frequency plane. The Fourier spectrum of a source image is interfered with the Fourier spectrum of a reference image in a common-path interferometer. Both the real and imaginary parts of the complex-valued spectrum are determined, and in addition, the source and reference images are easily matched in order to guarantee good fringe visibility. The power of the source image must be known in advance to supply the optimum reference image. Six interferograms, which were measured sequentially in this work but may also be measured in parallel, are post-processed to extract the real and imaginary parts of the Fourier spectrum of the image at a selected number of points on the spatial frequency plane. When the number of desired points is comparable to the number of pixels in the image, a digital two-dimensional (2-D) fast Fourier transform (FFT) is appropriate, however, when a relatively small number of frequency-domain points is desired, the proposed hybrid optical-digital technique can offer substantial computational savings. The number of operations required by the hybrid optical-digital Fourier processor 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 computationally efficient than a pruned 2-D FFT, especially when the input images are very large. 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 was employed to demonstrate the mixed-domain method of trajectory estimation proposed by Bruton and Knudsen. The mixed-domain technique is intended for the trajectory estimation of small, barely discernable, moving objects of unknown position and velocity. The interferometric optical Fourier processor is employed to determine the frequency-domain representation of moving objects embedded in an image sequence. The speed and direction of the objects are estimated from the sequence of Fourier spectra in the time domain using the modified forward-backward linear prediction (FBLP) method of spectral estimation. The trajectory estimation system, based on the novel joint-transform interferometric optical Fourier processor, estimated the trajectory of multiple objects moving over both stationary and white noise backgrounds. The system tracked and estimated the trajectories of $3\times3$ and single-pixel objects when the target-to-noise ratios were greater than -3 dB and 3 dB, respectively. The trajectories estimated using the optical Fourier processor compared favorably with those calculated using a digital FFT. The hybrid optical-digital calculation of the complex-valued Fourier transform at the selected frequency domain points required the measurement of six interferograms and approximately 200 floating-point operations. Calculation of the Fourier transform at the same points via the FFT required more than three orders of magnitude more floating-point operations. The number of moving objects in the image sequence was known *a priori* by the system and was not detected.

CHAPTER

# 1

# Introduction

This dissertation deals with analog optical computing and the optical Fourier transform. Two interferometric optical Fourier transform processors, one based on a traditional interferometer and the other based on a novel technique, are proposed and experimentally verified. The optical processors are demonstrated in a trajectory estimation system designed to estimate the speed and direction of multiple moving objects.

The objective of this first chapter is to provide a brief introduction to optical computing, summarize the main contributions of this dissertation, and describe the framework for their presentation. The first section describes the motivation behind optical computation and presents several examples of its application. The emphasis is on analog rather than digital optical processing. The next section summarizes the main contributions made to the field of analog image processing, specifically, interferometric optical Fourier processors. The objectives, scope and framework of the dissertation are presented in the last section.

## 1.1 OPTICAL COMPUTING

The ultimate goal of optical computing is not to replace traditional electronic computers, but rather to exploit the parallel-processing capabilities of optics to construct better computing machines. Electronic and optical computers have different architectures, modes of operation, and capabilities. They each have their own specific set of tasks which they perform most effectively. On a digital electronic computer, the numerical result of an arbitrary mathematical operation can be computed to any desired degree of accuracy. High accuracy requires repetitive processing, digitization, numerous machine cycles, and results in accurate but relatively slow computation. Optical computing is fast but less accurate. The two-dimensional (2-D) power spectrum (absolute value squared of the Fourier transform) of an arbitrarily large $M \times N$ image is calculated in less than $1$ ns on an analog optical computer (neglecting the electronic-to-optical and optical-to-electronic conversions), the time it takes an optical field to propagate across the processor. The result is calculated very quickly, literally at the speed of light, however it is much less accurate than that calculated electronically using a digital signal processor. The object of optical computing is not to replace electronics but rather to augment its abilities by employing it to do tasks that electronics perform poorly and inefficiently.

The real-time processing of synthetic aperture radar (SAR) data is generally regarded to be the first successful application of optical computing. The inventors of SAR realized early on that currently available analog and digital systems could not provide the computational horsepower required by the new technology. The optical SAR processor [Lei68, Brow69, Jen77] was the system of choice until the early 1980s when rapid advances in digital electronics made the transition to a fully digital system more attractive. The new Canadian surveillance satellite Radarsat-2, scheduled to be launched in the year 2001, will transmit raw SAR data in real-time at $105$ Mbps to ground stations for electronic post-processing

[MacD99]. The data is processed electronically by a digital computer at $1/10^{th}$ real time to produce images with a resolution comparable to those of US military satellites. Optical SAR processors, however, provide real-time imagery and consume far less power than electronic digital SAR processors.

During the 1980s and 1990s, digital SAR processors were supplanting the more traditional optical processors; more recently, there has been a renewed interest in optical technologies [Han94]. New breakthroughs in optical modulation and detection technologies now make it possible to consider satellite designs which employ the benefits of compact, power-efficient, fast optical computers for on-board, real-time SAR processing.

The fields of optical computing and optical information processing are still immature. The penetration of these technologies into mainstream consumer products for communication and computation applications has been largely transparent to the end user. The tremendous advantages of optical processors have so far been largely untapped. In stark contrast, electronic computing and information processing, based on very large scale integration (VLSI), particularly the complementary metal oxide silicon (CMOS) process, is a very mature technology which has almost reached its technical limits.

The field of optical computing is very broad and not very well defined. The most exhaustive review and taxonomy can be found in Feitelson's book, Optical Computing: A survey for computer scientists [Fei88]. There are several good review articles [Abr83, Bel86] and special issues [Jor94, Ira98] which are more technical in nature. The field is most easily characterized by dividing it into analog and digital optical computing. The two varieties have one feature in common – they both take advantage of the fundamental properties which differentiate photons from electrons.

### 1.1.1 Analog Optical Computing

Most analog optical computing is based on the so called Fourier transforming ability of a lens. It is often referred to as *optical signal processing* when the input to the optical processor is a 1-D temporal signal, and *optical image processing* when the input is a 2-D spatial signal or image. The method of phase-contrast, invented in 1934 by Fritz Zernike, and later applied to the phase-contrast microscope [Hec87, pp. 570], was one of the first and is probably the best-known example of spatial filtering in an analog optical processor. The marriage of classical physical optics with communication theory in the 1950's (giving birth to what is generally called *modern optics*) [Van74], and the invention of the laser in 1960, provided respectively, a necessary optical signal processing paradigm, and the required coherent light source for optical Fourier processing. The great success of optical computing in the fields of radar signal processing [Van92], specifically the invention of the SAR optical processor [Brow69, Jen77], and optical pattern recognition [Cas81, Cas94b], helped to encourage further research. Analog optical computers most often implement the Fourier transform, or some other linear (or nonlinear) transformation, to process their input signals.

#### *1.1.1.1 Optical Transform Processors*

The optical Fourier transform is the fundamental operation upon which the vast majority of all analog optical signal processors are based. A lens is easily employed to image an illuminated object onto a viewing screen, as in a $35\ \mathrm{mm}$ slide projector. By reducing the distance between the lens and the screen, the same lens maps the Fourier transform of the object onto the screen. The physical principle is quite simple: the far-field diffraction pattern of a coherently illuminated aperture (or slide) is equal to the Fourier transform of the aperture. A lens is employed to bring the far-field pattern into the near-field in order to make the optical processor space-efficient. A host of other integral transforms have also been implemented optically, including the Fresnel [Goo77], wavelet [Fry90],

chirp-Z [Van92, page 133], inverse Radon [Gmi80], and Mellin [Fei88, page 66] transforms.

Any space-invariant linear operation can be implemented optically. In a generic analog optical computer, two Fourier transform lenses, one each for the forward and inverse transforms, are employed to perform frequency domain spatial filtering. A space-invariant linear operation can be recast as a convolution which is easily implemented optically using a frequency domain filter. The phase-contrast microscope is simply a phase-filter located in the Fourier plane of the instrument's image path; the optical wavelet transform is generally implemented as a bank of wavelet filters. Quite complicated nonlinear transformations can also be implemented optically by cascading simple nonlinear point-wise transformations with more complex linear transformations [Häu77, Goo77]. Space-variant linear transformations can be implemented optically using an array of small micro-lenses (lenslet array) rather than a single lens [Fei88, page 103].

#### *1.1.1.2 Optical Correlators*

Optical correlation was one of the first applications of analog optical computing and remains an area of active research today. The very successful optical SAR processor is in fact a multichannel cross-correlator [Lei68]. The operation of most optical correlators is based on the convolution theorem: the product of two functions (or images) in the frequency domain is equivalent to their spatial-domain convolution. The theorem can easily be recast in terms of a cross-correlation: the frequency-domain product of a function and the complex-conjugate of a second function is equivalent to their spatial-domain cross-correlation. The latter describes the operational principle of the classical 4-$f$ optical correlator illustrated in Figure 1-1.

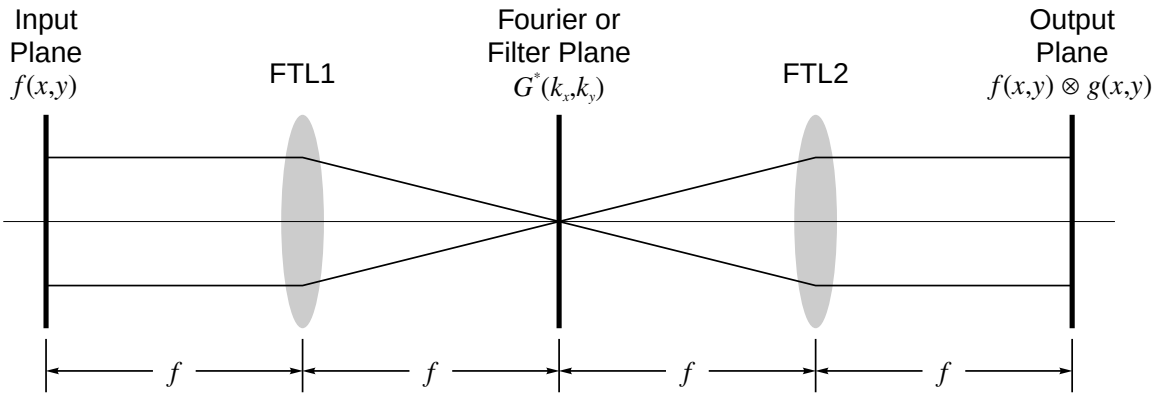



**Figure 1-1** Conventional 4-$f$ optical correlator.

Consider the problem of locating all occurrences of the capital letter E in a page of printed text. As shown in Figure 1-1, a transparency of the source image $f(x,y)$, a photographic positive of a page of text, is placed in the front focal-plane of a Fourier transform lens FTL1. The transparency is illuminated from the left with coherent light and the wavefronts passing through the transparency are modulated by the transparency. The optical distribution $f(x,y)$ in the plane of the transparency is Fourier transformed into the back focal-plane of FTL1. A transparency of a reference-image filter $G^*(k_x,k_y)$, a photographic positive of the *matched filter* of the letter E, is placed in the Fourier (also called the filter) plane, located in the back focal-plane of FTL1 (which is also the front focal-plane of FTL2). The reference-image filter is a matched filter because it is equal to the complex conjugate of the reference spectrum. The optical distribution $F(k_x,k_y)$ incident upon the second transparency is modulated by the filter to produce an optical distribution given by $F(k_x,k_y)G^*(k_x,k_y)$. This product is inverse Fourier transformed by FTL2 into its back focal-plane to provide the cross-correlation between source image $f(x,y)$ and reference image $g(x,y)$. The optical distribution mapped into the output plane is a filtered version of the input image where the transfer function is the matched filter $G^*(k_x,k_y)$. The output image is a

series of bright dots (δ-functions if you prefer) corresponding to each occurrence of the letter E in the page of text.

Two particular correlators, the VanderLugt correlator [Van64] and the joint transform correlator (JTC) [Wea66], deserve special attention. It is often very difficult to realize a general complex-valued spatial filter, particularly one which is dynamic, to use in the filter plane of real-time systems. The VanderLugt correlator employs a complex-valued spatial filter which is recorded using a holographic process (either photographic or computer generated) by encoding the amplitude and phase of the complex-valued filter on a spatial carrier. The output plane contains both the cross-correlation and the convolution of the two functions.

In a JTC (joint transform correlator), the source and target images are placed side-by-side on the same input transparency. An optically-addressed SLM (spatial light modulator) is employed in the filter plane to detect the intensity of their joint transform, and in addition, modulate a second coherent light beam which is inverse Fourier transformed to provide the cross-correlation. It is the square-law nature of the optically-addressed SLM which effectively calculates the product of the two Fourier spectra in the filter plane. The major advantage of the JTC is that a complex-valued filter does not have to be constructed – the source and reference images written to the same input transparency are Fourier transformed in parallel. The disadvantage, however, is that the space bandwidth product of the JTC is reduced because both the source and reference images are written to the same SLM. It is also difficult to implement an optically efficient correlator by employing a phase-only filter because the JTC lacks flexibility in its modes of optical modulation and filter coding.

The classical optical correlator performs matched filtering. The matched filter has been shown to be the best filter for detecting signals in stationary white noise [Tur60]. Targets and background clutter in typical images are not described well

by this model because the background (noise) is generally correlated and non-stationary. Classical correlators have trouble with *interclass discrimination*, for example, detecting the critically important difference between an E and an F, and *intraclass recognition*, ignoring the meaningless differences between a sans serif E and an serif E. The current trend is non-linear optical correlators in which the filter function is either binarized [Jav88], amplitude modulated [Fen91], or chirped [Tan93]. More recently, the joint transform has been detected and optically read out using a multiple quantum well (MQW) photorefractive SLM to introduce the non-linearity [Boo96].

Optical correlators developed in Canada have been available commercially for several years. A low cost optical correlator for industrial machine vision tasks such as pattern recognition, identification, quality control, feature extraction is available from the National Optics Institute (Sainte-Foy, Quebec). An automatic finger print identification and verification system is also available from Mytec Technologies (Don Mills, Ontario). Their system is based on optical Fourier processing and the measurement of either the power spectrum of a fingerprint [US5761330] or its Fourier hologram [US5740276].

#### *1.1.1.3 Automatic Target Recognition*

An active area of research in the field of analog optical computing is automatic target recognition (ATR) systems [Bha86]. The problem is to recognize and identify man-made artificial targets in a scene, for example, the classes of aircraft stationed on an airfield or the number and models of cars on a highway, which may appear in different sizes, orientations, and perspectives. Military and industrial applications finance and motivate a large part of the research into ATR systems. Real-time operation is essential and the systems must be small, lightweight, and rugged.

Generally, an ATR system consists of several pattern recognition and target classification stages. They often employ optical correlators for real-time pattern

recognition. The classical optical correlator reviewed in the previous section is invariant to target displacement (shift-invariant) because its operation is based on a linear matched filter, however, its performance is not invariant to target rotation, scale, and intensity distortions. Generally, a target appearing in an image scene is scaled according to its distance from the detector, and in addition, distorted due to any in-plane or out-of-plane rotation.

One approach to rotationally-invariant optical pattern recognition is the circular harmonic filter proposed by Hsu and Arsenault [Hsu82, Arsn84, Pre93]. In polar coordinates, the reference object is described by a sum of circular harmonics and a matched filter is constructed from one of the circular harmonic components to yield a rotationally-invariant filter. Another approach to both rotation and scale-invariance is the synthetic discriminant function (SDF) filter. In the SDF method, a training set of distorted reference images, which are assumed to be representative of all possible distortions, are used to construct a composite filter. The minimum variance synthetic discriminant filter (MVSDF) [Kum86] and minimum average correlation energy (MACE) filter [Mahal87] are two examples of these composite-filter approaches.

More recently, neural networks and the wavelet transform have been investigated for optical [Cas94a] and digital [Des98] pattern recognition. Bergeron *et al.* have proposed a optodigital neural network classifier, based on an optical correlator, to classify detected input objects into a digital format [Ber97]. Ahmed *et al.* have proposed a wavelet-feature-based composite reference formulation for distortion estimation and have described its implementation in a multireference synthetic JTC [Ahm98].

### 1.1.2 Digital Optical Computing

The current trend in digital optical computing is the use of free-space optical interconnects to alleviate the bottlenecks experienced in conventional electronic architectures. The advantages of photonic information carriers for free-space

optical interconnects are well known. Photons can pass through each other without any mutual interference, propagate easily through free space, and allow for massively parallel free-space optical interconnects. A matrix of several thousand optical channels can operate in parallel, each with a bandwidth much greater than that of any electronic circuit. Photons are charge neutral and are consequently immune to extraneous electromagnetic interference. Photons travel at the speed of light ($30\ \mathrm{cm \cdot ns^{-1}}$ in free space) while electrons are typically 1/3 as fast. More importantly, the velocity of light is independent of the number of interconnects branching from a node. The information velocity of an electronic transmission line is inversely proportional to its capacitance per unit length, and as such, the information velocity depends on the electronic load that must be driven.

Optoelectronic architectures which employ free-space optical interconnects for high-performance digital optical computing [Gui98] have been demonstrated by OptiComp Corporation (Lake Tahoe, Nevada). These optoelectronic processors employ an array of vertical-cavity surface-emitting lasers (VCSELs), a 2-D control mask or holographic diffractive optical element, and a detector array to perform Boolean logic functions. High fan-in and fan-out is achieved through optical interconnection while the digital logic function is achieved by summing the fields from several incident beams on a thresholding detector. The digital optical computer (DOC) II performs Boolean vector-matrix multiplication with a peak throughput of about $10^{12}$ bit operations per second [Gui91]. The high-performance optoelectronic computing (HPOC) module performs Boolean matrix-tensor multiplication at similar throughputs and may be cascaded to produce processors capable of executing complete instruction sets [Gui93].

The early success of analog optical processing laid the foundation for the field of digital optical computing. Digital optical computing, however, has received more attention in the last decade, judging by the number of papers published in the two fields. A good review of the current state of the art in digital optical

computing can be found in a special issue of *IEEE Computer* [Ira98]. Optical analog correlators can be purchased as peripherals for electronic computers attesting to the fact that analog optical processing continues to be a thriving field.

## 1.2 MAIN CONTRIBUTIONS

This dissertation deals with analog optical computing and the complex-valued optical Fourier transform. The Fourier transform (FT) is critically important to many signal processing systems. The FT is an extremely powerful tool, however, it is very computational intensive – the fastest algorithms based on the FFT (fast Fourier transform) [O&S75] often consume too many CPU clock cycles and prevent real-time processing in operational systems. It has been known for some time that optical Fourier processors calculate an FT at the speed of light. Unfortunately, only half of the FT is recoverable using conventional optical detectors: the FT has both magnitude and phase components – but only the former is easily detectable using 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 very difficult to realize a stable optical processor which provides both the magnitude and phase components required by most FT 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 can process a complex-valued $256$-point FFT in approximately $5000$ instruction cycles. A TMS320C80 running at $50\ \mathrm{MHz}$ can therefore calculate the FFT of a real-valued $256 \times 256$ pixel image in about $256 \cdot 5000 \cdot 20\ \mathrm{ns} = 25.6\ \mathrm{ms}$ which is equivalent to almost $40$ image frames per second.

The interferometric optical Fourier processor proposed in this dissertation, based on a novel technique called *joint transform interference* (JTI), is designed

to produce the complex-valued Fourier spectra of an arbitrarily large image at only a selected number of points on the spatial frequency plane. When the number of points desired is comparable to the number of pixels in the image, a digital 2-D FFT is appropriate, however, when a relatively small number of frequency-domain points are 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 $O\left(N^2\right)$. The number of operations required for the full 2-D FFT computation is $O\left(N^2 \log_2 N\right)$, and the computational complexity of a pruned 2-D FFT algorithm [Mar71] 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 dissertation 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 very 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 [Smit90] and moving object detection and trajectory estimation [Knu92a, Knu92b].

A novel interferometric technique called *joint-transform interference (JTI)* is developed in this dissertation to recover both the magnitude and phase in an optical Fourier processor. The interferometer is insensitive to vibration and temperature fluctuations due to its simple common-path design. The technique was successfully demonstrated in an optical Fourier processor and subsequently applied to the problem of moving-object trajectory estimation. A hybrid optical-digital signal processing system, based on the new optical Fourier processor,

successfully estimated the velocity and direction of multiple targets moving in a two-dimensional field.

The main contributions are presented in the following list. They are not itemized in order of importance but rather in the order as they appear in the dissertation.

1. A hybrid optical-digital trajectory estimation system based on the mixed-domain method of trajectory estimation is constructed and demonstrated using a point-diffraction based optical Fourier processor. The mixed-domain technique was developed by Knudsen and Bruton [Knu92a, Knu92b] and demonstrated by them in software. Their technique is implemented by the author in hardware using an optical Fourier processor and point-diffraction interferometer (PDI). The optical processing is performed at 1/3 video frame rates. The system is limited to the trajectory estimation of a single-pixel object of maximum contrast moving over a zero intensity background. The system correctly estimates the path and speed of the moving object, however, the direction with which the object moved along the path is indeterminate.

2. A novel interferometric technique called *joint-transform interference (JTI)* is developed to recover the complex-amplitude spectrum in an optical Fourier processor. The FT of a source image is interfered with the FT of a reference image in a common-path interferometer. Unlike the PDI, both the real and imaginary parts of the complex-valued spectrum are determined, and in addition, the source and reference images are easily matched in order to guarantee good fringe visibility. The interferometric optical Fourier processors 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 to 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.

3. A spectrum which is not affected by contrast-ratio noise is identified in the far-field diffraction pattern of the Texas Instruments digital micromirror device (DMD). Zero interleaving and spatial-domain modulation techniques, well known in the field of multirate digital signal processing, are employed to compress and shift the zero-order spectrum of the diffraction pattern. The spectrum of an image is compressed by a factor of two in the horizontal and vertical directions by inserting zeros between valid pixels in the original image. One of the replicated and compressed spectra is located in a region of the diffraction pattern which is not affected by contrast-ratio noise and which is centered on the DMD Nyquist frequency. The linear phase distribution introduced by the titled mirrors at each DMD pixel shifts the position of the diffraction envelope (region of maximum intensity) to the Nyquist frequency. The spectrum centered on the DMD Nyquist frequency has a high diffraction efficiency and is much less noisy than the spectrum at DC.

4. The JTI method of coherent detection is experimentally verified and applied to the problem of moving-object trajectory estimation. A hybrid optical-digital trajectory estimation system, based on the mixed-domain method of trajectory estimation, is constructed and demonstrated using a JTI-based interferometric optical Fourier processor. The optical processing is performed at 1/6 video frame rates. Unlike the PDI-based optical processor described earlier, the JTI based processor is able to estimate the trajectory (speed and direction over 360°) of multiple moving objects moving over both stationary and white noise backgrounds. The system is able to estimate the trajectories of $3\times3$ and single-pixel objects when the SNRs are greater than -3 dB and 3 dB, respectively.

5. Knudsen and Bruton's trajectory estimation algorithm is extended by substituting the systematic search for 3-tuples with a Hough transform technique. An ideal moving object is described in 3-D frequency space by a plane through the origin. Its trajectory is estimated by finding a set of frequency-domain triples

$(k_x, k_y, \omega)$ which lie on the plane. Knudsen and Bruton's algorithm searched a large set of frequency-domain triples for 3-tuples (sets of three frequency triples) which described the same object. In the approach described in this dissertation, the large set of frequency-domain triples is Hough transformed [Hou62, Bal82] into a velocity parameter space, $(k_x, k_y, \omega) \rightarrow (v_x, v_y)$, in order to reduce the size of the search space.

## 1.3 DISSERTATION FRAMEWORK

The objectives, approach, and a brief outline of the entire dissertation, are presented in the following three sections.

### 1.3.1 Objectives and Scope

This dissertation deals with analog optical computing. There are two primary objectives: 1) to develop an indirect interferometric means of detecting and measuring the complex-valued optical field mapped into the back focal-plane of a Fourier transform lens; and 2) to construct an optical Fourier transformer, based on the method of coherent detection, and employ it to demonstrate a trajectory estimation application using Knudsen and Bruton's mixed-domain algorithm.

The first objective requires an indirect method of detection. Optical detectors such a photographic film, photodiodes, and CCD cameras, are sensitive to the power of an incident optical signal, rather than its complex amplitude. The only way to measure the complex amplitude (i.e. magnitude and phase) of an optical distribution is to measure it indirectly by interfering the source distribution with a known reference. An interference pattern is detected and processed to recover the complex-valued information. An interferometer in which the source and reference beams follow a common path is very desirable because such a device is insensitive to mechanical shock, vibration, and temperature fluctuations. This necessitated a common-path architecture where the source and reference

beams originate from the same SLM (spatial light modulator) located in the front focal-plane of the Fourier transform lens.

The trajectory estimation application was employed to demonstrate the optical Fourier processor because it was a computationally intensive problem which had a simple frequency-domain solution. Estimating the trajectory of small, barely discernible, moving objects of unknown position and velocity is a particularly computationally intensive problem. It is a good candidate for an optical signal processor, because it requires the integration of several high-resolution image frames. The optical Fourier processor is very well suited to calculating the spatial 2-D Fourier transforms required by the Knudsen and Bruton mixed-domain algorithm. The trajectory estimation algorithm was employed only to demonstrate the optical processor. The purpose was not to investigate improvements to the trajectory estimation algorithm, propose new algorithms, or to compare the results obtained with the optical processor with those of other authors.

### 1.3.2 Approach and Methodology

The contributions presented in this dissertation were motivated by several initial experiments the author performed with a common-path type of interferometer known as the PDI (point-diffraction interferometer) [Lin33]. In these experiments, the PDI was employed to produce interferograms from which the real part of a complex-valued source distribution was extracted. The visibility of the interferograms depended on how well the power of the source distribution was matched with that of the PDI-generated reference distribution. The proposed approach was to extend the functionality of the PDI such that it would be possible to: 1) extract both the real and imaginary parts of the complex-valued source distribution; and 2) maximize the visibility of the interferograms by dynamically updating the total optical power of the reference distribution.

A new interferometric technique was developed and later employed in an optical Fourier processor to detect the real and imaginary parts of a complex-

valued distribution in the back focal-plane of a lens. The interferometric optical Fourier processor was developed using linear system theory and Fourier optics. It functions as a conventional optical Fourier transformer with the addition of coherent detection provided by the novel common-path interferometer. The interferometric optical Fourier processor was first simulated in software and then experimentally verified in hardware. Finally, the optical processor was demonstrated in a moving-object trajectory estimation system.

The optical hardware and experimental methods employed in this dissertation are simple and straight forward. All experiments involve a traditional optical Fourier processor and an interferometric (homodyne) detector. An SLM (spatial light modulator) is employed in the front focal-plane of a FTL (Fourier transform lens) to modulate the amplitude of incident plane waves. A CCD camera is used in the back focal-plane of the FTL to detect the intensity distribution. A Helium-Neon (He-Ne) laser is employed to illuminate the SLM. Source and reference images are combined and written to the SLM and their interferogram is measured by the CCD. The interferograms are electronically post-processed to provide real and imaginary spectra using algorithms developed in this dissertation. The algorithms used in the trajectory estimation system are based on those developed by Knudsen and Bruton [Knu92a, Knu92b, Knu92c].

Two different SLMs, one based on mature liquid-crystal technology, and the other based on microelectromechanical systems (MEMS) technology, were employed during the experimental work. A twisted-nematic liquid crystal display (TN-LCD) from Seiko-Epson was employed in the initial experiments while a digital micromirror device (DMD) from Texas Instruments was employed in the latter experiments. The TN-LCD is a low contrast ($20{:}1$ [Mor86]) and low speed ($30$ frames per second) SLM. Conversely, the DMD is a high-contrast ($250{:}1$ [TI96]) and high speed ($60 \times 24 = 1440$ frames per second, $1$ bit per pixel) SLM.

### 1.3.3 Organization and Overview

A concise review of the optical image processing theories and concepts employed in this dissertation is presented in Chapter 2. The chapter reviews the propagation of light through free-space using Fourier optics, the Fourier transforming property of a lens, and several interferometers which have been employed to measure the complex amplitude of an optical field. The optical Fourier transform and the point-diffraction method are married in the last section to construct an optical Fourier transformer with coherent detection.

The construction and experimental verification of a point-diffraction based optical Fourier processor for trajectory estimation is presented in Chapter 3. The frequency and mixed-domain trajectory estimation methods, an interpretation unique to the optical implementation, and the accuracy of the optical approach, are reviewed in the first sections of the chapter. The velocity components of 13 different objects, each moving on its own trajectory, are successfully estimated in two sets of experiments. The point-diffraction based optical Fourier processor, the post-processing algorithms, and the experimental results are also presented.

A new method of coherent detection, called joint transform interference (JTI), which extends and improves upon the point-diffraction method, is presented in Chapter 4. The main ideas appropriate for implementation on either a gray scale or binary SLM are presented in the first section followed by a detailed description for implementation on a binary SLM. Five classes of JTI are identified and the algorithms required to determine the real and imaginary parts of an arbitrary source image FT are presented. The effect of reference image symmetry is discussed and an example of the JTI technique is illustrated. The functional form of the interferogram modulation terms, as well as the conditions required for optimal source spectra SNR, are derived and presented in the final section.

A multiple-object trajectory estimation system, based on the JTI method of coherent detection, is presented in Chapter 5. The pixel structure, switching time,

and far-field diffraction pattern of the digital micromirror device (DMD) employed as the SLM are presented in the first section. The weighted three-component JTI method, which was developed in the previous chapter, is experimentally verified. The remainder of the chapter is devoted to the application of the JTI method of coherent detection to the problem of trajectory estimation. A single-pixel reference image is used to estimate the trajectories of single and multiple moving point-objects. A multipixel Bessel-function reference image is used to estimate the trajectory of multiple $3\times3$ pixel objects moving over a zero-intensity background, different synthetic backgrounds, and in the presence of Gaussian white noise. The trajectories of multiple single-pixel objects in white noise are estimated in the final section.

A brief summary of the contributions made in this dissertation and several suggestions for further research are made in Chapter 6.

CHAPTER

# 2

# Optical Signal and Image Processing

The purpose of this chapter is to provide a concise review of the optical image processing theories and concepts employed in this dissertation. It is not intended as a comprehensive review of the field, but rather, as a lucid account of the concepts specific to this dissertation. The propagation of light through free-space, and the linear systems approach to propagation known as Fourier optics, are reviewed in the first half of the chapter. These ideas lead to the optical Fourier transform, and the so called Fourier transforming property of a lens. The second half of the review describes several techniques of coherent optical detection which have been employed to measure the complex amplitude of an optical field. The optical Fourier transform and the point-diffraction method of coherent detection are married in the last section to construct an interferometric optical Fourier transformer. The review of optical signal processing presented in this chapter is by no means complete. More complete discussions can be found in the references cited in the appropriate sections.

## 2.1 INTRODUCTION

Optical spectrum analysis was one of the first recognized applications of optical signal processing. Acousto-optical spectrum analyzers have been employed in (synthetic aperture) radar applications and are still employed in radio astronomy. Acousto-optical spectrometers for real-time spectral analysis in millimeter radio astronomy are simple and inexpensive devices. They typically provide 2 GHz of instantaneous bandwidth at a resolution of 125 kHz [H&H89, page 1037]. Real-time digital spectrometers based on the Fast Fourier transform (FFT) are more expensive and generally more complicated. Digital spectrometers with an instantaneous bandwidth of 320 MHz and 3 bits of resolution appear to be practical (their bandwidth is generally limited by the sampling rate of the analog-to-digital converter) [Pay89].

Optical spectrum analyzers can be divided into two classes based on the variable of integration. The *space-integrating* class performs a Fourier transform with respect to one or more spatial variables, while the *time-integrating* class performs a Fourier transform with respect to a single temporal variable. Generally, the integration in the former is provided by a lens which collects a spatial distribution of light. In the latter, the integration is usually provided by a photodetector which 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 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 (after the Greek word *holos*, meaning the whole, referring to a recording of both amplitude and phase) was probably the first application of coherent optical detection. In 1967, King *et al.*

reported one of the first heterodyning time-integrating spectrum analyzers for the measurement of both amplitude and phase [Kin67]. 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.* 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 [Car69]. In a time-integrating interferometric spectrum analyzer proposed by VanderLugt, two acousto-optic modulators and a spatially modulated reference beam were employed to reduced the frequency of the temporal fringe pattern [Van81]. Turpin reported one of the first instances of coherent detection in a space-integrating spectrum analyzer [Turp81]. In his spectrum analyzer, a spatial fringe pattern was detected in the focal plane of a Fourier transform lens using a high resolution photodetector array.

An excellent review of heterodyne spectrum analysis based on acousto-optic modulators can be found in Optical Signal Processing [Van92]. A review of different interferometric Fourier transform processors (space and time integrating) designed for processing complex-valued SAR (synthetic aperture radar) data can be found in [Ale90]. Space-integrating 2-D interferometric Fourier processors have not generally been pursued or addressed in the literature due to the lack of fast, high-resolution, high contrast 2-D SLMs. Time-integrating 1-D interferometric Fourier processors are generally based on the more mature technology offered by acousto-optic SLMs.

The motivation behind interferometric optical Fourier processors has been either increased dynamic range [Van81] (the interference term is proportional to signal amplitude rather than its intensity) or complex-valued optical Fourier processing [Ale90, Lan96b]. The latter is the motivation here – this dissertation deals with the space-integrating type of interferometric Fourier processor for the purpose of computation.

This chapter is presented as follows. The propagation of light in free-space and the well known Fourier transforming property of a lens are reviewed in Section 2.2. A novel model which describes the Fourier spectrum of a pixelated SLM is presented in Section 2.3. Coherent detection and the operation of several traditional interferometers, including the Mach-Zehnder, Twyman-Green, and point-diffraction interferometers, are reviewed in Section 2.4. The optical Fourier transformer and point-diffraction interferometer are married in the last section to construct an interferometric optical Fourier processor.

## 2.2 FOURIER OPTICS

Fourier optics is a branch of physical optics which provides a description of the propagation of light through an optical system using harmonic analysis and linear system theory. It is particularly well suited to individuals in Electrical Engineering (and other disciplines) who have a good understanding of linear system theory and the basic principles of the Fourier transform. The standard text on the subject is Introduction to Fourier Optics, by J.W. Goodman [Goo68]. A good primer can be found in Chapter 4 of Fundamentals of Photonics, by B.E.A. Saleh and M.C. Teich [S&T91].

Fourier analysis permits one to represent an arbitrary 2-D wavefunction as the sum of several harmonic functions of different spatial frequency and complex amplitude. The function $f(x,y)$, for example, can be written as a superposition of harmonic functions $\exp\left[-i2\pi\left(\nu_x x+\nu_y y\right)\right]$, where $\left(\nu_x,\nu_y\right)$ are the spatial frequency components of the harmonic function (cycles per unit length), and the complex amplitude of each function is determined from the Fourier transform of $f(x,y)$.

The plane wave is the simplest form of a three-dimensional wave. It can be used to analyze a travelling wave of arbitrary complexity, as shown in Figure 2-1. The optical disturbance of a plane wave, at a given instant in time, can be described by a sequence of equally-spaced plane surfaces of constant phase

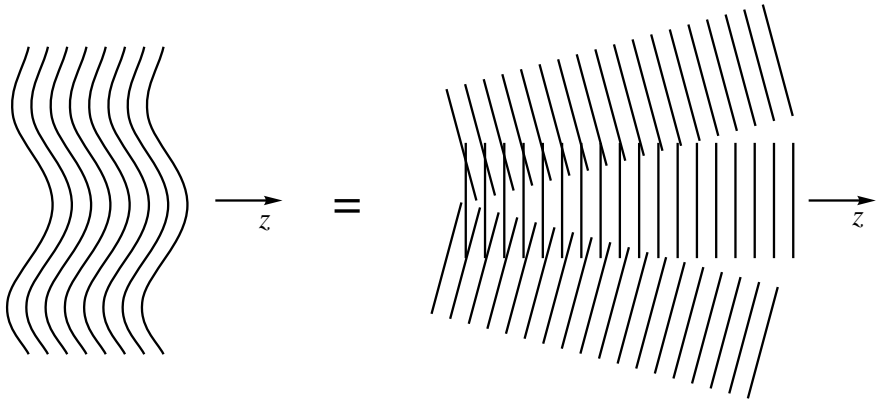



**Figure 2-1** An arbitrary wave shown as a superposition of plane waves.

called wavefronts. The wavefronts propagate through space in a direction parallel to the wavefront normal given by its wavevector $\mathbf{k} = (k_x, k_y, k_z)$. A plane wave has a spatial angular frequency equal to its wave number $k = \sqrt{k_x^2 + k_y^2 + k_z^2} = \frac{2\pi}{\lambda}$ and its complex amplitude is written $\exp[-i(k_x x + k_y y + k_z z)]$. In an arbitrary plane, $z = 0$ for example, the complex amplitude of the plane wave is identical to the 2-D harmonic function $\exp[-i2\pi(\nu_x x + \nu_y y)]$ if the first two components of the wave vector, $k_x = 2\pi\nu_x$ and $k_y = 2\pi\nu_y$, are interpreted as spatial angular frequencies (radians per unit length). An arbitrary function $f(x, y)$ can therefore be analyzed as a superposition of plane waves,

$$f(x,y) = (2\pi)^{-2} \iint F(k_x, k_y) \exp[-i(k_x x + k_y y)]\, dk_x\, dk_y \qquad (2\text{-}1)$$

where the complex amplitudes of the plane waves are determined by,

$$F(k_x, k_y) = \iint f(x,y) \exp[i(k_x x + k_y y)]\, dx\, dy \qquad (2\text{-}2)$$

This pair of equations should be recognized as the spatial 2-D Fourier transform pair defined in Equation 1-2 (see Section 1.4.2). The plane wave is the spatial building block used to synthesize 2-D wavefronts of arbitrary complexity in the same way that the sinusoid is used to synthesize 1-D temporal signals.

### 2.2.1 Propagation of Light in Free Space

Consider a plane wave with complex amplitude $\exp\left[-i\left(k_x x + k_y y + k_z z\right)\right]$ and wavevector $\mathbf{k} = \left(k_x, k_y, k_z\right)$. The wavevector makes angles $\theta_x = \sin^{-1}\left(\frac{k_x}{k}\right)$ and $\theta_y = \sin^{-1}\left(\frac{k_y}{k}\right)$ with the *y*-*z* and *x*-*z* planes, respectively. The complex amplitude in the $z = 0$ plane is a harmonic function $\exp\left[-i2\pi\left(\nu_x x + \nu_y y\right)\right]$ with spatial frequency components $\left(\nu_x, \nu_y\right)$. The angles of the wavevector are therefore related to the spatial frequencies of the harmonic function by,

$$\begin{aligned} \theta_x &= \sin^{-1}\left(\lambda \nu_x\right) \\ \theta_y &= \sin^{-1}\left(\lambda \nu_y\right) \end{aligned} \tag{2-3}$$

If the plane wave propagates almost parallel to the *z* axis, $k_x << k$ and $k_y << k$, and the wavevector is said to be paraxial. Consequently, the angles of inclination are small ($\sin\theta_x \approx \theta_x$ and $\sin\theta_y \approx \theta_y$) and they can be approximated by,

$$\begin{aligned} \theta_x &\approx \lambda \nu_x \\ \theta_y &\approx \lambda \nu_y \end{aligned} \tag{2-4}$$

in the paraxial approximation. A plane wave, propagating almost parallel to the *z* axis, is described by a wavevector whose angles of inclination are directly proportional to the spatial frequency components of the corresponding harmonic function in the *x*-*y* plane.

A plane wave of wavelength $\lambda$ is shown in Figure 2-2 incident upon a thin optical element whose amplitude transmittance is a harmonic function of spatial frequency $\nu_x$ (period $\Lambda_x = \nu_x^{-1}$). The plane wave is modulated by the harmonic function, and consequently, it is redirected through an angle $\theta_x = \sin^{-1}\left(\lambda \nu_x\right)$. The incident and transmitted waves are required to have the same wavelength; the transmitted wavefronts must match the periodic pattern of the transmittance

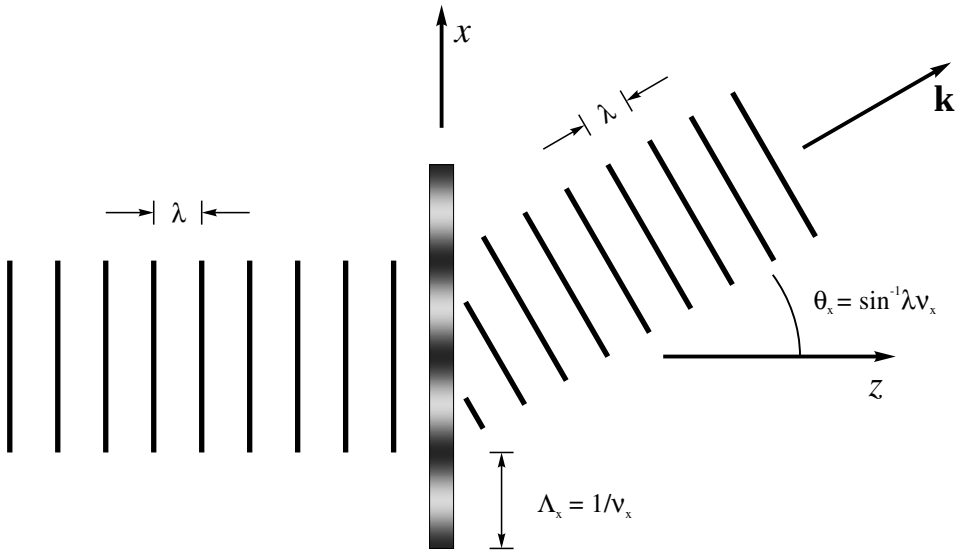



**Figure 2-2** Plane waves incident on a thin optical element whose amplitude transmittance is a harmonic function with spatial period $\Lambda_x = \nu_x^{-1}$.

function and this constraint causes the redirection. If the transmittance of the optical element consists of a sum of several harmonic functions of different spatial frequencies, the transmitted optical wave is also the sum of an equal number of plane waves directed into different directions. Each spatial frequency component is mapped into a corresponding direction and the amplitude of each wave is proportional to the amplitude of the corresponding harmonic component of the transmittance function.

The process of spatial spectral-analysis, as illustrated in Figure 2-2, is very similar to the temporal spectral-analysis, or angular dispersion, provided by a prism or diffraction grating. The component plane waves diffracted by an arbitrary element are spatially separated as they propagate through space. Free-space propagation appears to serve as a natural *spatial spectrum-analyzer* or *spatial prism*, sensitive to the spatial content of an optical wave.

### 2.2.2 Optical Fourier Transform

The propagation of light through free-space is conveniently described by Fourier analysis. Consider a paraxial wave of wavelength $\lambda$, whose complex amplitude at $z = 0$ is a function $f(x, y)$ composed of harmonic components of different spatial frequency. Each harmonic component corresponds to a different paraxial plane wave: the plane wave corresponding to spatial frequency component $(\nu_x, \nu_y)$ propagates at an angle $(\lambda\nu_x, \lambda\nu_y)$ and has a complex amplitude proportional to its harmonic component. The plane waves separate naturally as they propagate to produce a far-field complex-amplitude distribution proportional to the Fourier transform of $f(x, y)$.

The preceding discussion suggests that the Fourier transform of an arbitrary function $f(x, y)$ could be calculated by illuminating a transparency designed to have $f(x, y)$ as its amplitude transmittance function, and then simply measuring its far-field diffraction pattern. A more practical and space-efficient approach is to map the far-field diffraction pattern into the back focal-plane of a lens. The lens maps the angular spectrum produced by the transparency into a spatial spectrum in its back focal-plane. An optical Fourier transformer based on this principle is illustrated in Figure 2-3. A transparency of amplitude transmittance $f(x, y)$ is placed a distance $d$ in front of a Fourier transform lens (FTL). The optical field observed in the back focal-plane $g(x, y)$ is proportional the Fourier transform of $f(x, y)$. Plane waves inclined at an angle $(\theta_x, \theta_y)$ to the $z$ axis are mapped into the back focal-plane (or Fourier plane) of the lens at position $(\theta_x f, \theta_y f)$, where $f$ is the focal length of the lens.

An expression for the complex amplitude of the field in the Fourier plane is readily calculated using Fourier optics. The procedure is presented in most standard texts on optical signal processing including Goo68, pp. 77-90, S&T91,

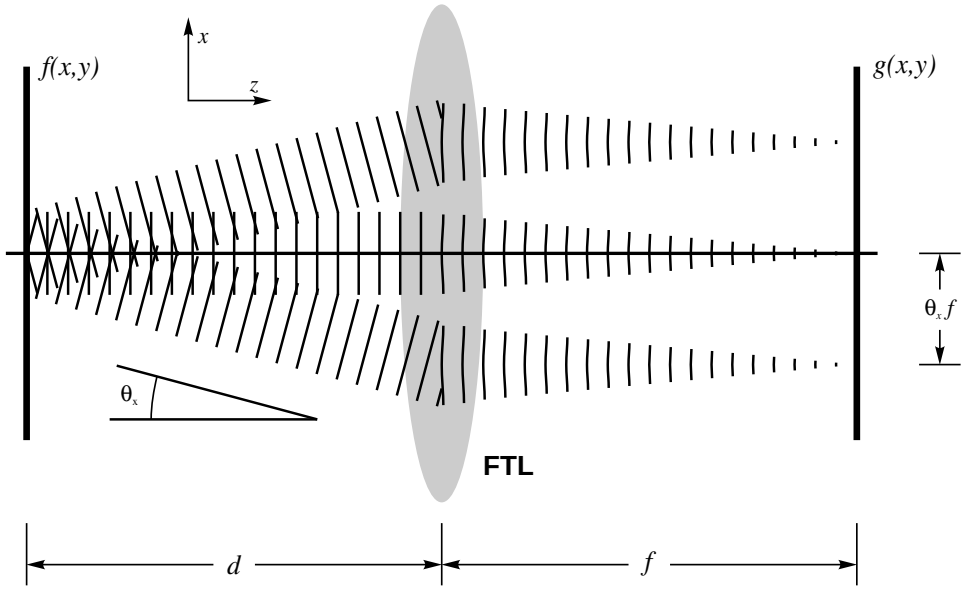



**Figure 2-3** Principle of the optical Fourier transform. The amplitude distribution $f(x,y)$ in the back focal-plane of the lens (FTL) is Fourier transformed into the front focal-plane $g(x,y)$.

pp. 124-126, Ban91, pp. 122-133, Van92, pp. 97-100, and Yu92, pp. 20-23. The expression is given here without proof,

$$g(x,y) = \frac{i\exp(-ik(d+f))}{\lambda f}\exp\left[\frac{i\pi(x^2+y^2)(d-f)}{\lambda f^2}\right]F\left(2\pi\frac{x}{\lambda f}, 2\pi\frac{y}{\lambda f}\right) \quad (2\text{-}5)$$

where $F(k_x,k_y)$ is the spatial 2-D Fourier transform of $f(x,y)$ defined in Equation A-2 (see Appendix A). If the transparency is placed in the front focal-plane of the lens, then $d=f$, and the quadratic phase term which pre-multiplies $F(k_x,k_y)$ disappears and the field is given by,

$$g(x,y) = \frac{i\exp(-i2kf)}{\lambda f}F\left(2\pi\frac{x}{\lambda f}, 2\pi\frac{y}{\lambda f}\right) \quad (2\text{-}6)$$

The optical field in the back focal-plane of the lens is typically measured with a square-law detector such as a CCD camera. Square-law detectors are sensitive

to incident optical flux or optical intensity rather than the complex amplitude of the field. The detector measures a quantity proportional to the square of the absolute value of the complex amplitude,

$$|g(x,y)|^2 \quad = \quad (\lambda f)^{-2}\left|F\left(2\pi\frac{x}{\lambda f}, 2\pi\frac{y}{\lambda f}\right)\right|^2 \quad = \quad (\lambda f)^{-2}\left|F\left(k_x, k_y\right)\right|^2 \qquad (2\text{-}7)$$

where $k_x = \frac{k\,x}{f}$, $k_x = \frac{k\,y}{f}$, and $k = \frac{2\pi}{\lambda}$. In general, the complex amplitude of the field observed in the back focal plane $g(x,y)$ is a function of distance $d$ (see Equation 2-5), however, the field intensity measured by a square-law detector is independent of $d$ even when $d \neq f$. The phase of the optical distribution in the back focal-plane of the FTL is very sensitive to the difference $d - f$. Coherent detection by processing a sequence of interferograms is relatively insensitive to the difference because the power spectra $|g(x,y)|^2$ detected by a square-law detector are not sensitive to the $d - f$ difference.

The function whose Fourier transform is required is written to some kind of spatial light modulator (SLM) in the front focal-plane of the Fourier transform lens. An SLM is any device that is able to modulate the complex amplitude (typically either the amplitude or the phase), of an optical field propagating through it, at addressable points or pixels on its cross-section. For example, a transparency on an overhead projector or a $35\text{mm}$ slide are common examples of static SLMs. Practical optical Fourier processors employ dynamic SLMs which may be addressed either electrically or optically. The experimental work in this dissertation employs two different electrically addressable SLMs: 1) a twisted-nematic liquid crystal display (TN-LCD); and 2) a digital micromirror device (DMD).

## 2.3 FOURIER TRANSFORM OF A PIXELATED SLM AS A DSFT

An optical Fourier transformer maps a continuous amplitude distribution from the front focal-plane of its lens into a continuous amplitude distribution in its back focal-plane. The continuous-space optical Fourier transform is appropriate to model the spectrum of an SLM which modulates the amplitude distribution in a continuous fashion. Electrically addressable SLMs are usually pixelated devices. The discrete amplitude distribution produced by a pixelated SLM is Fourier transformed into a continuous and periodic amplitude distribution. The pixelated structure is often modeled by multiplying the spatial domain function $f(x,y)$ written to the SLM by a 2-D *bed-of-nails* sampling function or a rectangular grating function based on the SLM pixel structure [Rey89, Gia92, Gia93b, Dav90, Kno95]. This spatial domain sampling leads to *spectrum replication* by virtue of the frequency domain convolution theorem. The spectrum of the function written to the SLM, $F(k_x,k_y)$, is replicated into each order $(p,q)$ of the Fraunhofer diffraction pattern. The total diffraction pattern is the sum of all replicated spectra $\sum_{p,q} F(k_x+2\pi p, k_y+2\pi q)$. The total spectrum observed in any diffraction order is equal to $F(k_x,k_y)$ plus the contributions aliased from the tails of spectra diffracted into all other orders. The aliased contributions are often ignored because describing the total spectrum as an infinite sum of individual spectra is mathematically cumbersome.

### 2.3.1 Spectrum of an SLM

An optical Fourier transformer which employs a pixelated input SLM produces the continuous FT of a discrete function. It is convenient to represent the pixelated (discrete) image written to an $M \times N$ pixel SLM as an infinite complex sequence $f = \{ f(m,n) \}$, where $m = 1, 2, \ldots, M$, $n = 1, 2, \ldots, N$, and all other elements are set to zero. The complex amplitude at SLM pixel $(m,n)$ is

represented by $f(m,n)$ which is also typically quantized in amplitude. When the SLM is illuminated with coherent plane waves, it produces a discrete amplitude distribution in continuous space. The amplitude distribution can be expressed as the sum of modulated and shifted rectangular functions. An SLM pixel $(m,n)$ can be represented in continuous *x*-*y* space as $f(m,n)\,\mathrm{rect}\left(\frac{x-m}{a},\frac{y-n}{b}\right)$, where $a=\frac{\Delta'_x}{\Delta_x}$ and $b=\frac{\Delta'_y}{\Delta_y}$ are the horizontal and vertical spatial duty cycles of the SLM (ratio of pixel-dimension to pixel-pitch). The SLM pixels are assumed to consist of a rectangular active area surrounded by an opaque dead zone. The complex amplitude across the active area is uniform and the opaque dead zone blocks all incident light. The physical pixel pitch (center-to-center spacing) is $\Delta_x \times \Delta_y$ and the active area of each pixel has dimensions $\Delta'_x \times \Delta'_y$. The SLM fill-factor (sometimes called the aperture ratio) is then given by $ab=\frac{\Delta'_x\Delta'_y}{\Delta_x\Delta_y}\times 100\%$. The amplitude distribution produced by the SLM can be written as a sum of modulated and shifted pixel functions, or as the convolution of the pixel function with a sum of modulated and shifted δ-functions,

$$\begin{aligned} f(x,y) &= \sum_m \sum_n f(m,n)\,\mathrm{rect}\left(\tfrac{x-m}{a},\tfrac{y-n}{b}\right) \\ &= \mathrm{rect}\left(\tfrac{x}{a},\tfrac{y}{b}\right) * \sum_m \sum_n f(m,n)\,\delta(x-m,y-n) \end{aligned} \tag{2-8}$$

The pixel aperture function $\mathrm{rect}\left(\frac{x}{a},\frac{y}{b}\right)$ represents the aperture of a single pixel with unit pixel pitch. The spectrum of the image sequence on the SLM is determined by calculating the continuous spatial FT of Equation 2-8. Using the convolution theorem, the FT can be expressed as the product of the FT of the two terms in the convolution, specifically, $ab\,\mathrm{sinc}\left(a\frac{k_x}{2\pi},b\frac{k_y}{2\pi}\right)$ and,

$$\begin{aligned}
&\mathcal{F}\left\{\sum_m\sum_n f(m,n)\,\delta(x-m,y-n)\right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_m\sum_n f(m,n)\,\delta(x-m,y-n)\exp\left[i\left(k_x x+k_y y\right)\right]dx\,dy \\
&= \sum_m\sum_n f(m,n)\exp\left[i\left(k_x m+k_y n\right)\right] \\
&= \mathcal{F}_{DSFT}\{f(m,n)\}
\end{aligned} \tag{2-9}$$

where $k_x$ and $k_y$ represent the normalized angular frequencies in radians per pixel. The (non-normalized) spatial frequencies are $k_x' = \frac{k_x}{\Delta_x} = \frac{2\pi x}{\lambda f}$ and $k_y' = \frac{k_y}{\Delta_y} = \frac{2\pi y}{\lambda f}$, where $f$ is the focal length of the Fourier transform lens, and $\lambda$ is the wavelength of light. The expression on the second last line of Equation 2-9 represents the discrete space Fourier transform (DSFT) of the 2-D sequence $f(m,n)$. The FT of the SLM-displayed sequence $f(m,n)$ is equal to its DSFT multiplied by the FT of the SLM pixel aperture function,

$$F\left(k_x,k_y\right) \;=\; ab\,\mathrm{sinc}\left(a\tfrac{k_x}{2\pi},b\tfrac{k_y}{2\pi}\right)\;\mathcal{F}_{DSFT}\{f(m,n)\} \tag{2-10}$$

The spectrum is described by the product of a modulation envelope and a periodic interference function. The modulation envelope is equal to the (coherent) transfer function of the SLM and the periodic interference function is equal to the DSFT of the 2-D sequence.

The modulation envelope describes the spatial frequency response of the SLM. It is proportional to $ab$ which suggests that an SLM with a large fill-factor is required for an optically efficient Fourier transformer. The modulation envelope $\mathrm{sinc}\left(a\frac{k_x}{2\pi},b\frac{k_y}{2\pi}\right)$ is called the (coherent) optical transfer function (OTF) of the SLM [Inf93, Goo68] (the coherent and incoherent OTFs are identical in this special case because the autocorrelation of the coherent OTF is equal to itself). The DSFT of a sequence is $2\pi$-periodic in $\left(k_x,k_y\right)$. Its properties follow directly from the continuous Fourier transform and are well documented [O&S75, pp. 30-34].

The DSFT has been borrowed from the field of discrete-time signal processing, where it is known as the discrete-time Fourier transform (DTFT). The DSFT represents a transformation from discrete space $(m,n)$ to continuous spatial frequency $(k_x,k_y)$.

The absolute value of the OTF of an SLM with spatial duty cycle $a=b=0.8$ and the periodic nature of the DSFT are illustrated in Figure 2-4. The zero-order spectrum is the shaded region at the origin of frequency space bounded by $-\pi \le k_x,k_y \le \pi$. The $(p,q)^{\text{th}}$ order spectrum is centered on $(k_x,k_y)=(2\pi p,2\pi q)$. The OTF acts as a passive frequency domain filter; it passes the zero-order spectrum (the pass band is not flat however) and attenuates higher-order spectra. The first zeros of the OTF occur at $k_x=\pm\frac{2\pi}{a}$ and $k_y=\pm\frac{2\pi}{b}$. They are closest to the origin when $a=b=1$ (100% fill-factor) and retreat from the origin as the fill-factor decreases. When the fill-factor equals 100%, the product of the transfer function and the DSFT equals the continuous FT of the sequence

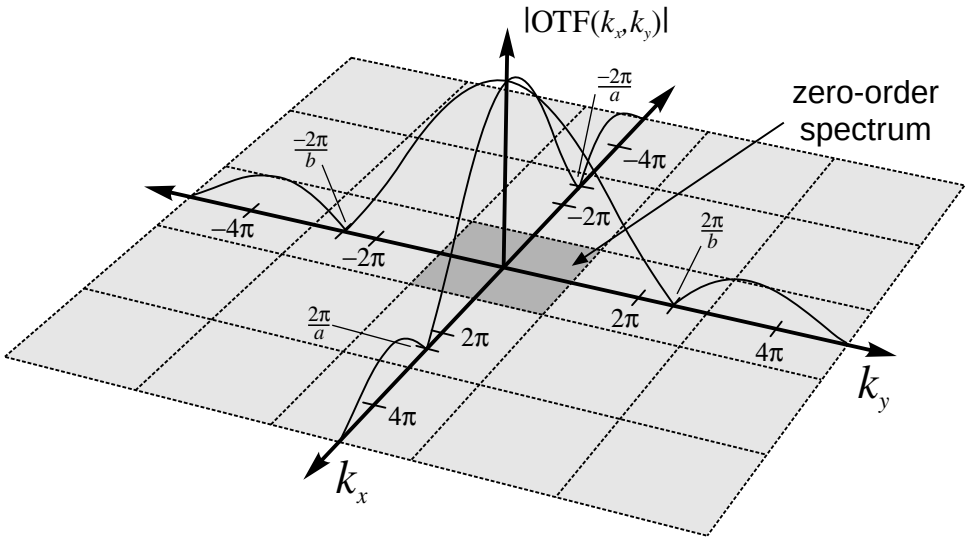



**Figure 2-4** Spectrum of a pixelated SLM. The OTF is determined by the pixel aperture and the periodicity is determined by the pixel pitch ($a=b=0.8$).

(elements of the sequence are represented by δ-functions in the continuous FT). In the limit as the fill-factor approaches $0\%$ ($a = b = 0$), the pixel aperture approaches a δ-function, and the OTF approaches unity for all frequencies (the efficiency also approaches zero because very little light is transmitted or reflected by the SLM).

### 2.3.2 DSFT of a 1-D Aperture

As an example, consider the spectrum of an $N$-pixel rectangular aperture displayed in the center of an SLM. The aperture is represented by a discrete rectangular sequence centered at the origin. The centered $\mathrm{rect}$-sequence is defined as

$$\mathrm{rect}_N(n) \quad = \quad \begin{cases} 1 & |n| \le \frac{N-1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2-11}$$

where its width $N$ must be odd. The DSFT of the $\mathrm{rect}$ sequence is presented in most references on discrete-time signal processing [O&S75] and is easily calculated using the DSFT definition from Appendix A (see Equation A-4),

$$\mathscr{F}_{DSFT}\{\mathrm{rect}_N(n)\} \quad = \quad N\frac{\sin\left(\frac{Nk}{2}\right)}{N\sin\left(\frac{k}{2}\right)} \quad = \quad N\,\mathrm{diric}_N(k) \tag{2-12}$$

The ratio of sinusoids in Equation 2-12 is called the Dirichlet function of order $N$, $\mathrm{diric}_N(k)$, to be analogous with $\mathrm{sinc}\left(\frac{k}{2\pi}\right)$, which is obtained in the continuous space FT analogue [Eva93] (the $\mathrm{diric}_N$ function is sometimes referred to as the Dirichlet kernel, or the periodic or aliased $\mathrm{sinc}$ function). The Dirichlet function, $\mathrm{diric}_N(k)$, has unit value at the origin and is $2\pi$ periodic for $N$ odd.

Assume a centered $\mathrm{rect}$-sequence of length $N = 9$ is written to a 1-D pixelated SLM with spatial duty cycle $a = 0.8$. The spectrum observed in the back focal-plane of a Fourier transform lens under coherent illumination is proportional to

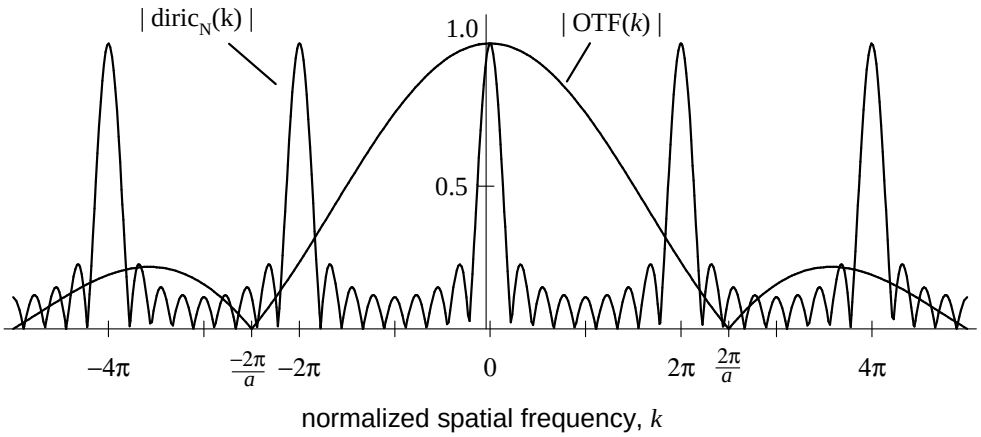



**Figure 2-5** Discrete space Fourier transform (DSFT) of a rect-sequence. The SLM spatial duty cycle (pixel width-to-pitch ratio) is $a = 0.8$ and the sequence is length $N = 9$.

the product of $N\,\mathrm{diric}_N(k)$ and the SLM transfer function $a\,\mathrm{sinc}\left(a\frac{k}{2\pi}\right)$. The normalized magnitude of the Dirichlet function, $\left|\mathrm{diric}_N(k)\right|$, and the absolute value of the OTF, $\left|\mathrm{sinc}\left(a\frac{k}{2\pi}\right)\right|$, are plotted in Figure 2-5. The Dirichlet function is $2\pi$ periodic and has $N-2=7$ *secondary maxima,* and $N-1=8$ *minima,* between each *principle maxima*.

## 2.4 COHERENT DETECTION

Optical detectors such as photographic film, photodiodes, photo-multiplier tubes, CCD cameras, and even the human eye, are sensitive to the power of an incident optical signal rather than its complex amplitude (these optical detectors are actually responsive to the incident photon flux rather than the incident power, however the distinction is not important here). Optical detectors produce an output signal which is proportional to the absolute value squared of the incident complex field, and for this reason, they are referred to as square-law detectors.

A square-law detector in the back focal-plane of an FTL measures the power spectrum of the optical field distribution in its front focal-plane. The phase of the complex valued Fourier transform is lost in the detection operation because the

square-law detector is sensitive to incident optical power. The complex amplitude of an optical field must be measured *indirectly*, by mixing it with a coherent reference of stable phase, and detecting the superposition with a square-law detector. In the field of optical communications, the reference wave is generated by a *local oscillator* and the method of coherent detection is referred to as *optical heterodyning*, *optical mixing*, or *light beating*. It is the optical equivalent to the superheterodyne receiver used in RF (radio frequency) communication systems.

In other (non-communication) applications where a local oscillator is not required, the complex amplitude of an optical field is measured by mixing it with a coherent reference in an interferometer. The interference is typically achieved by dividing a single coherent beam of light into two separate beams (by either division of wavefront or division of amplitude), modulating or processing one beam independently of the other, and then recombining the two beams in the plane of a square-law detector. The square-law detector is critically important because the optical fields must be multiplied rather than simply added.

The process of recording a hologram is a familiar example of coherent detection using an interferometer. Wavefronts from an object of interest (object beam) are interfered with a reference beam of known amplitude and phase in the plane of a photographic plate. The interference pattern recorded on the film contains both the amplitude and phase information of the object beam. The object beam can be reconstructed by illuminating the photographic plate with a beam identical to the reference called the readout beam.

The architecture and operational principles of three traditional interferometers, which have been used for the coherent detection of an optical signal, are reviewed in the following sections. The fundamental principle common to all three interferometers is the optical mixing or interference of two arbitrary waves.

### 2.4.1 Interference of Two Waves

The superposition of two optical fields present simultaneously in the same region of space is simply the sum of the individual fields. The intensity of their superposition is not necessarily equal to the sum of their intensities and the difference is due to the interference between the two waves.

Consider the interference of two monochromatic waves $U_1(\mathbf{r},t)$ and $U_2(\mathbf{r},t)$ where $\mathbf{r}=(x,y,z)$ is position and $t$ is time. If the fields exist simultaneously in time and space, their superposition is simply, $U(\mathbf{r},t)=U_1(\mathbf{r},t)+U_2(\mathbf{r},t)$. The intensity of the superposition is proportional to the square of the absolute value of its complex wavefunction,

$$\begin{aligned} |U|^2 &= |U_1|^2 + |U_2|^2 + U_1U_2^* + U_1^*U_2 \\ &= |U_1|^2 + |U_2|^2 + 2\Re(U_1U_2^*) \end{aligned} \tag{2-13}$$

where $\Re$ is the real-part operator and the spatial and temporal dependencies have been dropped for clarity. Suppose the waves have intensities $I_1$ and $I_2$, and phases $\varphi_1$ and $\varphi_2$, then their complex amplitudes can be written $U_1=\sqrt{I_1}\exp(i\varphi_1)$ and $U_2=\sqrt{I_2}\exp(i\varphi_2)$. Substitution of these expressions into Equation 2-13 yields,

$$I = I_1 + I_2 + 2\sqrt{I_1I_2}\cos(\varphi) \tag{2-14}$$

where $\varphi=\varphi_2-\varphi_1$ is the phase difference between the two waves. The relation described by Equation 2-14 is called the *interference equation*. Obviously the intensity of the superposition of the two waves is not equal to the sum of their individual intensities. The third term of the interference equation, called the *interference term*, accounts for the interference between the two waves, and may be positive or negative, corresponding to either *constructive* or *destructive* interference. When the intensity of both waves are equal, $I_1=I_2=I_0$, the total

intensity is given by $4I_0 \cos^2\left(\frac{\varphi}{2}\right)$. The total intensity is $4I_0$ when the two waves are exactly in phase, ie $\varphi = 0$ (constructive interference), and zero when they are exactly out of phase, ie $\varphi = \pm\pi$ (destructive interference). When the phase difference is a random variable, uniformly distributed between -$\pi$ and $\pi$ (incoherent or partially coherent light), or exactly $\varphi = \pm\frac{\pi}{2}$, the total intensity is the sum of the individual intensities because the interference term is zero.

The interference equation (Equation 2-14) predicts a sinusoidal intensity variation with phase for the interference of two plane waves. The interference pattern for two waves of unequal intensity is illustrated in Figure 2-6. The contrast, or modulation depth, of the interference pattern is described by its fringe visibility which is defined as,

$$\mathcal{V} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad (2\text{-}15)$$

where $I_{max}$ and $I_{min}$ are the maximum and minimum values of intensity for all phase shifts , $-\pi \le \varphi \le \pi$. Since $\cos(\varphi)$ varies between 1 and -1, the visibility of

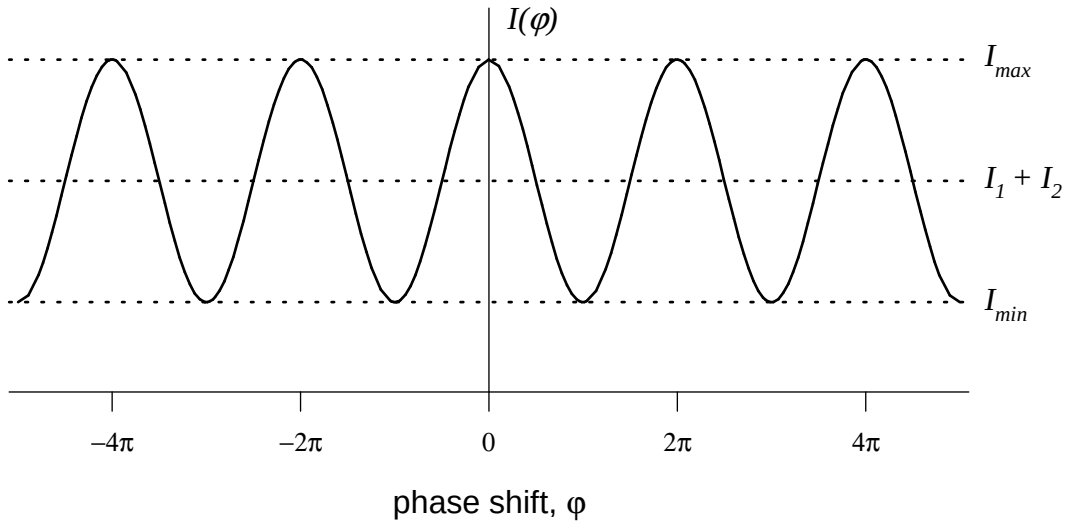



**Figure 2-6** Intensity distribution versus phase for the interference of two plane waves of unequal intensity.

the interference pattern described by Equation 2-14 is given by,

$$\mathcal{V} = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2} \tag{2-16}$$

When the intensities of the two interfering beams are equal, $I_1 = I_2 = I_0$, the visibility is maximum ($\mathcal{V} = 1$, $I_{max} = 4I_0$, and $I_{min} = 0$), and the interference pattern has maximum contrast.

### 2.4.2 Mach-Zehnder Interferometer

The Mach-Zehnder interferometer [Hec87, p.358], illustrated in Figure 2-7, is an example of a mirrored amplitude-splitting type of interferometer. It consists of two beam splitters and two totally reflecting mirrors. An incident wave $U_0$ from an extended source is split by the first beam splitter (BS1) into two waves, $U_1$ and $U_2$, of approximately equal intensity. The two waves traverse the interferometer along different physical paths: wave $U_1$ in the lower path is phase (and possibly amplitude) modulated by some kind of spatial light modulator (SLM) which modifies its optical path length (and reduces its amplitude); wave $U_2$ in the upper

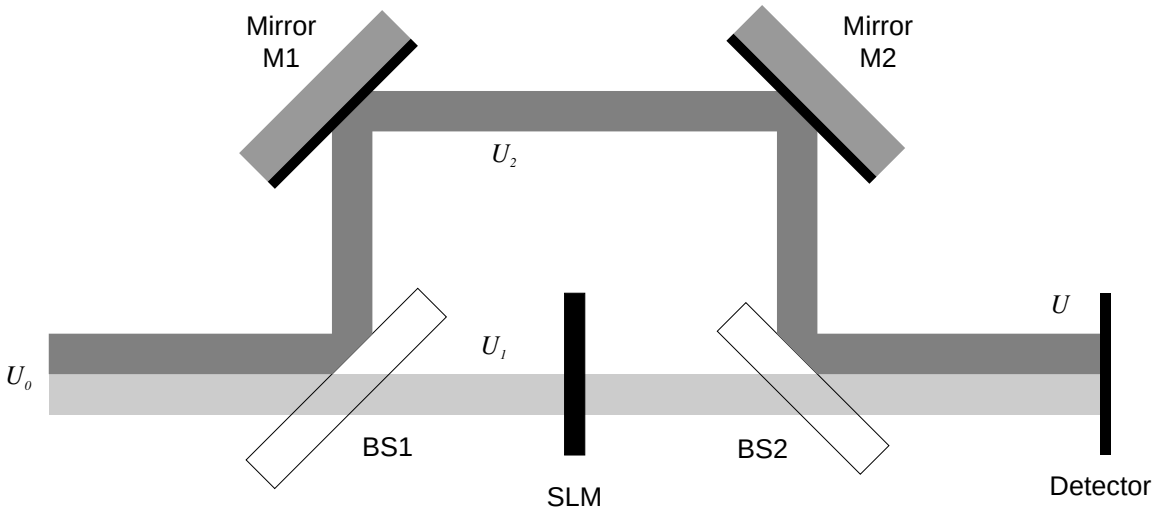



**Figure 2-7** Mach-Zehnder interferometer.

path is not modulated and acts as a reference . The two waves are recombined at the second beam splitter (BS2) to produce wave $U = U_1 + U_2$. The interference pattern of the superposition is recorded by a square-law detector.

The Mach-Zehnder interferometer's two-path design has both advantages and disadvantages. The spatial separation of the two paths makes the instrument applicable to a myriad of experiments. For example, a common application is to observe the density variations in gas-flow patterns within a wind tunnel. The primary disadvantage of the two-path design is its extreme sensitivity to mechanical shock, vibration, and temperature fluctuations. The Mach-Zehnder interferometer's extreme sensitivity to mechanical shock makes it very suitable for optical instruments designed to measure vibrations and very small (on the order of optical wavelengths) displacements. The interferometer is also relatively difficult to align.

### 2.4.3 Twyman-Green Interferometer

The Twyman-Green interferometer is a slight modification of the well known Michelson interferometer. Both instruments have a similar configuration, however their methods of illumination are different: plane wave illumination (typically from a laser) is employed in the Twyman-Green interferometer; the Michelson interferometer is illuminated by an extended source (a diffusing ground-glass plate illuminated by a discharge lamp).

The Twyman-Green interferometer [Hec87, pp. 385-386], also a mirrored amplitude-splitting device, is illustrated in Figure 2-8. An incident plane wave $U_0$ is split by the beam splitter (BS) into two waves, $U_1$ and $U_2$, of approximately equal intensity. Reflected wave $U_1$ propagates left to mirror M1 were it is reflected back along its original path and transmitted by the beam splitter. Transmitted wave $U_2$ propagates up to mirror M2 where it is reflected back along its path and subsequently reflected by the beam splitter. The superposition of the

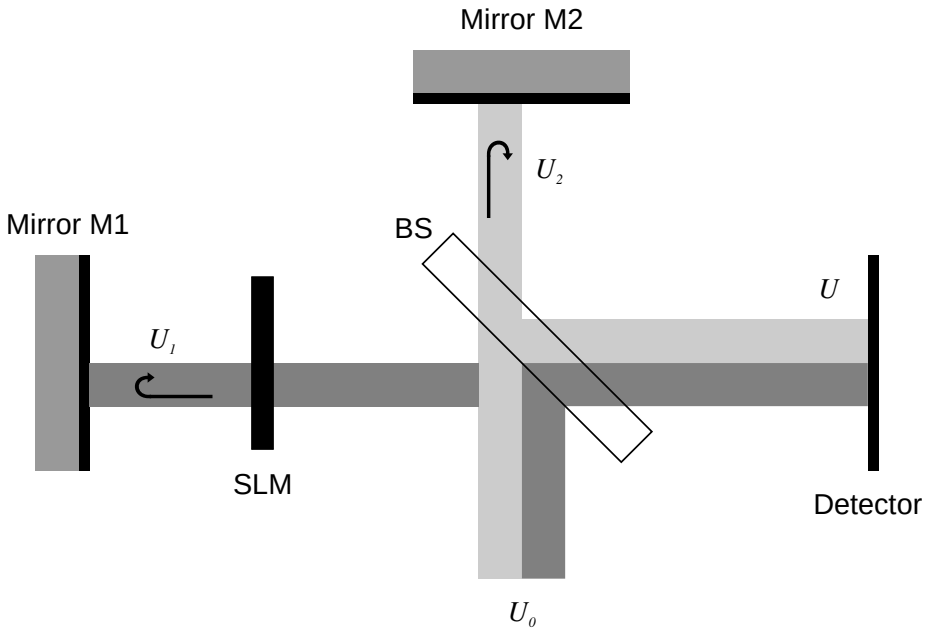



**Figure 2-8** Twyman-Green interferometer.

two waves produce an interference pattern in the plane of the detector. The phase (and possibly the amplitude) of wave $U_1$ is modified by a spatial light modulator in the first arm (the wave is actually modulated twice); wave $U_2$ in the second arm is not modulated and acts as a reference. The Twyman-Green interferometer is used extensively in modern optical testing to measure the quality of lenses, prisms, and other optical components. Like all two-path interferometers, it is sensitive to mechanical shock, vibration, and temperature fluctuations, however, it is relatively easy to align.

### 2.4.4 Point-Diffraction Interferometer

The point-diffraction interferometer (PDI) is a simple yet extremely powerful wavefront splitting common-path interferometer. It was originally invented [Lin33] by Linnik in 1933 and subsequently reinvented [Sma72, Sma74] and patented by Smartt and Strong in 1972 (the original paper by Linnik was published in Russian however an English translation is included in a paper by Speer [Spe79]).

The PDI is much less sensitive to mechanical shock, vibration and temperature fluctuations than the traditional Mach-Zehnder, Twyman-Green and Michelson interferometers due to its simple common-path design. In a common path interferometer, the object and reference waves follow similar paths through the same optical elements. The waves of the more traditional interferometers follow widely separated paths, and consequently, each beam is affected differently by shocks and fluctuations in temperature. The fringe pattern of a traditional interferometer is often very unstable and its measurement is difficult. The common path design also reduces the number and required quality of the optical elements thereby reducing the cost, size, and weight of the device, and simplifying its alignment.

The operational principle of the point-diffraction interferometer is illustrated in Figure 2-9. The wave of interest, $U_1$, typically a distorted plane wave generated by some external process, is focused onto a PDI mask in the back focal-plane of converging lens L1. The PDI mask consists of an absorbing film fixed to a transparent substrate. A small pinhole or small opaque disc is placed at the center of the mask to provide a discontinuity. Wavefronts incident on the mask are transmitted through the absorbing film with reduced amplitude and, in addition, diffracted by the discontinuity at the center of the mask. The transmitted

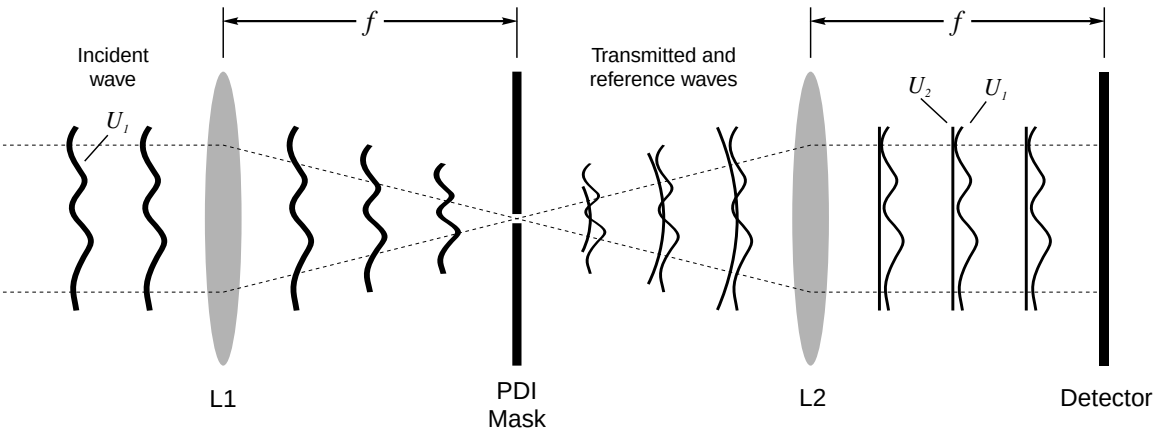



**Figure 2-9** Point-diffraction interferometer (PDI).

and diffracted wavefronts produce an interference pattern in the back focal-plane of the second lens L2. In essence, the reference wave $U_2$ is generated from the object wave $U_1$ by the PDI mask. The spherical waves diffracted by the discontinuity are transformed into plane waves by lens L2; the incident waves are imaged onto the detector through both lens. The superposition of both produces an interference pattern in the plane of the detector. The attenuation of the absorbing film on the PDI mask must be chosen carefully in order to maximize the visibility of the interference pattern. When the amplitude of the transmitted and diffracted waves are equal the fringe visibility is maximum.

The diameter of the pinhole in the PDI mask determines both the total optical power contained in the diffraction pattern and its spatial distribution. The central maximum of the diffraction pattern produced by the pinhole is an Airy disk with a diameter which is inversely proportional that of the pinhole. The total power and the maximum (on-axis) intensity of the diffraction pattern are proportional to, respectively, the area of the pinhole and the area of the pinhole squared. There is obviously a tradeoff between spatial extent and total optical power: a small pinhole is required for a broad Airy disk with an approximately uniform intensity distribution, however, a large pinhole is required for an optical efficiency interferometer with good fringe visibility. As shown in Figure 2-9, the pinhole is placed in the back focal-plane of a converging lens in order to maximize the light intensity incident on the pinhole and allow a smaller pinhole to be employed.

The PDI was originally designed for fluid flow diagnostics, however it has been used more extensively in the measurement of wavefront distortion for testing lenses and other optical elements. The methods of phase-shifting interferometry (PSI) are the most accurate and effective techniques to measure the distortion of an object wave. It has been difficult to combine the advantages of the PDI and the advanced technique of PSI because it is difficult to shift the phase of one beam with respect to the other in a common path design. Several liquid crystal based interferometers have been demonstrated which combine the advantages

of both. Mercer [Mer94, Mer96] used a liquid crystal layer to introduce an arbitrary phase shift and a microsphere embedded within the liquid crystal layer provided a locally generated reference. The phase of the object beam is shifted by changing the voltage applied across the nematic liquid crystal which alters the refractive index of the material. In a similar design, Kadono [Kad94] used a circular region etched in the liquid crystal electrode to prevent rotation of the liquid crystal molecules.

The PDI has also been applied to SLM phase-compensation and coherent detection in optical signal processing applications. The PDI was employed by Tam to compensate for the phase distortion introduced by a liquid crystal television (LCTV) [Tam90]. The interference pattern measured by the PDI was written back onto the LCTV in a closed-loop to compensate for the phase distortion. Zhang *et al.* have used the PDI to detect the complex amplitude of the optical wavelet transform of 1-D and 2-D input functions [Zha93a, Zha93b, Zha94]. In their experiments, a photographic mask containing a PDI reference pattern and a (complex) wavelet filter were placed in the filter plane of a modified VanderLugt correlator. The interference pattern at the output plane of the correlator contained a term proportional to the bipolar-valued wavelet transform.

## 2.5 PDI-BASED INTERFEROMETRIC OPTICAL FOURIER PROCESSOR

It is very desirable to detect the complex amplitude of the field in the back focal-plane of an optical Fourier transformer. The point diffraction interferometer (PDI) is easily incorporated into an optical Fourier transformer to enable its complex amplitude to be detected. The Fourier transformer with coherent detection is designed to measure the complex amplitude of the FT produced by a spatial light modulator (SLM). It represents a marriage of the optical Fourier processor introduced in Section 2.2.2 and the PDI discussed in Section 2.4.4.

The architecture of the combined optical Fourier processor and PDI is illustrated in Figure 2-10. It is very similar to the basic PDI configuration except:

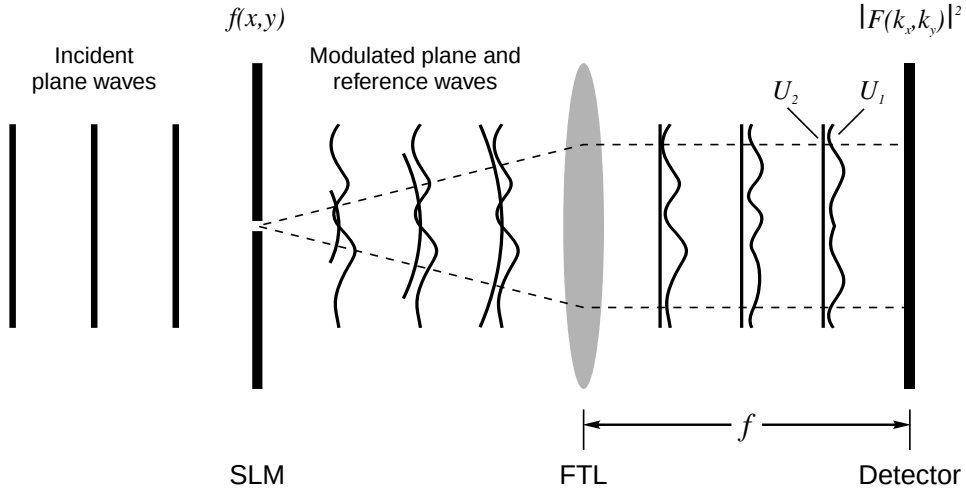



**Figure 2-10** Optical Fourier transformer with coherent detection.

1) the PDI mask of the interferometer is replaced by an SLM; and 2) the PDI mask (SLM) is illuminated by plane waves rather than converging spherical waves. A single pixel at the center of the SLM provides the discontinuity instead of the pinhole aperture in an absorbing film. Plane wavefronts incident on the SLM are modulated according to the function written to it and, in addition, diffracted by the single-pixel discontinuity at the center of the SLM. In essence, the SLM produces both source and reference waves: the modulated waves are Fourier transformed into the back focal-plane of the Fourier transform lens (FTL) to produce $U_1$; the (almost) spherical waves diffracted by the single-pixel discontinuity are Fourier transformed into the back focal-plane of the FTL to produce (almost) plane waves $U_2$. The superposition $U = U_1 + U_2$ produces an interference pattern in the plane of a square-law detector.

The real part of the Fourier transform of the function written to the SLM can be extracted from the interference pattern. Let $s(x, y)$ be the 2-D distribution written to the SLM and let $\mathrm{rect}\left(\frac{x}{\Delta'_x}, \frac{y}{\Delta'_y}\right)$ be the single-pixel reference function at its origin where $(\Delta'_x, \Delta'_y)$ represents the dimensions of the pixel. The reference pixel at the origin of the SLM is reserved for reference generation, and consequently, the

value at the origin of the function $s(x,y)$ must be zero. The amplitude distribution in the plane of the SLM is given by the sum of the source and reference functions,

$$f(x,y) = s(x,y) + \mathrm{rect}\left(\frac{x}{\Delta_x},\frac{y}{\Delta_y}\right) \tag{2-17}$$

In the limit as the pixel approaches a pinhole aperture, the pixel dimensions approach zero, and the rect-function approaches a δ-function. The reference pixel can never become an ideal point reference; its amplitude is limited by the incident light intensity and the passive SLM. The intensity transmitted by the reference pixel should ideally be matched to the total intensity transmitted by the source image. If the reference pixel and source image intensities can not be matched, the fringe visibility of the interference pattern is poor and there is limited dynamic range available to a detector.

Assuming the reference pixel can be represented as a δ-function, the power spectrum observed by a square-law detector in the back-focal plane of the FTL is given by,

$$\left|F(k_x,k_y)\right|^2 = 1 + \left|S(k_x,k_y)\right|^2 + 2\Re\left(S(k_x,k_y)\right) \tag{2-18}$$

and contains an interference term which is proportional to the real part of the Fourier transform $S(k_x,k_y)$. The interference term can be extracted by subtracting the power spectrum of $s(x,y)$, and that of the reference, from the joint power spectrum of both,

$$\Re\left(S(k_x,k_y)\right) = \frac{1}{2}\left(\left|F(k_x,k_y)\right|^2 - \left|S(k_x,k_y)\right|^2 - 1\right) \tag{2-19}$$

The optical Fourier processor with single-pixel PDI and the algorithm described by Equation 2-19 will be used in Chapter 3 to calculate the real part of the Fourier transform of a function written to an SLM.

## 2.6 CHAPTER SUMMARY

This chapter reviewed the optical signal and image processing concepts that will be employed in this dissertation. The propagation of light through free-space, and the linear systems approach to propagation known as Fourier optics, were reviewed. These ideas led to the optical Fourier transform, and the so called Fourier transforming property of a lens. It was shown that the spectrum (or far-field diffraction pattern) of an image displayed on a pixelated SLM (spatial light modulator) is equal to the DSFT (discrete space Fourier transform) of the image multiplied by the OTF (optical transfer function) of the SLM.

The second half of the chapter reviewed several methods of coherent optical detection. It was learned that an indirect method of detection must be employed because optical detectors are insensitive to the phase an optical field. The architecture and operational principles of the Mach-Zehnder, Twyman-Green (Michelson), and point-diffraction interferometers were briefly reviewed. It was discovered that the point-diffraction interferometer, due to its simple common-path design, is much less sensitive to mechanical shock, vibration and temperature fluctuations, than the other two interferometers. The optical Fourier transform and the point-diffraction method of coherent detection were married in the last section to construct an optical Fourier processor with coherent detection.

CHAPTER

# 3

# Single-Object Trajectory Estimation

This chapter describes the construction and experimental verification of a moving object trajectory estimation system. The mixed-domain algorithm for trajectory estimation, developed by Knudsen and Bruton, is implemented in hardware to estimate the trajectory (speed and direction) of point-objects moving in a sequence of images. The 2-D spatial FT (Fourier transform) required by the algorithm is calculated using an optical Fourier processor and coherently detected using a point diffraction interferometer. The sequence of optically processed spectra is post-processed electronically to estimate the trajectory of the moving objects. The velocity components of 13 identical point-objects, each moving on their own trajectory, are successfully estimated in two sets of experiments.

## 3.1 INTRODUCTION

We take for granted the ability to discern objects, estimate their motion, and navigate in three-dimensional spaces using our biological vision system. Duplicating these capabilities in a machine vision system, even to a limited degree, has proved to be a very difficult task [Agg88]. The computational power required for even a basic vision system seems to transcend the capabilities of traditional electronic processors. This has motivated the present hybrid optical-digital approach which exploits and combines the parallel-processing capabilities of optical processing and the accuracy of electronic digital processing.

The specific problem addressed in this dissertation is the detection and trajectory estimation of small, barely discernible, moving objects of unknown position and velocity. An optical implementation is of particular computational interest as typically more than ten high resolution image frames must be integrated to detect and track targets imbedded in a noisy background. 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.

Porat and Friedlander have identified two general approaches to trajectory estimation: *detect-before-track* and *track-before-detect* [Por90]. In the first approach, frame-by-frame differencing followed by thresholding is typically used to detect and segment the moving objects. The 3-D image sequence is then projected onto a single 2-D image called a track or streak map. Object trajectories can be identified from the streak map [Cow83] using a Hough transform [Bal82, Hou62].

The track-before-detect approach represents an exhaustive search for all candidate trajectories per pixel per frame. The decision as to whether a target

exists on a particular trajectory is deferred until the candidate trajectory has been observed for a predetermined number of frames. The approach can be computationally prohibitive, depending on the number of candidate trajectories, the image resolution, and the observation interval.

The matched filtering operation of an exhaustive search can be implemented in the spatiotemporal domain, the frequency domain, or the mixed domain. In terms of geometry, the spatiotemporal-domain approach represents a search for all possible line segments in space-time, while the frequency domain approach represents a search for all possible planes through the origin of spatiotemporal frequency space.

Spatiotemporal-domain approaches often use statistical tests to reduce the number of candidate trajectories. Bruton and Bartley used 3-D recursive filters, based on the concept of network resonance, to enhance and track moving objects [Bru85, Bru86]. Mohanty detects objects by searching for a trajectory which maximizes a likelihood ratio computed from an estimation of the background noise statistics [Moh81]. Barniv proposed a dynamic programming solution as a substitute for the prohibitive exhaustive search [Bar85]. Blostein and Huang used sequential-hypothesis testing to reduce the size of the search space [Blo91].

Frequency-domain approaches can take advantage of the computational efficiency of the FFT (fast Fourier transform). Reed *et al.* derived an expression for the optimum frequency-domain 3-D matched filter in terms of the optical and detector transfer functions and the spectral densities of the background and detector noise [Ree83]. Mahmoud *et al.* proposed a projection-based algorithm which uses either the FFT [Mah88] or the fast Hartley transform (FHT) [Mah91] to determine the trajectory of a large moving object. Porat and Friedlander proposed a bank of 3-D frequency-domain filters for all possible candidate trajectories [Por90]. Choi *et al.* demonstrated a time-recursive (Kalman) filter in

parallel with the bank of 3-D frequency-domain filters to solve the problem of spectral aliasing due to very fast moving objects [Cho97].

The mixed-domain approach to trajectory estimation has several advantages over a purely spatiotemporal or frequency-domain approach. The theory of mixed-domain signal processing [Knu91, Knu92c] has recently been formalized by Knudsen and Bruton. They have successfully applied their mixed-domain signal-processing algorithms to the trajectory estimation of linear [Knu92a, Knu93a] and nonlinear targets [Knu92b]. Linear targets move at constant velocity and describe a line in space-time parameterized by a velocity and initial position; nonlinear targets move at constant acceleration and describe a parabola in space-time parameterized by an acceleration, initial velocity, and initial position. In the mixed-domain approach, trajectories are estimated by Fourier transforming the two spatial dimensions of the image sequence, followed by time-domain spectral estimation using a high-resolution algorithm. The trajectories estimated in the 3-D frequency-domain approach often suffer from poor accuracy and resolution due to the sampled nature of the image sequence; a large number of image frames is usually required for accurate estimates. The mixed-domain estimates are more accurate than the frequency estimates because the high-resolution spectral-estimation technique used in the mixed-domain approach is not as dependent on the duration of the image time-series.

The hybrid optical-digital trajectory estimation system discussed in this chapter extends that of Knudsen's system by calculating the 2-D spatial component of the computationally intensive FT using an optical Fourier processor rather than a conventional electronic processor. The objective of this present chapter is to demonstrate a hybrid optical-digital moving-object trajectory estimation system which implements the Knudsen and Bruton mixed-domain algorithm. The purpose of this experimental system is to illustrate the fundamental principles involved and provide proof-of-concept only – it is not a deployable trajectory estimation system. Practical systems which accommodate multiple moving

objects, background imagery, and the effects of noise will be considered later in Chapter 5. The trajectory estimation system, interferometric optical Fourier processor, and the preliminary results, have been published by the author [Lan98]; complete results will appear in a full paper which was in press at the time of writing [Lan99a].

The trajectory estimation systems presented in this thesis deal with the trajectory estimation of constant-velocity targets moving on a linear trajectory. A real-world moving object does not describe a constant-velocity linear trajectory with infinite extent in space-time. The trajectory of a real-world moving object may, however, be approximated by a sequence of short piecewise-linear trajectories. In this dissertation it is assumed that the object is moving with constant velocity during an observation interval of 25 frames during which time its velocity is estimated. A new piecewise-linear trajectory is estimated every 25 frames (more than once every second at video frame rates) and it is assumed that the velocity of the target does not change significantly during the observation interval. Any target acceleration or velocity fluctuations during the observation interval will introduce chirp and phase jitter into the sinusoidal data vectors of the mixed-domain algorithm. The chirp and phase jitter will hinder the performance of the spectral estimation method which expects pure sinusoids and will degrade the overall performance of the trajectory estimation system.

The frequency and mixed-domain approaches to trajectory estimation are reviewed in Section 3.2. The spatiotemporal, mixed, and frequency-domain representations of an ideal moving object are derived. An interpretation of the mixed-domain representation, unique to an optical implementation of the mixed-domain algorithm, and the accuracy of the algorithm are presented. The hybrid optical-digital trajectory estimation systems and its post-processing algorithm are presented in Section 3.3. The results of two experiments are presented in Section 3.4 and their accuracy and limitations are discussed in Section 3.5.

## 3.2 TRAJECTORY ESTIMATION

The frequency-domain method of trajectory estimation is well established. The principles were originally developed for a model of human visual-motion sensing based on human psychophysics [Wat85, Ade85] and have been applied more recently to machine-vision trajectory estimation. The frequency-domain approach to trajectory estimation requires a transformation from 3-D spatiotemporal space to 3-D frequency space. In the mixed-domain approach, developed by Knudsen and Bruton [Knu92a, Knu92b, Knu93a], the 3-D transformation is reduced to a cascade of two simpler transformations of reduced dimensionality. The spatial dimensions of the 3-D spatiotemporal data set are processed first using a 2-D spatial FT and the temporal dimension is processed using a high-resolution time-domain spectral-estimation technique.

### 3.2.1 Frequency-Domain Trajectory Estimation

Consider a single point-object with velocity vector $\mathbf{v} = \begin{bmatrix} v_x & v_y \end{bmatrix}^{\mathrm{T}}$, where T is the transpose operator, and initial position $(x_0, y_0)$, moving along a line in the $x$-$y$ plane. The position $(x_n, y_n)$ of the object at a particular point in time, or during a particular frame $n$, is described by a set of two parametric equations, $\{ x_n = v_x n + x_0, y_n = v_y n + y_0 \}$, where the $n$ subscript indicates the integer frame dependence on position. Each equation in the set represents a plane in 3-D space-time and their intersection describes the trajectory of the object. The trajectory is described by a 3-D function,

$$s(x, y, n) \quad = \quad \delta(x - v_x n - x_0)\,\delta(y - v_y n - y_0) \tag{3-1}$$

which represents the intersection of the two planes. The object's spatiotemporal representation is illustrated in Figure 3-1.

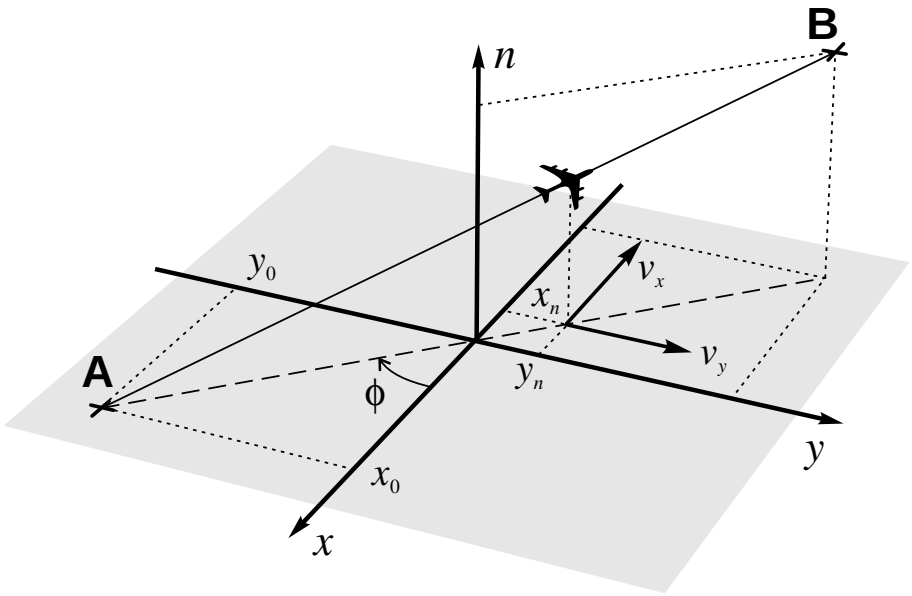



**Figure 3-1** Spatiotemporal-domain representation of a moving point-object with velocity vector $\mathbf{v} = \begin{bmatrix} v_x & v_y \end{bmatrix}^T$ and angle $\phi$.

The mixed-domain representation of the moving object is determined by Fourier transforming the 3-D spatiotemporal representation with respect to only its two spatial dimensions. Application of the continuous-space Fourier transform to Equation 3-1 yields,

$$\begin{aligned} S(k_x, k_y, n) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(x,y,n)\exp\left[i(k_x x + k_y y)\right]dx\,dy \\ &= \exp\left[i(k_x x_n + k_y y_n)\right] \\ &= \exp\left[i(\omega_t n + \phi_t)\right] \end{aligned} \tag{3-2}$$

where $\omega_t = k_x v_x + k_y v_y$ is a temporal frequency and $\phi_t = k_x x_0 + k_y y_0$ is a phase offset.

The mixed-domain representation of Equation 3-2 can be interpreted as either 1) a time-sequence of sinusoidal images; or 2) a spatial-array of 1-D sinusoidal time-sequences. At a particular moment in time, $n$ is constant, and the mixed-

domain representation is described by a complex-valued sinusoidal image (the real part of a complex-valued image sequence is illustrated in Figure 3-4 and will be discussed in the next section). This interpretation is suggested by the second line of Equation 3-2. The spatial frequency components of the sinusoidal image in the $k_x$-$k_y$ plane are determined by the object's displacement from the origin in the $x$-$y$ plane. Alternatively, at a particular point on the $k_x$-$k_y$ plane, the temporal variation in the mixed-domain representation can be described by a complex-valued 1-D sinusoidal sequence in $n$. This interpretation is suggested by the third line of Equation 3-2. The temporal frequency $\omega_t$ of the complex-valued time-sequence is determined by the speed of the moving object and its phase off-set $\phi_t$ is determined by the object's initial position.

The frequency-domain representation of the moving object is determined by Fourier transforming the mixed-domain representation with respect to its temporal dimension. Application of the discrete-space Fourier transform (DSFT) to Equation 3-2 yields,

$$\begin{aligned} S(k_x,k_y,\omega) &= \mathscr{F}_{DTFT}\{S(k_x,k_y,n)\} \\ &= \sum_{n=0}^{\infty} S(k_x,k_y,n)\exp(-i\omega n) \\ &= 2\pi\exp(i\phi_t)\delta(\omega-\omega_t) \end{aligned} \tag{3-3}$$

The frequency-domain representation is a complex-valued function of three variables; its absolute-value squared at some point $(k_x,k_y,\omega)$ describes the energy associated with that particular frequency triple. The energy of the moving object in the frequency domain is confined to a plane in 3-D frequency space defined by, $k_x v_x + k_y v_y - \omega = 0$. The plane passes through the origin and the components of its slope along the two spatial frequency axes are equal to the velocity components of the moving object. A plane is said to exist in frequency space if the energy distribution is such that points on the plane have high energy

(an ideal moving object has infinite energy as described by a δ-function) and all other points have low energy (zero energy in the ideal case). A moving object is therefore represented as a line in space-time and a plane in frequency space.

### 3.2.2 Optical Implementation

A traditional optical Fourier processor calculates the power spectrum of an input image. An interferometric optical Fourier processor calculates the real (or imaginary) part of an image's complex-valued spectrum, however, a phase-shifting interferometer is required to measure both the real and imaginary parts at the same time. The architecture of an interferometric optical Fourier processor is much simpler if its reference does not have to be phase-shifted. In this section the frequency-domain representation of a moving object is calculated assuming that only the real part of the complex-valued spatial transform is calculated by the optical processor.

Application of the discrete-space Fourier transform (DSFT) to the real part of the mixed-domain representation of Equation 3-2 yields,

$$\begin{aligned} S'(k_x, k_y, \omega) &= \mathscr{F}_{DTFT}\{\Re(S(k_x, k_y, n))\} \\ &= \cos(\phi_t)\,\mathscr{F}_{DTFT}\{\cos(\omega_t n)\} - \sin(\phi_t)\,\mathscr{F}_{DTFT}\{\sin(\omega_t n)\} \\ &= \pi\exp(-i\phi_t)\,\delta(\omega+\omega_t) \;+\; \pi\exp(i\phi_t)\,\delta(\omega-\omega_t) \end{aligned} \qquad (3\text{-}4)$$

The energy of the Fourier transform is now distributed evenly between a pair of planes in 3-D frequency space, $k_x v_x + k_y v_y \pm \omega = 0$. The two planes are illustrated in Figure 3-2. Both planes pass through the origin and one is the reflection of the other in the $k_x$-$k_y$ plane; the slope of one plane is equal to the negative slope of the other. One plane represents an object moving along a trajectory with velocity vector $\mathbf{v} = [v_x \; v_y]^T$ and the other represents an object moving along the same path but in the opposite direction, $\mathbf{v}' = [-v_x \; -v_y]^T = -\mathbf{v}$.

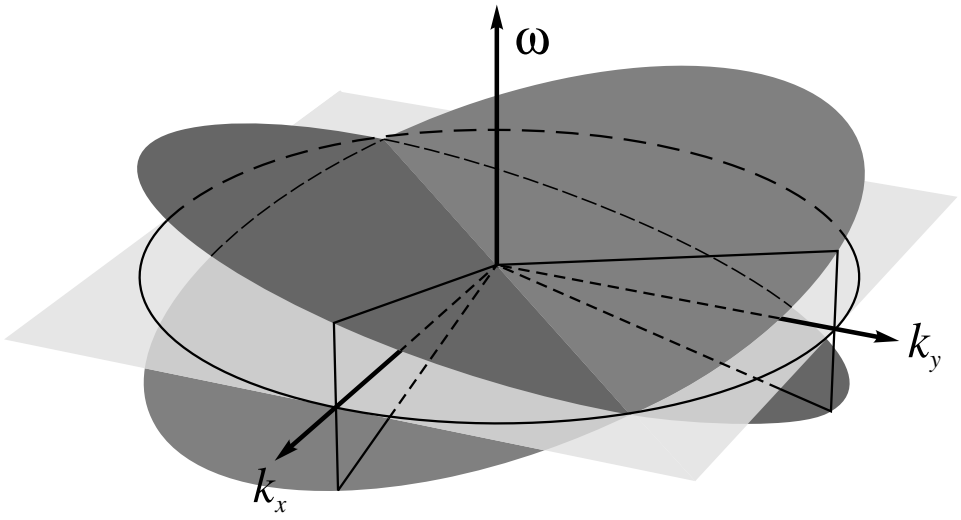



**Figure 3-2** Frequency-domain representation of a moving point-object when the temporal Fourier transform is applied to the real part of the mixed-domain representation only.

An object will therefore have two temporal frequencies $\pm\omega$ associated with each point on the $k_x$-$k_y$ plane. The two frequency-domain planes completely define the moving object's path however its direction along the path is indeterminate. It is not possible to determine the object's direction because the complex part of the spatial transform was lost.

### 3.2.3 Interpretation of the Mixed Domain

Consider two moving point-objects with the same velocity but different initial positions, as shown in Figure 3-3. One object moves along a path from point A, through points B, C, and D, to a final point E. The second object moves along a primed path at the same velocity. Both objects move with a velocity of $\sqrt{2}$ units per frame, along a path at $-45°$ to the $x$ axis. The non-primed object crosses the origin of the $x$-$y$ plane at point C, and the primed object crosses the $x$ and $y$ axes at points D' and B' respectively.

It is illustrative to consider the real part of the mixed-domain representation as a sequence of cosinusoidal images in time. These images are observable on the

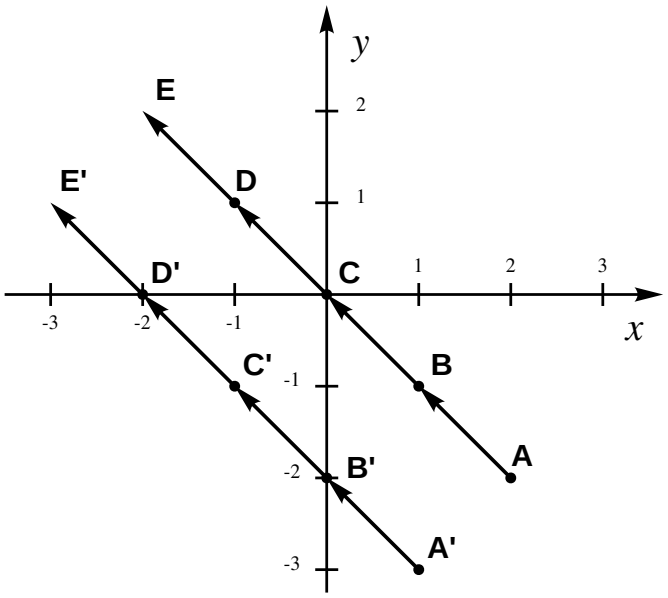



**Figure 3-3** Moving object trajectories.

frequency space CRT (cathode ray tube) of an optical processor equipped with a point-diffraction interferometer (PDI). The gray level of pixel $(k_x, k_y)$ during frame $n$ corresponds to $\Re(S(k_x, k_y, n))$ where white represents a large positive value, black represents a large negative value, and $50\ \%$ gray represents zero. The lines of constant phase make an angle $\theta = -\text{atan}\left(\frac{x_n}{y_n}\right)$ with the positive $k_x$ axis and the period of oscillation perpendicular to these lines is $\frac{2\pi}{r_n}$, where $r_n = \sqrt{x_n^2 + y_n^2}$ is the moving object's displacement from the origin. The real part of the mixed-domain representation of the primed and non-primed moving objects from Figure 3-3 are illustrated in Figure 3-4. The top sequence represents the object moving along the non-primed trajectory and the bottom sequence represents the object moving along the primed trajectory.

The non-primed object, illustrated in the top image sequence of Figure 3-4, passes through the origin of the $x$-$y$ plane. Because the derivative $\frac{d\theta}{dn}$ is zero, the slope of the lines of constant phase does not change with time and $\theta = 45°$ for all

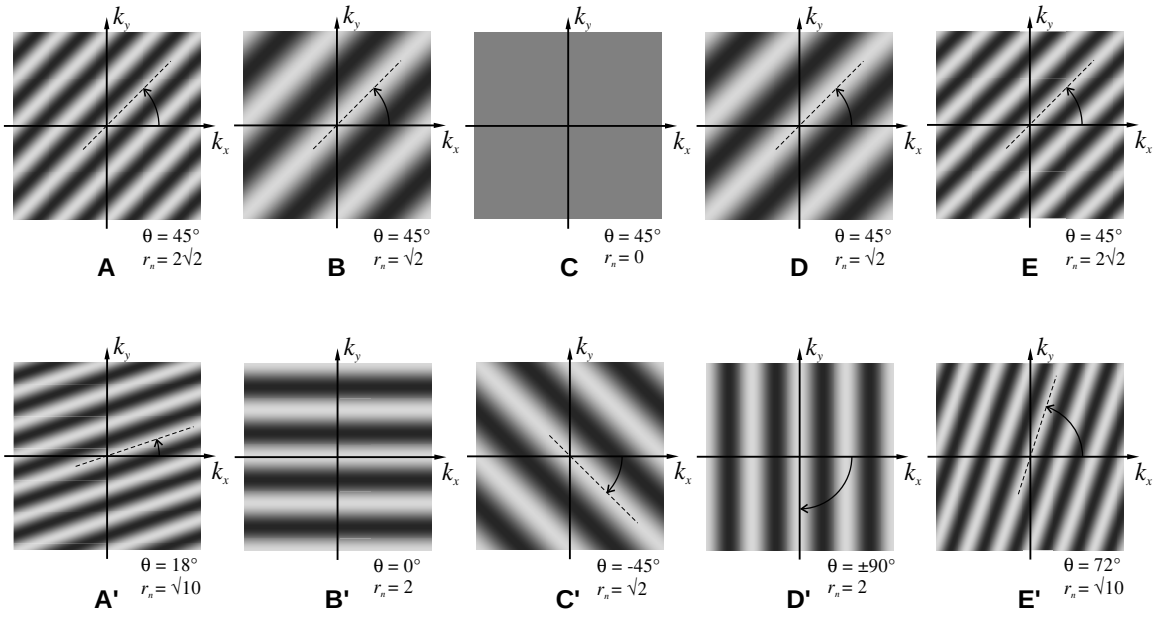



**Figure 3-4** Real part of the mixed-domain representation of a moving object. The top and bottom images sequences represent the trajectories of the two objects shown in Figure 3-3.

images in the top sequence. The frequency of the 2-D cosinusoids across the image sequence do, however, change with time. The angular frequency along a line perpendicular to the lines of constant phase is equal to the object's displacement $r_n$ from the origin. At position A, the object's displacement from the origin is large, and consequently, the frequency of the cosinusoid is large. As the object approaches the origin, the cosinusoid's frequency decreases, and eventually equals zero at position C, when the object is at the origin. The frequency increases again as the object moves away from the origin. A moving object at $(x_n, y_n)$, and it's reflection in the origin $(-x_n, -y_n)$, have the same real mixed-domain representation.

The primed object, illustrated in the bottom image sequence of Figure 3-4, does not passes through the origin of the $x$-$y$ plane. Now, both the cosinusoid's angular frequency and the angle of its lines of constant phase are functions of frame number. Two effects can be observed as the object moves from its initial

position A' to its final position E'. As with the non-primed object, the 2-D cosinusoid appears to expand as the object approaches the origin and then appears to compress as the object retreats (this time the cosinusoid's frequency doesn't reach zero because the object doesn't cross the origin). In addition to this expansion and compression, the lines of constant phase will appear to rotate about the origin of the spatial frequency plane. The rate at which the lines rotate about the origin is inversely proportional to the square of the object's displacement from the origin. As the object moves from its initial position A' to its final position E', the lines of constant phase appear to rotate in a clockwise direction about the origin.

A moving point-object is represented in the real mixed-domain by a sequence of 2-D cosinusoidal functions in spatial frequency space. The cosinusoidal function can be interpreted as the interference pattern generated by a point-object and a second point or pinhole at the origin. The frequency of the cosinusoid along the direction perpendicular to the lines of constant phase is equal to the object's displacement from the origin and the change in frequency with time is equal to the object's velocity. Similar relationships between frequency and position, and their first derivatives, hold for the frequency components along the $k_x$ and $k_y$ directions. If the object passes through the origin of the $x$-$y$ plane, the lines of constant phase remain at a constant angle with the positive $k_x$ axis for all $n$. If the object doesn't pass through the origin, the lines of constant phase appear to rotate about the origin. The speed of rotation is inversely proportional to the square of the object's displacement from the origin.

### 3.2.4 Accuracy and Aliasing

Under the mixed-domain approach to trajectory estimation, the images, from an image sequence describing the motion of a point-object, are Fourier transformed to produce a sequence of FT spectra. A set of frequency-domain points, selected as uniformly spaced pixels on a circle of radius $k_r^{circle}$ on the $k_x$-$k_y$ plane, are

processed in the time-domain using a high resolution spectral estimation technique called modified forward-backward linear prediction (FBLP). The radius of this *frequency-domain observation circle* determines the accuracy of the velocity estimates and the maximum velocity that can be estimated without causing the temporal sequence to be aliased in time.

The energy of a point-object is uniformly distributed across the 2-D spatial frequency plane. If the noise in the image sequence is assumed to be white, the signal-to-noise ratio (SNR) for temporal signals observed at all pixels in the frequency domain will be identical. The variance of all temporal frequency estimates over the spatial frequency plane will also be identical. The most accurate estimate of velocity is therefore determined when $k_r^{circle}$ is large and the temporal frequency estimate is also large. Knudsen has also shown that the velocity and direction resolving capability of the algorithm are better when $k_r^{circle}$ is large [Knu92c].

A pixel located at some critically large distance from the origin will observe a signal with temporal frequency greater than the Nyquist rate (half of the frame rate). This critical distance represents an upper bound on $k_r^{circle}$ such that the signal is not aliased in time. The temporal signal observed at a point $(k_x, k_y)$ on the spatial frequency plane, due to an object moving with velocity $[v_x \; v_y]^T$, will not be aliased in time if the absolute value of the normalized temporal frequency $\omega_t = k_x v_x + k_y v_y$ is less than $\pi$. The maximum possible velocity of the target sets an upper limit on $k_r^{circle}$, which maximizes the accuracy and resolving capability of the trajectory estimate, and guarantees that the temporal signal is not aliased in time. The velocity of the target is not known *a priori*, however, it is assumed that the maximum possible velocity is known. The maximum absolute value of the normalized temporal frequency must be less than or equal to $\pi$ in order to avoid

aliasing. This sets a limit on the sum of the products of the spatial frequency components and the velocity components,

$$\left|\omega_t\right| \;=\; \left|k_x v_x + k_y v_y\right| \;\leq\; \pi \tag{3-5}$$

The spatial frequency-domain Cartesian coordinates can be recast in polar coordinates, $k_x = k_r \cos\theta_k$ and $k_y = k_r \sin\theta_k$, where $k_r$ represents the radial distance from the origin of frequency space to the spatial frequency $(k_x, k_y)$ and $\theta_k$ represents the azimuthal angle measured from the $k_x$ axis. Similarly, the velocity components of the moving object can be recast in polar coordinates, $v_x = v\cos\phi$ and $v_x = v\sin\phi$, where $v$ is the object's speed and $\phi$ is its direction. Rewriting Equation 3-5 in polar coordinates yields,

$$\left|\omega_t\right| \;=\; k_r v\left|\cos(\phi - \theta_k)\right| \;\leq\; \pi \tag{3-6}$$

The absolute value of $\omega_t$ is maximum on the $k_x$-$k_y$ plane for spatial frequencies on a line inclined at $\theta_k = \phi$. Spatial frequencies on this line will have associated with them a temporal frequency equal to $k_r v$ while spatial frequencies not on the line will be associated with frequencies less than $k_r v$. The product of radial spatial frequency and object velocity must therefore be bound by, $k_r v \leq \pi$, in order to ensure than the absolute value of the temporal frequency is never aliased. In general, for the most accurate estimate of speed, $k_r$ should be selected to be as large as possible. The maximum possible value of an object, $v_{max}$, therefore determines the optimum radius of the spatial domain observation circle,

$$k_r^{circle} \;=\; \frac{\pi}{v_{max}} \tag{3-7}$$

The spatial frequency-domain radius $k_r^{circle}$ is selected using Equation 3-7 based on the *a priori* maximum possible velocity for a particular class of moving targets. The relationship between spatial frequency-domain radius and maximum target velocity is illustrated in Figure 3-5. The accuracy and resolving capability of the trajectory estimate is proportional to $k_r^{circle}$, however, the maximum detectable velocity is inversely proportional to $k_r^{circle}$.

The situation is somewhat more complicated for large objects. The frequency-domain representation of a large object is expressible as the product of its 2-D spatial FT and the frequency-domain representation of a point object [Cho97, Knu92a]. The energy of a large object is therefore concentrated around the origin of the 2-D spatial-frequency plane rather than being uniformly distributed over the entire plane. Specifically, the energy of a disc-shaped object of radius $\gamma$ is concentrated in a region of radius $\gamma^{-1}$, centered around the origin. The frequency-plane observation points must lie within the region of concentrated

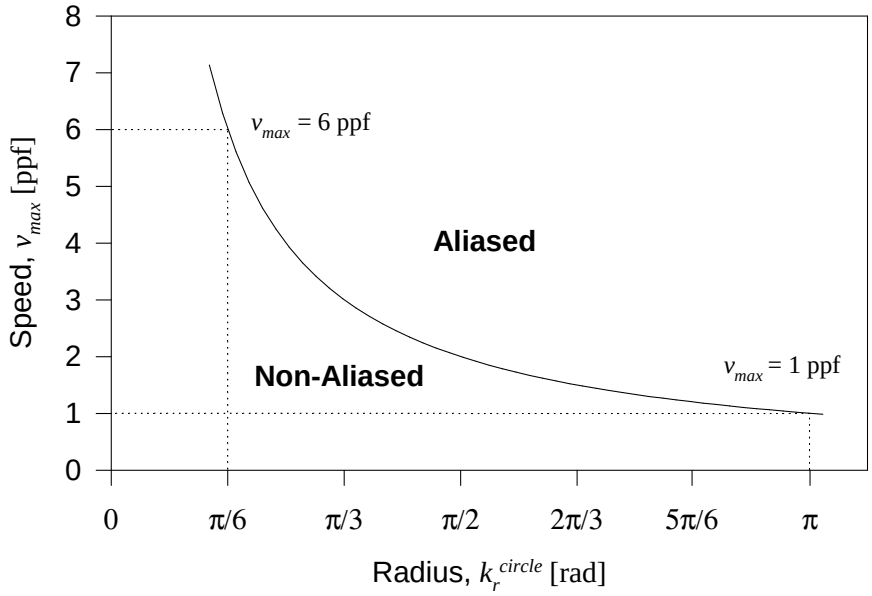



**Figure 3-5** Relationship between frequency-plane radius $k_r^{circle}$ and maximum target speed $v_{max}$.

energy in order to observe a temporal signal of high SNR, and consequently, the object's spatial extent also sets an upper limit on the choice of radius $k_r^{circle}$. It is desirable to chose $k_r^{circle}$ such that $k_r^{circle} < \frac{\pi}{\gamma}$. The majority of a typical object's energy is concentrated within a circle of radius $\frac{\pi}{3}\,\mathrm{rad}$ [Knu92a]. In general, it is desirable to choose radius $k_r^{circle}$ to be as large as possible without causing temporal aliasing and without leaving the region of concentrated energy.

### 3.3 EXPERIMENTAL

The hybrid optical-digital signal processing system for trajectory estimation is shown in Figure 3-6. An image sequence describing the linear motion of a single point-object is generated by a personal computer (PC). The images are written in sequence to a binary SLM in the front focal-plane of a Fourier transform lens (FTL) and the power spectrum of the image sequence is detected by a CCD camera in the back focal-plane. The real part of the mixed-domain signal in the back focal-plane is extracted by interfering it with a plane-wave reference generated using the point-diffraction method (see Section 2.5). A set of $T$ points from the real spectrum is processed in the time domain using a high-resolution spectral-estimation technique.

The SLM is an electrically-addressable $320\times200$ pixel twisted-nematic liquid crystal display (TN-LCD) with an $80\times90\,\mu\mathrm{m}$ pixel pitch and $30\,\mathrm{Hz}$ frame rate. The device was removed from a model LC500 Epson Crystal Image video projector [Mor86]. The TN-LCD has been characterized by several research groups [Kir91, Kir92, Gor94, Sou94a, Dou96, Kno96, Lan96a, Sta98] for various SLM optical computing applications. Its has been employed in both the input and filter planes of several optical correlators [Ama90, Cla92, Gor94, Sta98] and by the author in an optical Fourier transform processor [Lan96a, Lan97a].

The He-Ne laser beam shown in Figure 3-6 is passed through a spatial filter and beam expander assembly, composed of objective lens L1 ($20\mathrm{X}$ objective,

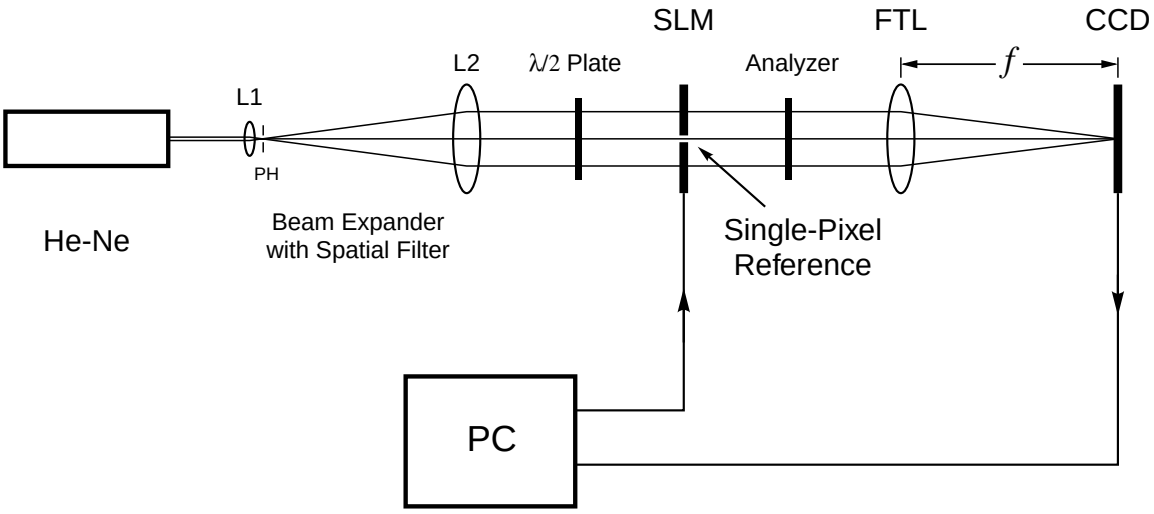



**Figure 3-6** Block diagram of the optical Fourier processor with point-diffraction interferometer (single-pixel reference).

$f_1$=8.3 mm), pinhole aperture PH ($\phi$=25 μm diameter), and collimating lens L2 (achromatic doublet, $f_2$=175 mm), to produce an expanded and filtered beam approximately $2w_1$=10 mm in diameter. The SLM is sandwiched between a λ/2 phase plate and a linear polarizer (analyzer) which are rotated such that the SLM operates as an amplitude modulator. The modulated light distribution in the plane of the SLM is Fourier transformed by lens FTL ($f$=125 mm) and its resulting intensity spectrum is measured by a CCD camera located in the back focal plane of the lens.

The SLM was used as a phase modulator in order to simplify the architecture of the optical processor and reduce the number of components. It is well known that the optical efficiency of an LCD is maximum when it is operated in a phase rather than an amplitude modulating mode. A pure phase correlation filter can have an optical efficiency of 100% in an optical correlation system [Horn82]. Operating the SLM as a phase rather than an amplitude modulator in the front focal-plane of the optical Fourier processor offers both advantages and disadvantages. It can be shown that the intensity of the interferogram produced by two pixels, modulated to have maximum phase contrast with respect to the

background, is 4 times larger than if the pixels were amplitude modulated with infinite contrast. The efficiency of the Fourier processor is increased by a factor of 4 and its sensitivity to noise is reduced. While the intensity of the interferogram is increased by a factor of 4, the DC component of the spectrum is increased by several orders of magnitude. The enormous increase in the intensity of the DC component would necessitate some kind of DC block in the back-focal plane of the lens and additional relay optics to image the intensity spectrum onto the plane of a detector. A four-fold increase in interferogram intensity could not be justified by the additional hardware required for a DC block and relay optics, and consequently, the LCD was operated in an amplitude rather than phase modulating mode.

The trajectories of several single-pixel objects were estimated experimentally using a two-stage process. The image sequence describing the motion of a moving object was optically processed first at 1/3 video frame rate by writing a series of three images derived from each image frame to the SLM and recording their power spectra. For each frame in the sequence, the power spectrum of the image frame, that of the reference, and the interferogram produced by combining the two, were measured by the CCD camera (a 1/2 video frame rate could have been achieved by measuring the power spectrum of the reference image only once at the start of the optical processing). The sequence of spectra were then post-processed off-line using MATLAB to estimate the trajectory of the object. The power spectra and interferograms were measured sequentially during the experiments, however, a parallel implementation would be more attractive on an actual real-time system.

A target animation and data acquisition flow chart is shown in Figure 3-7. All trajectories were $64$ frames in length, however, the initial position $(x_0, y_0)$ and velocity $\mathbf{v} = [v_x \; v_y]^{\mathsf{T}}$ changed with each experiment. The complex FT of the image sequence was extracted from the power spectra by interfering the signal with a

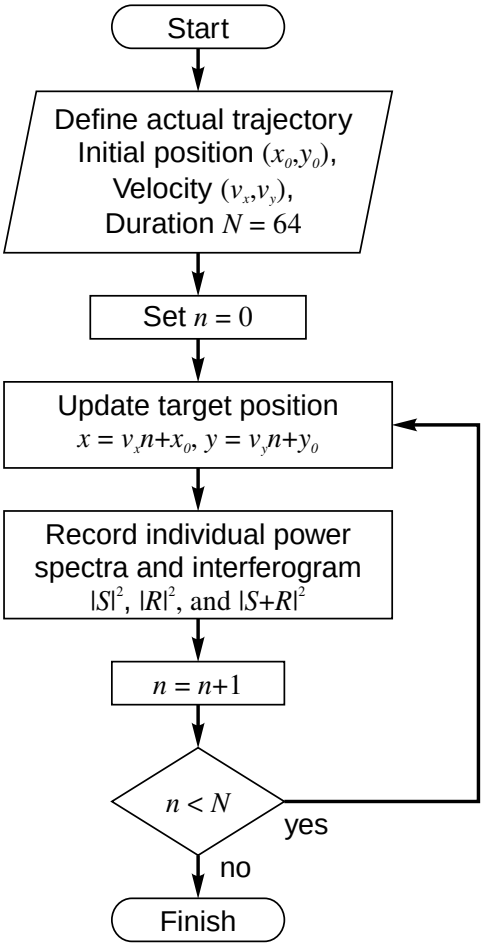



**Figure 3-7** Target animation and data acquisition flow chart.

plane-wave reference. The power spectrum of each frame was recorded by the CCD camera, with and without the plane-wave reference, so that the real part of the FT could be extracted from the power spectra. The reference was generated using the point-diffraction technique by employing Equation 2-19 (see Section 2.5) where the center pixel on the SLM was reserved for reference generation. When the reference pixel was on, it generated an approximation to a plane-wave reference in the frequency domain.

A post-processing and trajectory estimation flow chart is shown in Figure 3-8. The trajectory of the targets were estimated every $5^{\text{th}}$ frame starting with the $25^{\text{th}}$

frame and the estimation interval was $D = 25$ frames. Only selected points from the power spectra were used in the estimation. A set of $T = 12$ uniformly spaced points $Q = \{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_T \mid \mathbf{k}_t = (k_x, k_y)\}$ were chosen on a circle of radius $k_r^{circle}$ around the origin. The real part of the FT at each point $\mathbf{k}_t$ for all $N$ frames is calculated by subtracting both the power spectrum of the signal, and that of the reference, from the power spectrum of the signal and the reference.

A series of $T$ length-$N$ data vectors $\mathbf{u}_t$, one for each point $\mathbf{k}_t$, were constructed from the sequence of real spectra. The temporal frequency of the sinusoid contained in each data vector was estimated using the modified forward-backward linear prediction (FBLP) algorithm [Tuf82]. A summary of the algorithm is presented in Appendix B. It was employed in this dissertation to be consistent with the spectral estimation technique employed by the authors of the original trajectory estimation algorithm and therefore facilitate comparisons between it and the present optical implementation. Investigation of other parametric spectral estimation methods, in particular the MUSIC technique [Schm86] proposed by Schmidt, was outside the scope of this dissertation.

The frequencies of length $D$ vectors extracted from the length-$N$ vectors were estimated every $5^{\text{th}}$ frame. The signs of the temporal frequency estimates are undetermined because only the real part of the optical signals were processed. One of two possible directions for the object is arbitrarily chosen (this corresponds to choosing one of the planes described by Equation 3-4). The signs are assigned to the $T$ frequency estimates such that each point $(\mathbf{k}_t, \omega_t)$ lies on the same plane through the origin. The $T$ data points are used in a least squares fit for the plane $k_x v_x + k_y v_y - \omega = 0$ parameterized by the velocity $\mathbf{v} = [v_x \; v_y]^{\mathsf{T}}$. The speed and direction are then estimated as $\hat{v} = \sqrt{v_x^2 + v_y^2}$ and $\hat{\phi} = \operatorname{atan}\left(\frac{v_y}{v_x}\right)$.

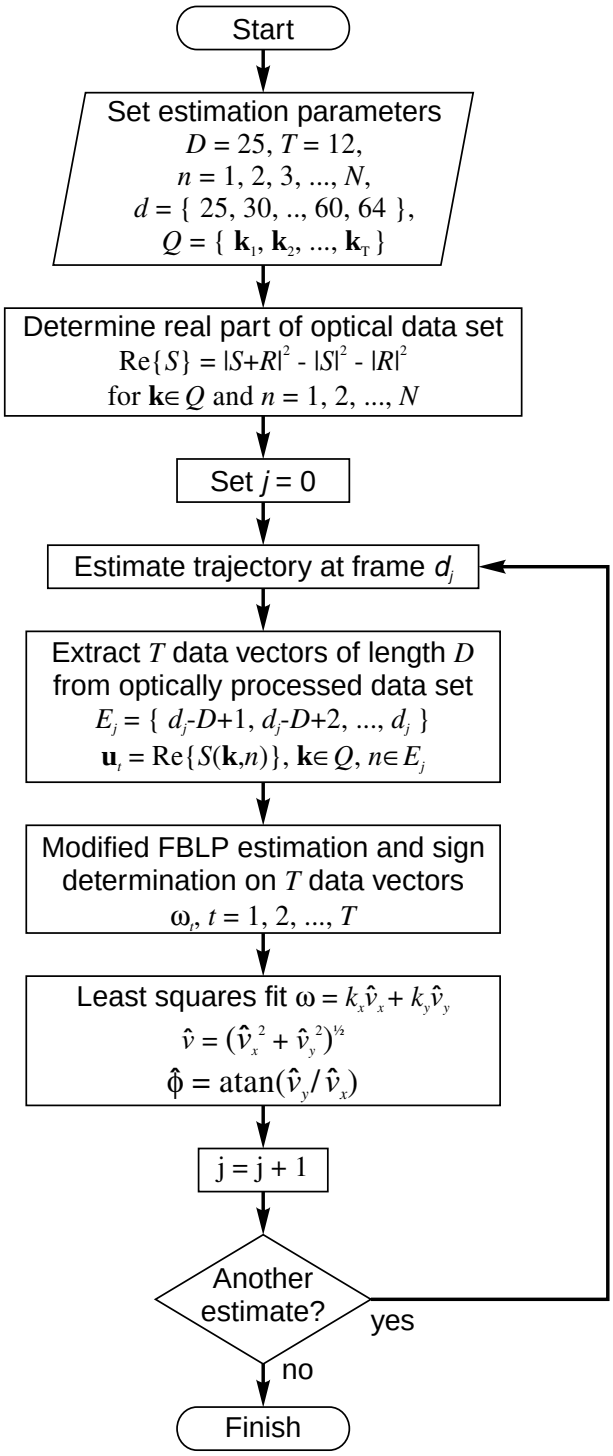



**Figure 3-8** Post processing and trajectory estimation flow chart.

## 3.4 RESULTS AND DISCUSSION

The velocity components of 13 trajectories were estimated in two sets of experiments. During the first set of experiments, targets with different speeds were moved along a straight line at -27° to the $x$ axis. During the second set of experiments, targets with a speed of either 1 or $\sqrt{2}$ ppf (pixels per frame) were moved along paths at $0$, $\pm 45$, $\pm 90$, $\pm 135$, and $180°$ to the $x$ axis. The targets were single pixels of maximum contrast on a uniform black background. The duration of each trajectory was $64$ frames. The velocity components were estimated every $5^{th}$ frame starting with the $25^{th}$ frame and the observation interval was $25$ frames in order to be consistent with estimation frequency and interval employed by Knudsen and Bruton.

### 3.4.1 Speed Estimation

The trajectories used in the speed experiment are shown in Figure 3-9. Objects were moved along a line at -27° to the $x$ axis for a duration of $64$ frames. The slowest object moved at $0.56$ ppf along the shortest trajectory A-A' ($36$ pixels) and

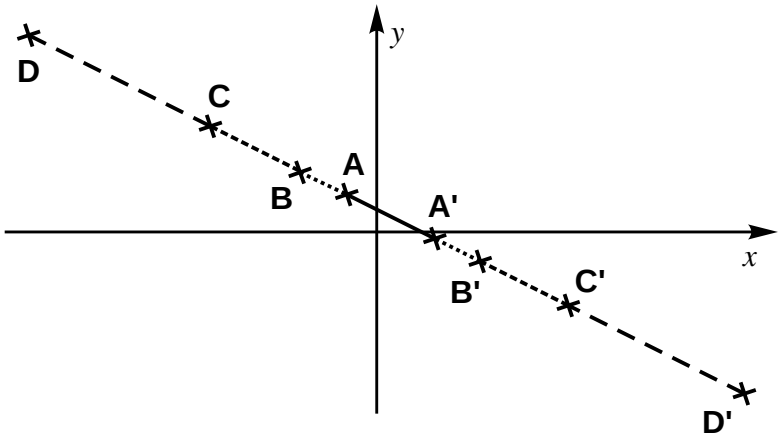



**Figure 3-9** Speed experiment actual trajectories. The duration of each trajectory is $64$ frames and the lengths ranged from $36$ pixels (trajectory A-A') to $282$ pixels (trajectory D-D').

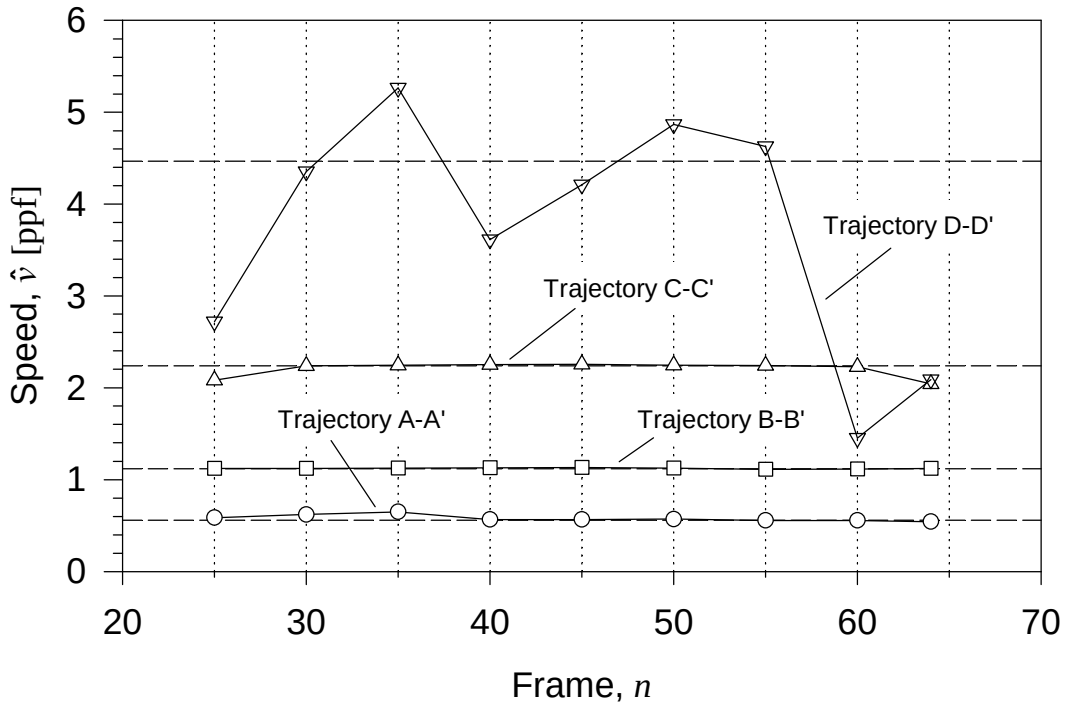



**Figure 3-10** Speed estimation for targets moving with velocity $0.56$, $1.12$, $2.24$, and $4.47$ ppf.

the fastest object moved at $4.47$ ppf along the longest trajectory D-D' ($282$ pixels). The objects' speeds were estimates every $5^{\text{th}}$ frame. These estimates, starting with the $25^{\text{th}}$ frame, are shown in Figure 3-10. The actual target speed, the average speed estimated over the duration of motion (average of $9$ estimates), and the percent errors are summarized in Table 3-1 for each trajectory. The radius $k_r^{circle}$ of the frequency-domain observation circle was set to $0.75\pi$, $0.57\pi$, $0.26\pi$, and $0.09\pi$ rad for the $0.56$, $1.12$, $2.24$, and $4.47$ ppf targets, respectively, in order to maximize the accuracy of the estimation. The speed of the fastest target could not be accurately estimated with the smallest radius $k_r^{circle} = 0.09\pi$ rad.

The average target speed was correctly estimated to within $4$ % of its actual value except for the fastest moving object which could not be accurately estimated. The standard deviation of the average speed estimate for the fastest moving object ($4.47$ ppf) in Table 3-1 is more than an order of magnitude larger

**Table 3-1** Actual and Estimated Speeds

| Trajectory | Radius | Actual and Estimated Speeds | | |
|---|---|---|---|---|
| | $k_r^{circle}$ [rad] | $v$ [ppf] | $\hat{v}$ [ppf] | % Error |
| A-A' | $0.75\pi$ | 0.56 | $0.58 \pm 0.03$ | 3.67 |
| B-B' | $0.57\pi$ | 1.12 | $1.12 \pm 0.01$ | 0.50 |
| C-C' | $0.28\pi$ | 2.24 | $2.20 \pm 0.08$ | 1.42 |
| D-D' | $0.09\pi$ | 4.47 | $3.69 \pm 1.32$ | 17.50 |

than that for the slower moving objects. Figure 3-10 shows that the speed estimates at the end-points (frames 25, 60, and 64) of trajectory D-D' are much less accurate than the mid-point estimates (frames 30 through 55). This effect is also observable to a lesser extent in trajectory C-C'. The first and last speed estimates (frames 25 and 64) are significantly lower than the other estimates (frames 30 through 55).

The trajectory of the fastest moving object (trajectory D-D', 4.47 ppf) could not be accurately estimated due to the poor contrast ratio of the SLM. Equation 3-7 suggests a radius $k_r^{circle} \leq 0.19\pi$ rad for targets moving at 4.47 ppf. The DC component of the light distribution incident on the SLM, due to its relatively poor contrast, is concentrated in a region about the origin of the spatial frequency plane. The DC component is much larger than the signal of interest and tends to dominate the intensity measurements even $0.09\pi$ rad (2 detector pixels) from the origin. Incorrectly collimated light exiting from the beam expander may have caused the DC component to be larger than necessary. The large DC spike at the origin reduces the SNR of the temporal frequency estimates and causes a large deviation in speed estimates at all frames.

Two factors can be identified which cause the accuracy of the speed estimates to be reduced during the end-points of trajectories C-C' and D-D'. The SLM is illuminated with a beam of Gaussian intensity profile. Targets near the middle of

their trajectory generate a stronger signal in the spatial-frequency plane than targets at the end-points of their trajectory. The SNR of the temporal-frequency estimates will be largest for a target on its closest approach to the origin and results in the most accurate estimate of its speed.

The accuracy of the speed estimates at the end-points of the trajectory also depends on the MTF (modulation transfer function) of the CCD camera. Figure 3-4 showed that the frequency of the 2-D cosinusoidal signals observed in the spatial frequency plane is proportional to the object's displacement from the origin of the SLM. It becomes increasingly difficult for a CCD camera with finite pixel size to resolve the cosinusoidal intensity variation as the object moves farther from the origin of the SLM. The SNR of the temporal frequency estimates and the speed estimates therefore decrease as the object moves farther from the origin.

### 3.4.2 Direction Estimation

Objects were moved along paths at $0$, $\pm 45$, $\pm 90$, $\pm 135$, and $180°$ to the $x$ axis at either $1$ or $\sqrt{2}$ ppf. Trajectories parallel to a coordinate axes were $64$ pixels long ($1$ ppf) and diagonal trajectories were $90$ pixels long ($\sqrt{2}$ ppf). Direction estimates at every $5^{\text{th}}$ frame starting with the $25^{\text{th}}$ frame are shown in Figure 3-11. The actual target direction, the average direction estimated over the duration of motion (average of $9$ estimates), and the absolute error are summarized in Table 3-2 for each trajectory. The radius of the frequency domain observation circle was $k_r^{circle} = 0.57$ rad.

As illustrated by Table 3-2, the average target direction (up to a reversal of $180°$) was correctly estimated to within $3°$ for each direction experiment. The trajectory estimation system is capable of estimating the trajectory of an object (or path described by an object) but not the direction in which the object moves along the path. The present implementation is unable to determine the direction

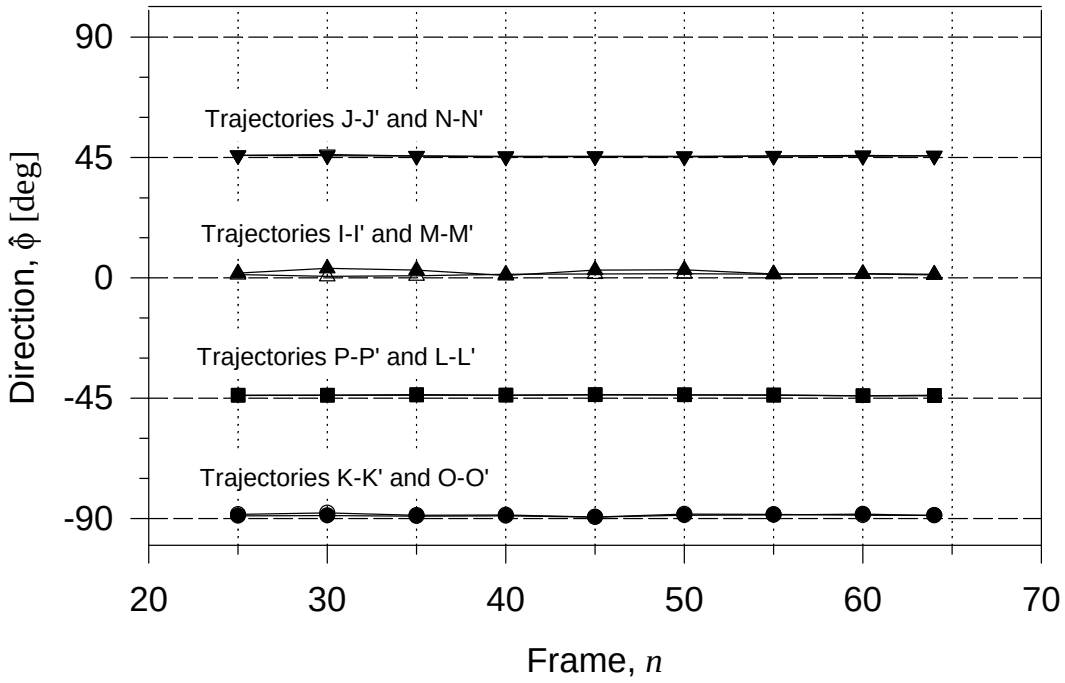



**Figure 3-11** Direction estimation for targets moving on paths $0$, $\pm 45$, $\pm 90$, $\pm 135$, and $180^{o}$ to the $x$ axis.

because only the real part of the mixed-domain signal is extracted from the spatial frequency plane. The accuracy of the direction estimations appears to be constant for all frames. The length of the direction trajectories (either $64$ or $90$ pixels) falls between the length of speed trajectories B-B' and C-C' ($71$ and $141$ pixels respectively) which also show little or no variation.

### 3.4.3 Estimation Accuracy

The estimation accuracy is largely determined by the observation radius $k_r^{circle}$. In general, the larger the observation radius, the more accurate the estimate; the choice of $k_r^{circle}$, however, is limited. The lower bound of $k_r^{circle}$ is determined by the contrast ratio of the SLM and the upper bound is determined by the maximum speed of the target. The speed and direction of a target moving along trajectory C-C' ($-27°$ to the $x$ axis with a velocity of $2.24$ ppf) were estimated using various observation radii $0 \le k_r^{circle} \le \pi$ rad. The relationship between frequency-domain

**Table 3-2** Actual and Estimated Directions

| Trajectory | Speed | Actual and Estimated Directions | | |
|---|---|---|---|---|
| | $v$ [ppf] | $\phi$ [deg] | $\hat{\phi}$ [deg] | $\lvert\phi - \hat{\phi}\rvert$ [deg] |
| I-I' | 1 | 0 | $1.22 \pm 0.31$ | 1.22 |
| J-J' | $\sqrt{2}$ | 45 | $45.58 \pm 0.26$ | 0.58 |
| K-K' | 1 | 90 | $-88.63 \pm 0.41$ | 1.37 |
| L-L' | $\sqrt{2}$ | 135 | $-43.91 \pm 0.14$ | 1.09 |
| M-M' | 1 | $\pm 180$ | $2.20 \pm 0.93$ | 2.20 |
| N-N' | $\sqrt{2}$ | -135 | $45.61 \pm 0.14$ | 0.61 |
| O-O' | 1 | -90 | $-88.91 \pm 0.30$ | 1.09 |
| P-P' | $\sqrt{2}$ | -45 | $-43.90 \pm 0.16$ | 1.10 |

radius $k_r^{circle}$ and observed estimation accuracy is illustrated in Figure 3-12. The absolute speed and direction errors appear to be minimized when $k_r^{circle} = 0.28 \text{ rad}$ which corresponded to $6$ detector pixels.

The maximum radius suggested by Equation 3-7 for a target moving at $2.24$ ppf is $0.45$ rad. The figure shows that the absolute speed and direction error increase gracefully when $k_r^{circle} > 0.38 \text{ rad}$. As the radius increases beyond $0.38$ rad, the temporal frequencies estimated at specific points on the spatial-domain observation circle begin to alias; all points on the circle do not alias at once. As $k_r^{circle}$ increases further, the number of aliased points on the circle increases, and consequently, more of the triples $(k_x, k_y, \omega)$ used to fit the plane through the origin of 3-D frequency space become outliers. The trajectory estimation accuracy therefore degrades gracefully as $k_r^{circle}$ increases and more triples become outliers. An adaptive algorithm could be implemented to flag the outliers (those that are aliased) or choose alternate frequency-domain points such that the temporal frequency is not aliased. The absolute errors also increase gracefully

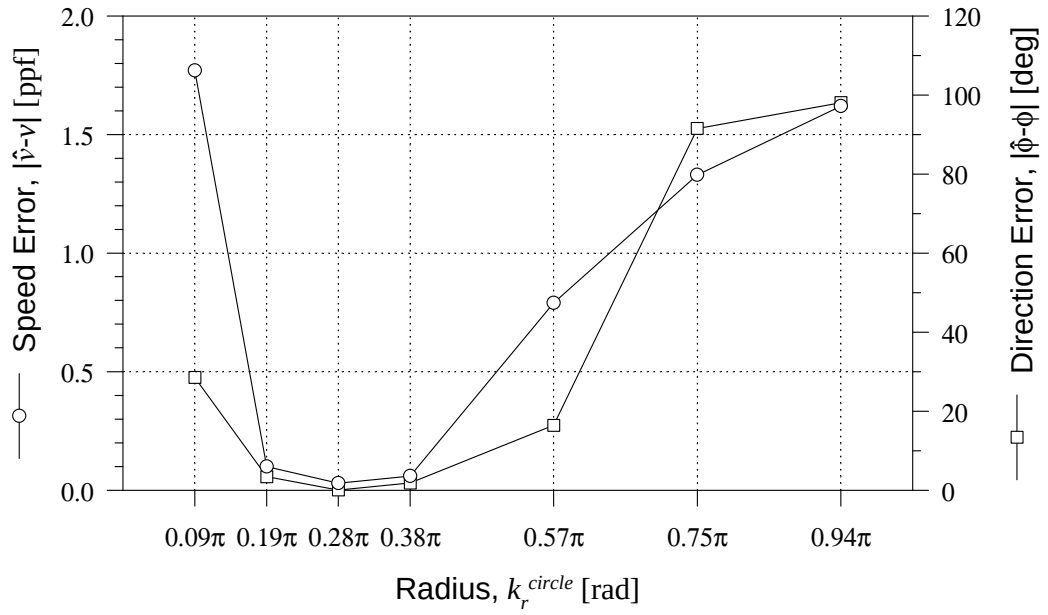



**Figure 3-12** Relationship between absolute estimation error and frequency-domain radius $k_r^{circle}$.

for $k_r^{circle} < 0.19 \text{ rad}$ because of the large DC spike centered on $k_r^{circle} = 0$ due to the poor contrast ratio of the SLM.

### 3.4.4 Multiple-Object Trajectory Estimation

The problem of multiple moving-object trajectory estimation is addressed in Chapter 5. The extension to multiple objects does not change the requirements of the spatial-domain (optical) processor, however, the time-domain (electronic) processor must detect and estimate the frequencies of an arbitrary number of sinusoidal components. The number of moving objects can be detected by estimating the median number of sinusoids present in the temporal data vectors. The original Knudsen and Bruton trajectory estimation algorithm employed the minimum descriptor length (MDL) criterion [Wax85] to detect the number of sinusoids, however, other techniques have also been proposed, most notably, the information-theoretic criterion (AIC) [Aka74] of Akaike. The number of moving objects was known *a priori* by the trajectory estimation systems demonstrated in

this dissertation (the number of sinusoids were not detected using either the MDL or AIC criteria). Except for the case of very high signal-to-noise ratios, detecting the number of sinusoids in a noisy time series is not a trivial problem and an investigation into the most effective approach was outside the scope of this dissertation. After the number of sinusoidal components has been determined, their frequencies are estimated using the modified FBLP technique [Tuf82] or a similar parametric spectral estimation method.

## 3.5 CHAPTER SUMMARY

The frequency and mixed-domain methods of trajectory estimation were reviewed in the first part of this chapter. Under the frequency-domain approach, a 3-D spatiotemporal data set is Fourier transformed into the 3-D frequency domain. A single point-object moving in space time is represented as a plane through the origin in 3-D frequency space. Multiple objects moving along different trajectories each have their own plane in frequency space. The velocity of a moving object can be determined by estimating the slope of its frequency-domain plane. Under the mixed-domain approach, the 3-D Fourier transform is reduced to a 2-D spatial Fourier transform (calculated by an analog optical processor in this work) followed by a time-domain spectral estimation technique (digital electronic processor).

A hybrid optical-digital trajectory estimation system was presented. The system was designed for the trajectory estimation of small, high-contrast, moving objects of unknown position and velocity. The system used an optical Fourier processor and point-diffraction interferometer to calculate the frequency domain representation of moving objects from which their trajectory was estimated using conventional electronic processing techniques. The capabilities of the processor were verified in two different sets of experiments. The velocity components of several targets were estimated by moving them at different velocities along the same path, and then at (almost) the same velocity along different paths.

The frequency-domain observation circle radius determines the maximum detectable speed and the overall accuracy of the estimate. When the observation radius was chosen to maximize the accuracy, the target velocity was estimated to within 4 % of its actual value, except for the fastest target moving at 4.47 ppf. Target direction was estimated to within 3° of its actual value. It was observed that the resolution and MTF of the CCD array in the Fourier plane determine the maximum displacement of a target in the $x$-$y$ plane.

The objective of this chapter was to demonstrate a hybrid optical-digital moving-object trajectory estimation system which implemented the Knudsen and Bruton mixed-domain algorithm. The system was limited to single point-objects moving over a maximum-contrast background in a noise-free environment. The system provided a useful interpretation of the mixed-domain approach as a dynamic interferogram produced by a moving target pixel and a stationary reference pixel. This interpretation is not intuitively clear when the algorithm is implemented in software alone, or when more than one target is present, however, it is very natural to an optical implementation with a single moving object. The practical disadvantages of this system are obvious and they are addressed in the following chapters.

CHAPTER

# 4

# Joint Transform Interference

A moving-object trajectory estimation system was constructed and demonstrated in the previous chapter. The system employed an optical Fourier processor and point-diffraction interferometer to calculate the Fourier spectra of a sequence of image frames. The point-diffraction method of coherent detection recovered only the real-part of the complex-valued Fourier transform spectra and restricted each image frame in the sequence to a single pixel.

The subject of this chapter is a new method of coherent detection called *joint transform interference* (JTI). The novel technique extends and improves upon the point-diffraction method and provides two significant advantages: 1) the power of the source and reference images can be easily matched in order to maximize the SNR of the Fourier transform spectra; and 2) both the real and imaginary parts of the complex-valued amplitude spectrum are recoverable. An improved trajectory estimation system which employs the JTI method of coherent detection is demonstrated in Chapter 5. The new system correctly estimates the full trajectory (speed and direction over $360^{o}$) of multiple objects of arbitrary size and shape moving over cluttered and noisy backgrounds.

## 4.1 INTRODUCTION

The JTI (joint transform interference) technique is a multi-point extension of the point-diffraction method of coherent detection. It is based on the interference between the Fourier spectra of two images – the Fourier transform (FT) of a reference image is interfered with the FT of a source image in the plane of a square-law detector. This is in contrast to the point-diffraction method where the FT of a single pixel is interfered with the FT of a source image; in JTI the reference is an image rather than a single pixel. The amplitude of the reference image can be increased in order 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 reference powers are matched. When the source and reference image powers are matched, the fringe visibility of the interferogram is optimized, the most effective use is made of the detector's limited dynamic range, and the SNR is maximum. A journal paper describing the novel JTI method of coherent detection is in preparation and will be submitted for publication shortly [Lan99d].

### 4.1.1 Phase-Shifting Interferometry

The real (or imaginary) part of an optical wavefront can be determined by interfering it with a known reference, however, a phase-shifting technique is required 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 measuring the phase distortion added by the element to an undistorted source wavefront [Mal91]. A time-varying phase shift is generally introduced between the source and reference wavefronts in an interferometer to produce a time-varying interferogram. The interferogram time-sequence is then post-processed on a point-by-point basis to determine the phase and magnitude of the distorted wavefront. The time-varying phase shift is usually introduced by mechanically translating one of the mirrors in a two-path interferometer using a piezoelectric transducer. The technique of PSI has also been demonstrated in a

common-path point-diffraction interferometer [Kad94, Mer94, Mer96] using a liquid crystal modulator to introduce the phase shift (see Section 2.3.4).

The JTI method introduced in this chapter employs a phase-stepping technique to determine the real and imaginary parts of a source wavefront. The phase shift is introduced by translating the source and 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 wavefronts are interfered with phase-shifted reference wavefronts and the phase shift is introduced directly. Under 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 mutually displacing the source and reference images. A quadrature reference is required 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.

### 4.1.2 Spatial Modulation

The interferograms generated by the JTI technique are very similar to Gabor holograms, specifically Fourier-transform Gabor holograms. The holograms first recorded by Gabor were in-line holograms as opposed to the more practical side-band holograms recorded later by Leith and Upatneiks [Ban91]. 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 virtual images along different axes, and the two images are spatially separable.

The interferograms produced by the JTI technique are very similar to Gabor holograms, however, the source and reference beams incident on the detector are typically inclined at a very small angle (the angle subtended by one or two pixels). 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 a 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 demodulating the JTI interferograms.

The fundamental principles of JTI, appropriate to either a gray scale or binary SLM, are presented in Section 4.2 of this chapter. Section 4.3 describes how the JTI technique is implemented on a binary SLM. Five classes of JTI are identified and the algorithms required to determine the real and imaginary parts of a source image FT are presented. The effect of reference image symmetry is discussed in Section 4.4 and a simple example of the JTI method is presented graphically in Section 4.5. The functional form of the interferogram modulation terms and their zeros are presented and discussed in Section 4.6. The conditions required and expressions for optimal source spectra SNR are derived and presented in the final section.

## 4.2 PRINCIPLE OF JTI

The basic principle of JTI (joint transform interference) is simple: the interference pattern produced by a source image and reference image, mutually displaced on the same SLM, is detected by a square law detector in the back focal plane of a lens. The Fourier transform (FT) of the source and reference images produces an interferogram which contains terms proportional to the real and imaginary parts of the source image FT. The real and imaginary parts can be isolated by adding and subtracting 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 Figure 4-1. The source is written to the SLM at position

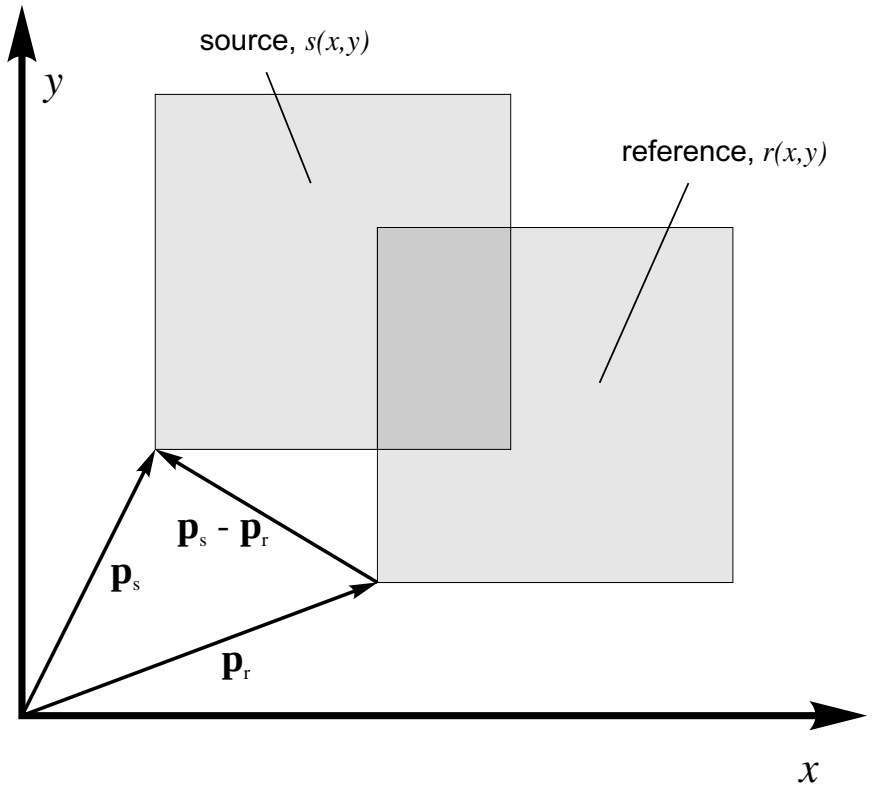



**Figure 4-1** Construction of a joint image from a source and reference image.

$\mathbf{p}_s = (x_s, y_s)$, 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,

$$s(\mathbf{p}-\mathbf{p_s}) \quad + \quad r(\mathbf{p}-\mathbf{p_r}) \qquad \Leftrightarrow \qquad \exp(i\,\mathbf{k}\cdot\mathbf{p_s})\,S(\mathbf{k}) \quad + \quad \exp(i\,\mathbf{k}\cdot\mathbf{p}_r)\,R(\mathbf{k}) \qquad (4\text{-}1)$$

where $\mathbf{p} = (x, y)$ and $\mathbf{k} = (k_x, k_y)$ represent the spatial, and spatial angular frequency dependence, respectively. The joint power spectrum detected by a square-law detector in the back focal-plane of the lens is given by,

$$\begin{aligned} I_\phi(\mathbf{k}) &= I_s(\mathbf{k}) + I_r(\mathbf{k}) + 2\Re\left(\exp(i\phi)\,S(\mathbf{k})R^*(\mathbf{k})\right) \\ &= I_s(\mathbf{k}) + I_r(\mathbf{k}) + 2\left[\cos\phi\,\Re\left(S(\mathbf{k})R^*(\mathbf{k})\right) - \sin\phi\,\Im\left(S(\mathbf{k})R^*(\mathbf{k})\right)\right] \end{aligned} \qquad (4\text{-}2)$$

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 phase difference introduced by the displacement, and the interference term has been expanded using the relationship $\Re(z_1 z_2) = \Re(z_1)\Re(z_2) - \Im(z_1)\Im(z_2)$. The $\phi$ subscript on $I_\phi(\mathbf{k})$ denotes the mutual displacement dependence between source and reference images. The joint power spectrum is the sum of three terms: 1) the power spectra of the source image; 2) the power spectrum of the reference image; and 3) an interference term which contains the real and imaginary parts of the source and reference images. The real and 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 reference images.

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

$$\Re\left(S(\mathbf{k})R^*(\mathbf{k})\right) = \frac{I_\phi(\mathbf{k}) + I_{-\phi}(\mathbf{k}) - 2I_s(\mathbf{k}) - 2I_r(\mathbf{k})}{4\cos\phi} \quad (4\text{-}3)$$

Similarly, the imaginary part is extracted by subtracting the two interference spectra and dividing by the sine modulation term,

$$\Im\left(S(\mathbf{k})R^*(\mathbf{k})\right) = \frac{I_\phi(\mathbf{k}) - I_{-\phi}(\mathbf{k})}{4\sin\phi} \quad (4\text{-}4)$$

If the reference image is designed such that $R^*(\mathbf{k})=1$, then the real and imaginary parts of the source image are described by the right-hand sides of Equations 4-3 and 4-4 exactly. When $R^*(\mathbf{k})$ is a function of spatial frequency, it must be moved to the right-hand side of the two equations in order to isolate the real and 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 Equations 4-3 and 4-4 to the respective denominator of the right-hand side. If the frequency domain reference is complex the equations are more involved (see Section 4.4).

## 4.3 JOINT TRANSFORM INTERFERENCE ON A BINARY SLM

The JTI principle discussed in the preceding section is directly applicable to a grayscale 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 reference images can be directly added and written to a grayscale 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 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. In order to realize the JTI method on a binary SLM, the source and

reference images must be interleaved rather than summed. The disadvantage to this simple solution is the reduction in the effective space-bandwidth product of the SLM.

On a binary SLM with square pixel structure, the source and reference images may be interleaved either horizontally, vertically, or diagonally. Consider, for example, a horizontal row-interleaved (or interlaced) configuration analogous to the even and odd fields of an NTSC (national television systems committee) video frame. When the source and reference images are displayed as the even and 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 reference images are displayed as the odd and 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 odd-interlaced images. Similarly, the imaginary part of the FT of the source image can be extracted from the difference of the two joint power spectra.

Expressions for the JTI spectra, produced by various source and reference image configurations, are derived and presented in the remainder of this section. 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 either 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 constructing the appropriate joint image such that it contains exactly four components. The JTI spectra for each major class of joint image and the algorithms to extract the real and imaginary parts of the source are presented in this section. Expressions for the JTI spectra of all possible joint images and their algorithms are summarized in Appendix C.

In order to accommodate all possible single-pixel mutual displacements between source and reference images, the $2M \times 2N$ pixels of a binary SLM are grouped into an $M \times N$ array of $2 \times 2$ pixels cells. The set of corresponding pixels from each cell are assigned its own $M \times N$ component image which can be either a source, reference, or 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 Figure 4-2.

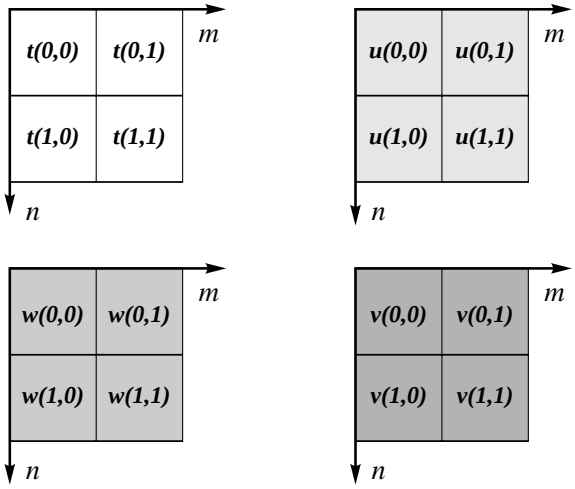
The joint image is constructed mathematically by upsampling each component image and then summing unit-shifted versions of the upsampled images. The component images are upsampled by inserting rows and columns of zeros between each row and column of valid pixel data. If $t(m,n)$ is an $M \times N$ image sequence, it's twice upsampled representation is a $2M \times 2N$ image sequence described by,

$$\tilde{t}(m,n) = \begin{cases} t\left(\frac{m}{2}, \frac{n}{2}\right) & n,m \text{ even} \\ 0 & \text{otherwise} \end{cases} \tag{4-5}$$

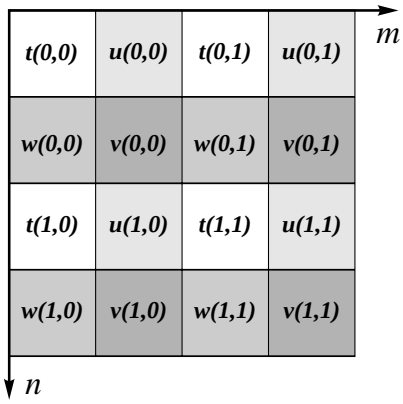
where $m = 0, 1, 2, \ldots, 2M-1$, and $n = 0, 1, 2, \ldots, 2N-1$. The $2M \times 2N$ composite image sequence $f(m,n)$ illustrated in Figure 4-2(b) is then described by,

$$f_{\substack{t \; u \\ w \; v}}(m,n) = \tilde{t}(m,n) + \tilde{u}(m-1,n) + \tilde{v}(m-1,n-1) + \tilde{w}(m,n-1) \tag{4-6}$$

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 $m$ axes increase from left-to-right and the $y$ and $n$ axes increase from top-to-bottom). If the discrete space Fourier transforms (DSFTs) 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 DSFTs of the




(a) component images, $t(m,n)$, $u(m,n)$, $v(m,n)$, and $w(m,n)$




(b) joint image, $f(m,n)$

**Figure 4-2** Component images shown in part (a) are assembled to construct the joint image shown in (b).

upsampled images are given by $T(2\mathbf{k})$, $U(2\mathbf{k})$, $V(2\mathbf{k})$, and $W(2\mathbf{k})$, respectively [Str96, p. 93]. The upsampling operation stretches the component images in the spatial domain by a factor of two; the spatial-domain expansion corresponds to a frequency-domain contraction by a factor of two. The DSFT of the joint image is then given by,

$$\begin{aligned} F_{\substack{t\ u \\ w\ v}}(\mathbf{k}) &= T(2\mathbf{k}) + \exp(ik_x)U(2\mathbf{k}) \\ &\quad + \exp\left[i(k_x + k_y)\right]V(2\mathbf{k}) + \exp(ik_y)W(2\mathbf{k}) \end{aligned} \tag{4-7}$$

and its power (or intensity) spectrum is given by,

$$I_{\substack{t\ u\\ w\ v}}(\mathbf{k}) = \left| F_{\substack{t\ u\\ w\ v}}(\mathbf{k}) \right|^2 \tag{4-8}$$

The general intensity spectrum described by Equation 4-8 consists of $16$ complex-valued terms which sum to a real-valued intensity. Expressions for various JTI spectra can be derived by setting each of the four general component spectra in Equation 4-7 to either source spectrum $S(2\mathbf{k})$, reference spectrum $R(2\mathbf{k})$, or zero, and then reducing Equation 4-8 to an *interference equation* form. For example, the expression for the JTI spectrum produced by a non-shifted 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,

$$\begin{aligned} I_{\substack{s\ r\\ 0\ 0}}(\mathbf{k}) &= \left| S(2\mathbf{k}) + \exp(ik_x)\, R(2\mathbf{k}) \right|^2 \\ &= I_s + I_r + 2\left[\cos(k_x)\,\Re\left(SR^*\right) + \sin(k_x)\,\Im\left(SR^*\right)\right] \end{aligned} \tag{4-9}$$

where $I_s$ and $I_r$ represent the power spectra of the source and 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 imaginary parts of the complex function $SR^*$ to improve the clarity of the interference equations.

The JTI spectrum in Equation 4-9 represents the interference pattern produced by a joint image composed of a non-shifted 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; and
3) and interference term given by two times the sum of a cosine modulated $\Re\left(SR^*\right)$ and sine modulated $\Im\left(SR^*\right)$.

The modulating functions which pre-multiply the real and imaginary parts of the complex function $SR^*$ are due to the mutual source-reference displacement in the spatial domain.

All JTI spectra presented in this section will have a simple three component form characteristic of the interference equation. In the example presented in Equation 4-9, 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 imaginary parts of $SR^*$ are always modulated by terms composed of cosine and sine functions, respectively.

Several classes of JTI spectra can be identified depending on the mutual source-reference displacement (horizontal, vertical, or diagonal) and the number of interfering source and reference image components. Five classes of JTI spectra and the algorithms used to extract the real and imaginary parts of $SR^*$ are derived and presented in the following sections. Expressions for the JTI spectra of all possible joint images and their algorithms are summarized in Appendix C.

### 4.3.1 Horizontal and Vertical JTI

A single source image is interfered with a single reference image. The source and reference images are interfered with respect to a line which can be either horizontal or vertical. If the source is displaced to the left or right of the reference, the interference is called vertical JTI, however, 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 presented below,

$$
\begin{aligned}
I_{\substack{s\ r\\0\ 0}} &= I_{\substack{0\ 0\\s\ r}} = I_s + I_r + 2\left[\cos(k_x)\,\Re\left(SR^*\right) + \sin(k_x)\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\ s\\0\ 0}} &= I_{\substack{0\ 0\\r\ s}} = I_s + I_r + 2\left[\cos(k_x)\,\Re\left(SR^*\right) - \sin(k_x)\,\Im\left(SR^*\right)\right] \\
I_{\substack{s\ 0\\r\ 0}} &= I_{\substack{0\ s\\0\ r}} = I_s + I_r + 2\left[\cos(k_y)\,\Re\left(SR^*\right) + \sin(k_y)\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\ 0\\s\ 0}} &= I_{\substack{0\ r\\0\ s}} = I_s + I_r + 2\left[\cos(k_y)\,\Re\left(SR^*\right) - \sin(k_y)\,\Im\left(SR^*\right)\right]
\end{aligned}
\tag{4-10}
$$

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

$$
\begin{aligned}
4\sin(k_x)\,\Im\left(SR^*\right) &= I_{\substack{s\ r\\0\ 0}} - I_{\substack{r\ s\\0\ 0}} \\
4\sin(k_y)\,\Im\left(SR^*\right) &= I_{\substack{s\ 0\\r\ 0}} - I_{\substack{r\ 0\\s\ 0}}
\end{aligned}
\tag{4-11}
$$

and the cosine-modulated real part can be determined using one of,

$$
\begin{aligned}
4\cos(k_x)\,\Re\left(SR^*\right) &= I_{\substack{s\ r\\0\ 0}} + I_{\substack{r\ s\\0\ 0}} - 2I_s - 2I_r \\
4\cos(k_y)\,\Re\left(SR^*\right) &= I_{\substack{s\ 0\\r\ 0}} + I_{\substack{r\ 0\\s\ 0}} - 2I_s - 2I_r
\end{aligned}
\tag{4-12}
$$

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

$$
I_s = I_{\substack{0\ 0\\s\ 0}} = I_{\substack{0\ 0\\0\ s}} = I_{\substack{0\ s\\0\ 0}} = I_{\substack{s\ 0\\0\ 0}} \tag{4-13}
$$

and similarly for the power spectrum of the reference image.

### 4.3.2 Diagonal JTI

A single source image is diagonally interfered with a single reference image. The source and reference images may be interfered with respect to a line which is diagonal with either positive or negative slope. If the source and reference are displaced along a diagonal line with 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 presented below,

$$\begin{aligned}
I_{\substack{s\ 0\\0\ r}} &= I_s + I_r + 2\left[c_2\,\Re\left(SR^*\right) + s_2\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\ 0\\0\ s}} &= I_s + I_r + 2\left[c_2\,\Re\left(SR^*\right) - s_2\,\Im\left(SR^*\right)\right] \\
I_{\substack{0\ r\\s\ 0}} &= I_s + I_r + 2\left[c_1\,\Re\left(SR^*\right) + s_1\,\Im\left(SR^*\right)\right] \\
I_{\substack{0\ s\\r\ 0}} &= I_s + I_r + 2\left[c_1\,\Re\left(SR^*\right) - s_1\,\Im\left(SR^*\right)\right]
\end{aligned} \tag{4-14}$$

where the cosine and sine modulation functions are given by $c_1 = \cos\left(k_x - k_y\right)$, $c_2 = \cos\left(k_x + k_y\right)$, and $s_1 = \sin\left(k_x - k_y\right)$, $s_2 = \sin\left(k_x + k_y\right)$, respectively. As before, the change in sign of the sine modulating function can be utilized to extract the real and imaginary parts of $SR^*$. The sine-modulated imaginary part can be determined using one of the following algorithms,

$$\begin{aligned}
4s_2\,\Im\left(SR^*\right) &= I_{\substack{s\ 0\\0\ r}} - I_{\substack{r\ 0\\0\ s}} \\
4s_1\,\Im\left(SR^*\right) &= I_{\substack{0\ r\\s\ 0}} - I_{\substack{0\ s\\r\ 0}}
\end{aligned} \tag{4-15}$$

and the cosine-modulated real part can be calculated using one of,

$$\begin{aligned} 4c_2\,\Re\{SR^*\} &= I_{\substack{s\ 0\\0\ r}} + I_{\substack{r\ 0\\0\ s}} - 2I_s - 2I_r \\ 4c_1\,\Re\{SR^*\} &= I_{\substack{0\ r\\s\ 0}} + I_{\substack{0\ s\\r\ 0}} - 2I_s - 2I_r \end{aligned} \tag{4-16}$$

### 4.3.3 Three-Component JTI

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. Expressions for the JTI spectra of four representative cases are presented below. Appendix C contains a summary of all possible three-component JTI spectra and algorithms to extract the real and imaginary parts.

$$\begin{aligned} I_{\substack{0\ s\\s\ r}} &= 2(1+c_1)I_s + I_r + 2\left[c_4\,\Re\{SR^*\} + s_4\,\Im\{SR^*\}\right] \\ I_{\substack{r\ s\\s\ 0}} &= 2(1+c_1)I_s + I_r + 2\left[c_4\,\Re\{SR^*\} - s_4\,\Im\{SR^*\}\right] \\ I_{\substack{r\ 0\\s\ r}} &= I_s + 2(1+c_2)I_r + 2\left[c_4\,\Re\{SR^*\} + s_3\,\Im\{SR^*\}\right] \\ I_{\substack{r\ s\\0\ r}} &= I_s + 2(1+c_2)I_r + 2\left[c_4\,\Re\{SR^*\} - s_3\,\Im\{SR^*\}\right] \end{aligned} \tag{4-17}$$

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 real and imaginary parts of $SR^*$ can be determined by adding and subtracting pairs of JTI spectra. The sine modulated imaginary part can be extracted using one of the following algorithms,

$$\begin{aligned} 4s_4\,\Im\left(SR^*\right) &= I_{\substack{0\ s\\ s\ r}} - I_{\substack{r\ s\\ s\ 0}} \\ 4s_3\,\Im\left(SR^*\right) &= I_{\substack{r\ 0\\ s\ r}} - I_{\substack{r\ s\\ 0\ r}} \end{aligned} \tag{4-18}$$

and the cosine modulated real part can be extracted using one of the following algorithms,

$$\begin{aligned} 4c_4\,\Re\left(SR^*\right) &= I_{\substack{0\ s\\ s\ r}} + I_{\substack{r\ s\\ s\ 0}} - 2I_{\substack{0\ s\\ s\ 0}} - 2I_r \\ 4c_4\,\Re\left(SR^*\right) &= I_{\substack{r\ 0\\ s\ r}} + I_{\substack{r\ s\\ 0\ r}} - 2I_s - 2I_{\substack{r\ 0\\ 0\ r}} \end{aligned} \tag{4-19}$$

where $I_{\substack{0\ s\\ s\ 0}}$ and $I_{\substack{r\ 0\\ 0\ r}}$ represent the modulated power spectra of the double-source and double-reference images, $2(1+c_1)I_s$ and $2(1+c_2)I_r$, respectively.

### 4.3.4 Symmetric Four-Component JTI

Symmetric JTI is a special case of 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 due to the special symmetry of the joint image. Expressions for the JTI spectra of both possible configurations are presented below,

$$\begin{aligned} I_{\substack{s\ r\\ r\ s}} &= 2(1+c_2)I_s + 2(1+c_1)I_r + 4c_4\,\Re\left(SR^*\right) \\ I_{\substack{r\ s\\ s\ r}} &= 2(1+c_1)I_s + 2(1+c_2)I_r + 4c_4\,\Re\left(SR^*\right) \end{aligned} \tag{4-20}$$

The cosine modulated real part of the complex function $SR^*$ can be calculated using one of the following algorithms,

$$
\begin{aligned}
4c_4\,\Re\{SR^*\} &= I_{\substack{s\ r\\ r\ s}} - I_{\substack{s\ 0\\ 0\ s}} - I_{\substack{0\ r\\ r\ 0}} \\
4c_4\,\Re\{SR^*\} &= I_{\substack{r\ s\\ s\ r}} - I_{\substack{0\ s\\ s\ 0}} - I_{\substack{r\ 0\\ 0\ r}}
\end{aligned}
\tag{4-21}
$$

and the imaginary part can not be determined.

### 4.3.5 Four-Component JTI

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. Expressions for four representative cases are presented below and a summary of all possible four-component JTI spectra and the corresponding algorithms are summarized in Appendix C.

$$
\begin{aligned}
I_{\substack{s\ s\\ s\ r}} &= (3+2c_5)I_s + I_r + 2\left[c_6\,\Re\{SR^*\} + s_6\,\Im\{SR^*\}\right] \\
I_{\substack{r\ s\\ s\ s}} &= (3+2c_5)I_s + I_r + 2\left[c_6\,\Re\{SR^*\} - s_6\,\Im\{SR^*\}\right] \\
I_{\substack{r\ r\\ s\ r}} &= I_s + (3+2c_6)I_r + 2\left[c_5\,\Re\{SR^*\} + s_5\,\Im\{SR^*\}\right] \\
I_{\substack{r\ s\\ r\ r}} &= I_s + (3+2c_6)I_r + 2\left[c_5\,\Re\{SR^*\} - s_5\,\Im\{SR^*\}\right]
\end{aligned}
\tag{4-22}
$$

where the cosine modulation functions are defined,

$$
\begin{aligned}
c_5 &= \cos(k_x)+\cos(k_y)+\cos(k_x-k_y) \\
c_6 &= \cos(k_x)+\cos(k_y)+\cos(k_x+k_y)
\end{aligned}
\tag{4-23}
$$

and the sine modulation functions are defined,

$$
\begin{aligned}
s_5 &= \sin(k_x)-\sin(k_y)+\sin(k_x-k_y) \\
s_6 &= \sin(k_x)+\sin(k_y)+\sin(k_x+k_y)
\end{aligned}
\tag{4-24}
$$

The modulated real and imaginary parts of $SR^*$ can be determined by adding and subtracting appropriate pairs of JTI spectra. The sine modulated imaginary part can be extracted using one of the following algorithms,

$$\begin{aligned} 4s_6\,\Im\left(SR^*\right) &= I_{\substack{s\,s\\s\,r}} - I_{\substack{r\,s\\s\,s}} \\ 4s_5\,\Im\left(SR^*\right) &= I_{\substack{r\,r\\s\,r}} - I_{\substack{r\,s\\r\,r}} \end{aligned} \tag{4-25}$$

and the cosine modulated real part can be extracted using one of the following algorithms,

$$\begin{aligned} 4c_6\,\Re\left(SR^*\right) &= I_{\substack{s\,s\\s\,r}} + I_{\substack{r\,s\\s\,s}} - 2I_{\substack{0\,s\\s\,s}} - 2I_r \\ 4c_5\,\Re\left(SR^*\right) &= I_{\substack{r\,r\\s\,r}} + I_{\substack{r\,s\\r\,r}} - 2I_s - 2I_{\substack{r\,r\\0\,r}} \end{aligned} \tag{4-26}$$

where $I_{\substack{0\,s\\s\,s}}$ and $I_{\substack{r\,r\\0\,r}}$ represent the modulated power spectra of the triple-source and triple-reference images, $(3+2c_5)I_s$ and $(3+2c_6)I_r$, respectively.

Expressions for five different classes of JTI spectra have been presented in this section. The algorithms presented with the spectra may be used to extract the modulated real and imaginary parts of the complex product $SR^*$ from the JTI spectra. If the FT of the source image is required explicitly, the extracted terms must be demodulated and then normalized with respect to the reference. This corresponds to isolating the real and imaginary parts of the source-image FT from the sinusoidal modulation and known reference images. The convenience of employing a uniform reference of constant value at all spatial frequencies is obvious. It is also advantageous to design reference images with either odd or even symmetry.

## 4.4 SYMMETRIC AND NON-SYMMETRIC REFERENCE IMAGES

When a reference image is engineered to possess even symmetry, its frequency-domain representation $R^*$ is real valued. In this case, the real and imaginary parts of the complex spectrum $SR^*$ can be simplified to $R\,\Re(S)$ and $R\,\Im(S)$, respectively. Similarly, when a reference image possesses odd symmetry, its frequency-domain representation $R^*$ is pure imaginary, and the real and imaginary parts of $SR^*$ can be simplified to $-iR\,\Im(S)$ and $iR\,\Re(S)$ (the real part of $SR^*$ contains the imaginary part of $S$ and vice versa). In general, when the reference image has both even and odd components, it is neither odd or even symmetric, its frequency domain representation is complex-valued, and the expressions for the real and imaginary parts of $S$ are more complicated.

More generally, $\Re(SR^*)$ and $\Im(SR^*)$ can be expanded into the real and imaginary parts of the source and reference spectra,

$$\begin{aligned} \Re(SR^*) &= \Re(S)\Re(R^*) - \Im(S)\Im(R^*) \\ \Im(SR^*) &= \Re(S)\Im(R^*) + \Im(S)\Re(R^*) \end{aligned} \tag{4-27}$$

Solving Equation 4-27 for the real and imaginary parts of the source spectra $\Re(S)$ and $\Im(S)$, in terms of the spatial domain reference $R$, and the real and imaginary parts of $SR^*$, yields,

$$\begin{aligned} \Re(S) &= \frac{\Re(R)\Re(SR^*) - \Im(R)\Im(SR^*)}{|R|^2} \\ \Im(S) &= \frac{\Im(R)\Re(SR^*) + \Re(R)\Im(SR^*)}{|R|^2} \end{aligned} \tag{4-28}$$

When the reference $R$ is pure real, the expressions for the real and imaginary parts of $S$ are given by $\Re(SR^*)R^{-1}$ and $\Im(SR^*)R^{-1}$, respectively. Similarly, when $R$ is pure imaginary, the real and imaginary parts of $S$ are given by $-i\,\Im(SR^*)R^{-1}$

and $i\,\Re\left(SR^{*}\right)R^{-1}$, respectively. A symmetric reference image greatly simplifies the calculations required to isolate the real and imaginary parts of the source image FT.

## 4.5 JTI EXAMPLE

It is illustrative to step through the calculation of the real and imaginary parts of a source image FT using the JTI technique. Consider a source image which consists of a single pixel at $(m,n)=(1,0)$ and a $3\times1$ pixel reference centered on the origin. The source and reference images were chosen to have single-pixel dimensions in the $y$ direction so their spectra have no $k_y$ dependence. The trivial images were chosen to facilitate their representation and interpretation on 2-D axes and is by no means a limitation of the JTI technique.

This example employs horizontal JTI to determine the complex spectrum of the source image. Expressions for the appropriate interference patterns were presented in Equation 4-10 and the algorithm required to extract the real and imaginary parts were described by Equations 4-11 and 4-12, respectively. The calculation is shown graphically in Figure 4-3. The frequency-domain images are plotted on 2-D axes with $k_y=0$. The power spectra of the source, $I_s(2k_x)$, and the reference, $I_s(2k_x)$, are shown in Figure 4-3(a). The power spectra of the left-displaced joint image $I_{\substack{s\,r\\0\,0}}(2k_x)$, and the right-displaced joint image, $I_{\substack{r\,s\\0\,0}}(2k_x)$, are illustrated in Figure 4-3(b). The four images illustrated in Figures 4-3(a) and (b) constitute the ensemble of images (power spectra) required by the JTI method to determine the source image complex spectrum. It is convenient to define the sum and difference of the two joint images,

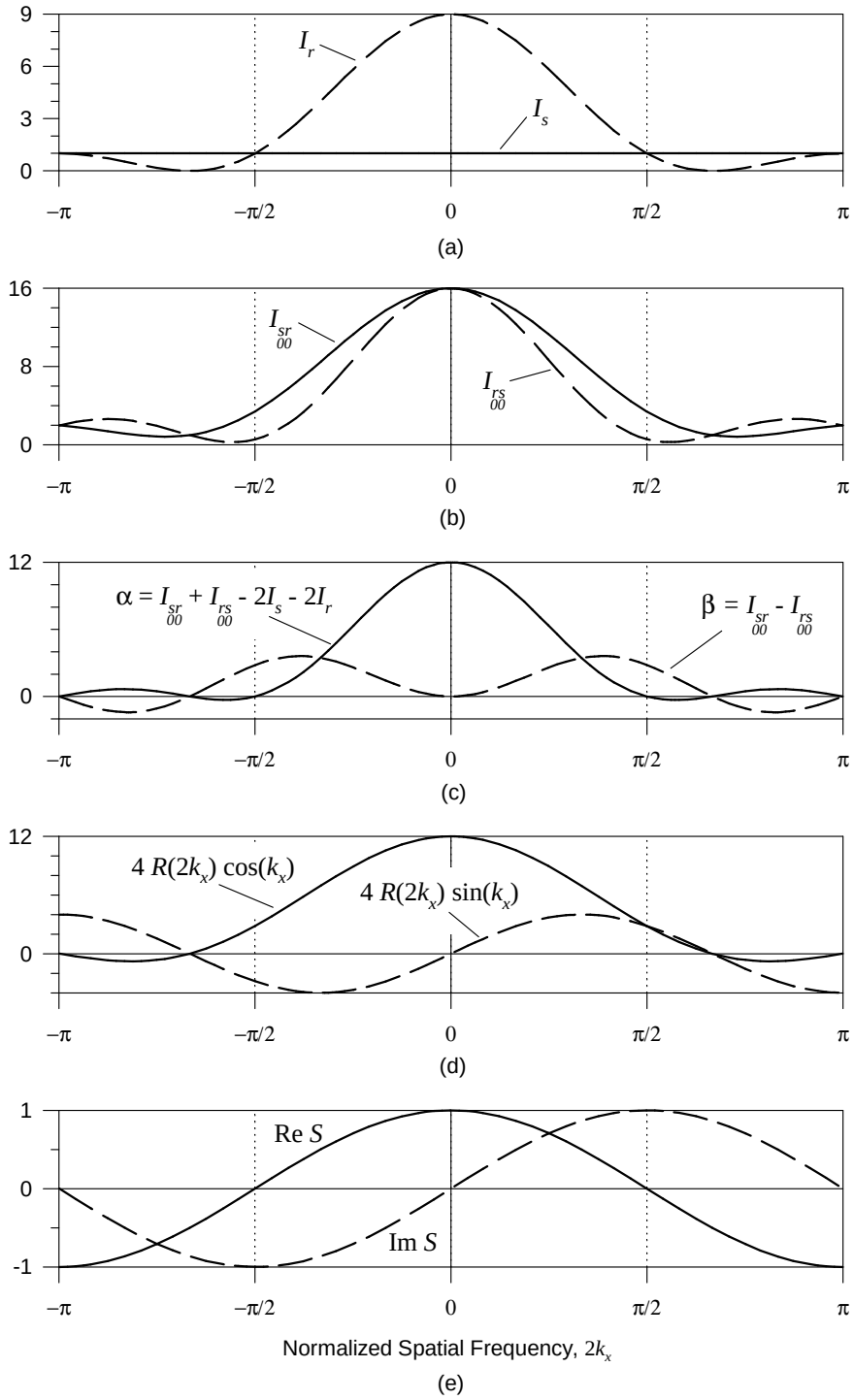



**Figure 4-3** JTI example: (a) individual source and reference intensity spectra; (b) joint intensity spectra; (c) spectra sum and difference; (d) modulation and normalization functions; (e) real and imaginary parts of source amplitude spectra.

$$\begin{aligned} \alpha(2k_x) &= I_{\substack{sr \\ 00}}(2k_x) + I_{\substack{rs \\ 00}}(2k_x) - 2\left[I_r(2k_x)+I_s(2k_x)\right] \\ \beta(2k_x) &= I_{\substack{sr \\ 00}}(2k_x) - I_{\substack{rs \\ 00}}(2k_x) \end{aligned} \tag{4-29}$$

as described by the algorithms presented in Equations 4-11 and 4-12. The two joint spectra, $I_{\substack{sr \\ 00}}$ and $I_{\substack{rs \\ 00}}$, and the two individual spectra, $I_s$, and $I_r$, are measured using a CCD camera. The two joint spectra, and the individual source spectrum, must be measured for each new source image, however the spectrum of the reference must only be measured the first time it is used. The sum and difference functions described by Equation 4-29 are illustrated in Figure 4-3(c).

The sum and difference images must be demodulated and normalized with respect to the reference spectrum. The reference image has even symmetry, its FT is real valued, and consequently, the demodulation and normalization is performed by dividing the sum image $\alpha(2k_x)$ by a cosine correction image $4R(2k_x)\cos(k_x)$, and dividing the difference image $\beta(2k_x)$ by a sine correction image $4R(2k_x)\sin(k_x)$. These correction images are illustrated in Figure 4-3(d).

The sine correction image has zeros at $2k_x = 0, \pm\frac{2\pi}{3}$ (three zeros total) and the cosine correction image has zeros at $2k_x = \pm\frac{2\pi}{3}, \pm\pi$ (four zeros total). The zeros at $2k_x = 0, \pm\pi$ are due to the cosine and sine modulation terms in the correction functions while the zeros at $2k_x = \pm\frac{2\pi}{3}$ are due to the zeros of the reference. The zeros of the correction functions will introduce errors into the determination of the source image complex spectrum when the power spectra are quantized with a finite number of levels.

The complex-valued spectrum of the source is illustrated in Figure 4-3(e). The FT of a single pixel at $(m,n)=(1,0)$ is given by $S(\mathbf{k}) = \exp(i\,k_x) = \cos(k_x) + i\sin(k_x)$ in a non-compressed coordinate system. In $k_x$ compressed frequency space, its FT is given by $S(2\mathbf{k}) = \cos(2k_x) + i\sin(2k_x)$. The complex spectrum of the source

image calculated using the JTI method is identical to its FT. A single-pixel source image was used to illustrate the JTI technique in this example, however, the technique is not limited to single-pixel images. When the source image contains finite extended objects, the reference image must be power matched to that of the source, in order to guarantee good fringe visibility, and reduce the quantization error introduced by the detector.

The JTI technique illustrated in this example requires the measurement of two intensity spectra and two interferograms to determine the real and imaginary parts of a single source image. The intensity spectra and interferograms could be measured sequentially by a CCD camera, however, the preferred method would involve parallel measurement. The CCD camera detects an averaged optical intensity at each pixel, due to the finite extended aperture of each pixel, rather than an ideal sampled point intensity. A subpixel interpolation algorithm may be employed to reduce the affect of this averaging over the spatial frequency plane.

## 4.6 MODULATION-FUNCTION ZEROS

The JTI spectra presented in Section 4.4 represent the interferograms produced by mutually displaced source and reference images. The interference term of a JTI spectrum is modulated by trigonometric functions which depend on the direction and magnitude of the displacement. The $\Re\left(SR^{*}\right)$ term is modulated by a cosinusoidal function and the $\Im\left(SR^{*}\right)$ term is modulated by a sinusoidal function. The period of the modulating function is much larger than the period of the term it modulates (for horizontal or vertical JTI, the period of modulation $4\pi$ is twice that of a single order of spatial frequency space $2\pi$). The modulation reduces the amplitude of the interference terms. The interference terms, and consequently the desired quantities $\Re\left(SR^{*}\right)$ and $\Im\left(SR^{*}\right)$, are reduced to zero when the appropriate modulation function equals zero, modified from their correct value when the absolute value of the modulation lies between zero and one, and

unmodified when the modulation function is equal to one. The real and imaginary parts of $SR^*$ are indeterminate in the region of frequency space where the appropriate modulation function equals zero. In general, the accuracy of the desired spectra at a point $\mathbf{k} = (k_x, k_y)$ in frequency space will be proportional to the absolute value of the modulation function at that point.

The zeros of the various modulation functions, for the five different classes of JTI, are illustrated in Figures 4-4 through 4-7. In each figure, the real part of the frequency domain, and the corresponding cosinusoidal modulation functions, are shown on the left; the imaginary part of the frequency domain and the sinusoidal modulation functions are shown on the right. The zero-order spectrum is indicated in each plot as the shaded region centered at the origin.

The modulation-function zeros for horizontal and vertical JTI are illustrated in Figure 4-4. The solid lines represent the zeros for horizontal JTI ($\cos(k_y) = 0$ for

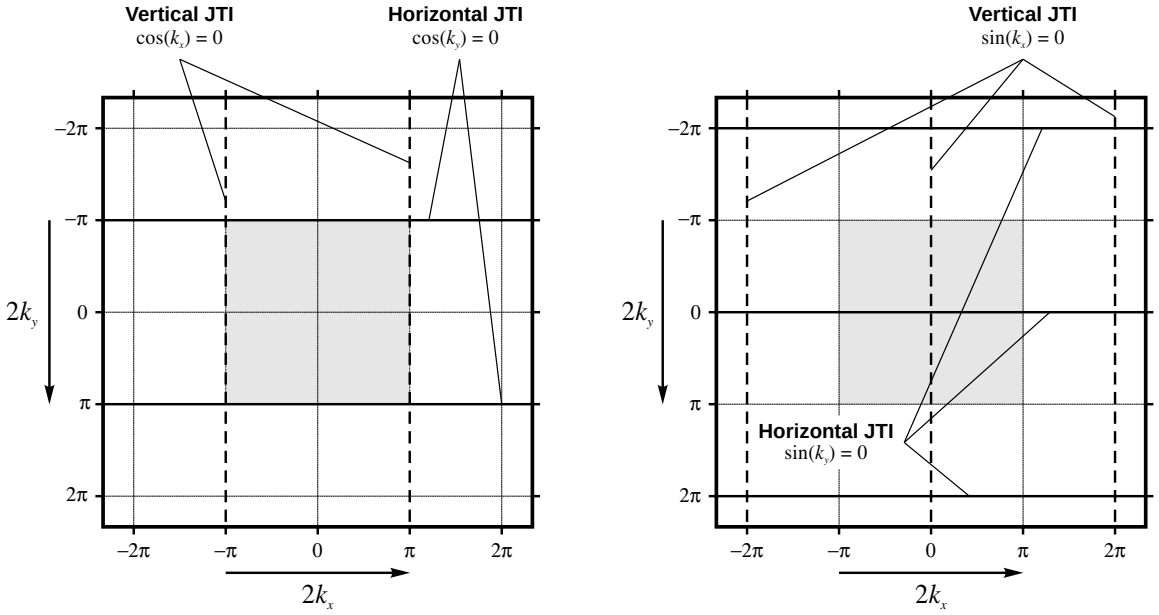



**Figure 4-4** Modulation function zeros for horizontal and vertical JTI: (a) real part of spectrum; and (b) imaginary part of spectrum.

the real part and $\sin(k_y)=0$ for the imaginary part) and the dashed lines represent the zeros for vertical JTI ($\cos(k_x)=0$ for the real part and $\sin(k_x)=0$ for the imaginary part). The cosine modulation functions are maximum at the origin and fall to zero at the boundaries of the zero order spectrum ($2k_y=\pm\pi$ and $2k_x=\pm\pi$ for horizontal and vertical JTI respectively). The sine modulation functions are maximum at the boundaries of the zero order spectrum and fall to zero along the frequency space axes ($2k_y=0$ and $2k_x=0$ for horizontal and vertical JTI respectively). The cosine functions have unit magnitude where the sine functions are zero and vice versa.

The zeros for diagonal JTI are illustrated in Figure 4-5. Diagonal JTI may be either positive or negative and is named after the slope of the interference fringes produced by a single-point source and a single-point reference. The dashed lines represent the zeros for negative diagonal JTI ($\cos(k_x+k_y)=0$ for the real part and

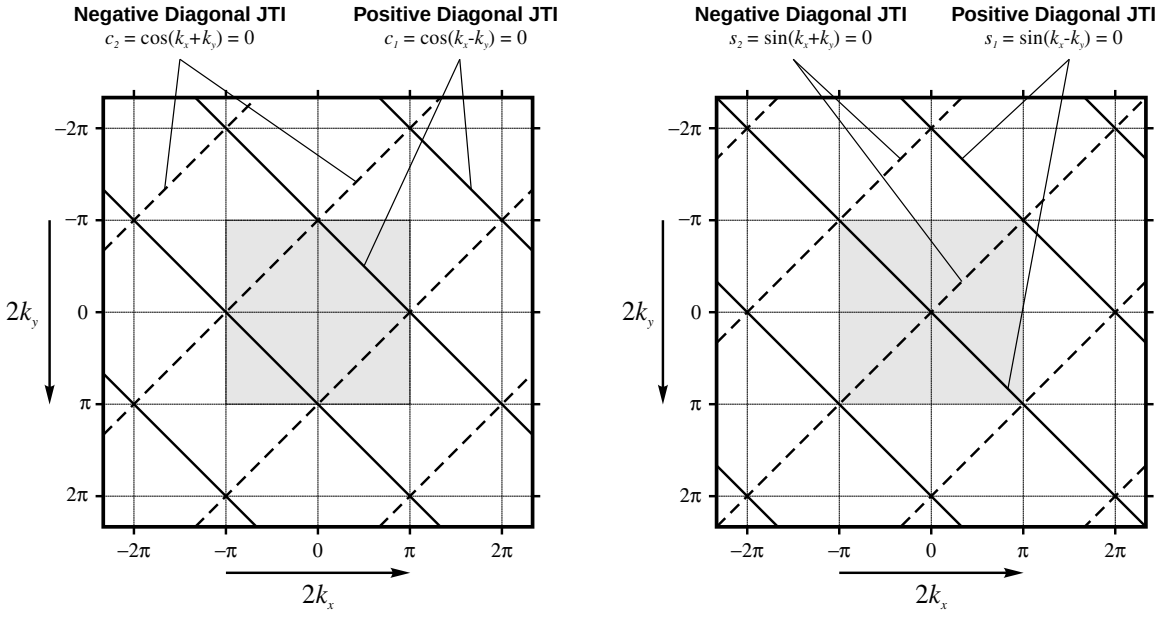



**Figure 4-5** Modulation function zeros for diagonal JTI: (a) real part of spectrum; and (b) imaginary part of spectrum.

$\sin(k_x + k_y) = 0$ for the imaginary part) and the solid lines represent the zeros for positive diagonal JTI ($\cos(k_x - k_y) = 0$ for the real part and $\sin(k_x - k_y) = 0$ for the imaginary part). The cosine functions are zero along diagonal lines which cut through the zero-order spectrum ($2k_x + 2k_y = \pm\pi$ for negative diagonal JTI and $2k_x - 2k_y = \pm\pi$ for positive diagonal JTI). The sine functions are zero along diagonal lines which pass through the origin and four corners of the zero order spectrum ($2k_x + 2k_y = 0, \pm 2\pi$ for negative diagonal JTI and $2k_x - 2k_y = 0, \pm 2\pi$ for positive diagonal JTI). As before, the cosine functions have unit magnitude where the sine functions are zero and vice versa.

The zeros for three-component JTI are illustrated in Figure 4-6. A single cosine modulation function describes the modulation for both the positive and negative cases. The dashed lines in the Figure 4-6(b) represent the sine modulation zeros for negative 3-component JTI and the solid lines represent the

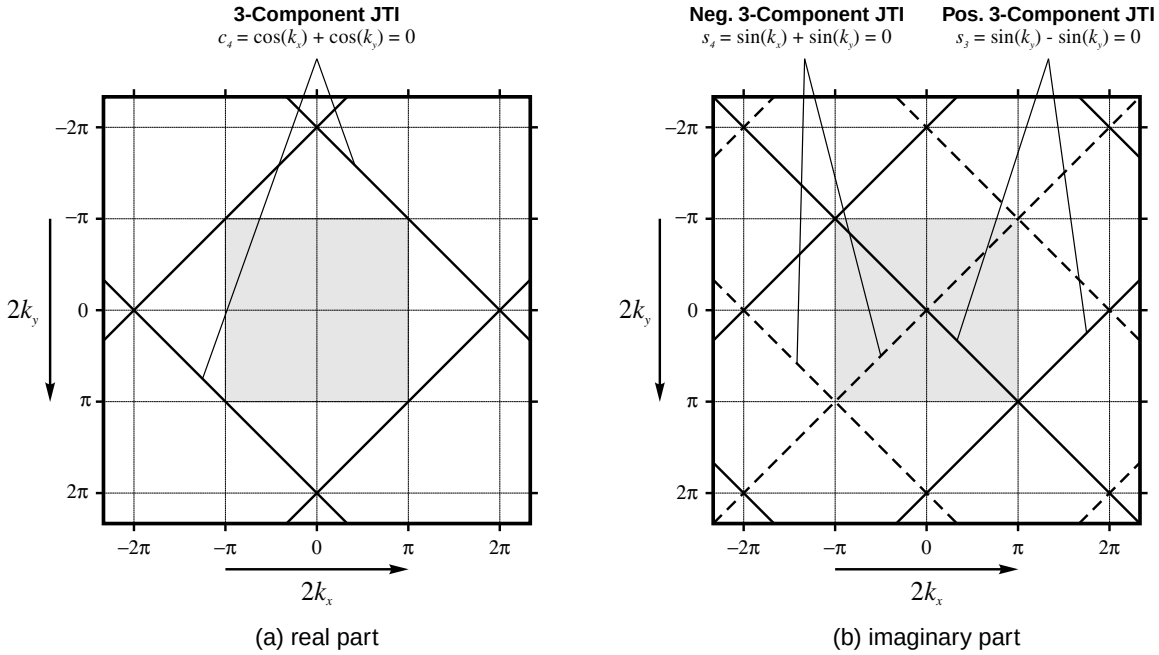
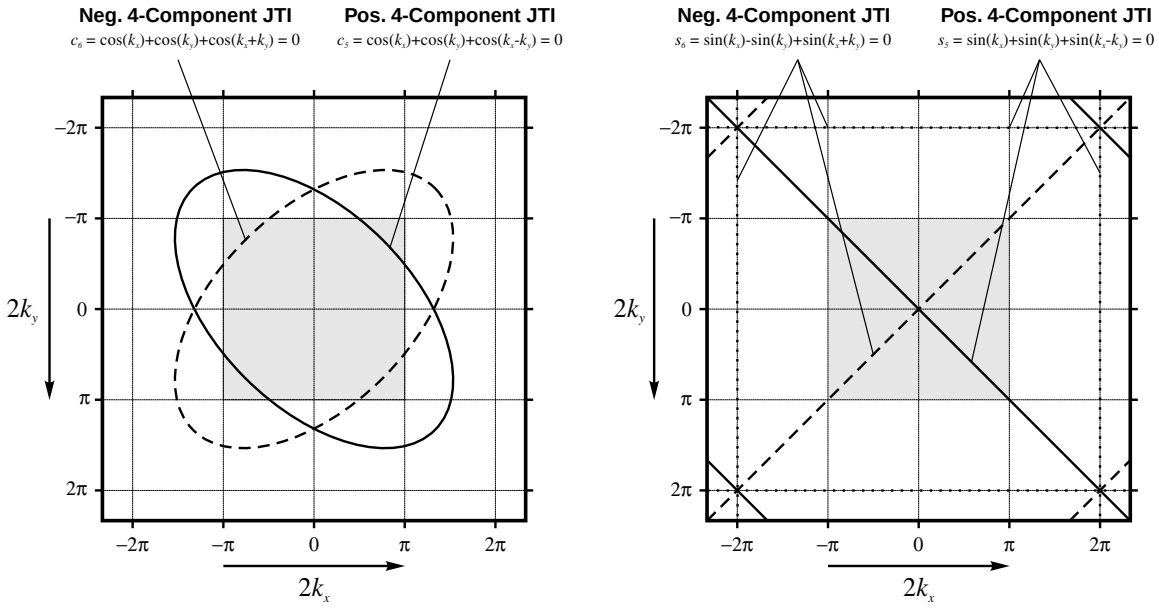



**Figure 4-6** Modulation function zeros for three-component JTI: (a) real part of spectrum; and (b) imaginary part of spectrum.

zeros for the positive case. The cosine functions are zero along diagonal lines which pass through the corners of the zero-order spectrum ($2k_x \pm 2k_y = \pm 2\pi$). For negative three-component JTI, the sine function is zero along a diagonal line through the origin, $2k_x + 2k_y = 0$, and on two perpendicular lines, $2k_x - 2k_y = \pm 2\pi$, through opposite corners of the zero order spectrum. For the positive case, the zeros occur on three similar lines rotated about the origin by $90°$. The real and imaginary modulation functions can be zero along the same lines in frequency space.

The zeros for four-component JTI are illustrated in Figure 4-7. The zeros of the special symmetric case are identical to the zeros illustrated in Figure 4-6(a) for three-component JTI. The dashed lines in Figure 4-7 represent the sine modulation zeros for negative four-component JTI and the solid lines represent the zeros for the positive case. The cosine functions for the negative and positive cases are zero along curves which resemble the loci of points described by two




(a) real part (b) imaginary part

**Figure 4-7** Modulation function zeros for 4-component JTI: (a) real part of spectrum; and (b) imaginary part of spectrum.

ellipses inclined at $\pm 45^{\circ}$ (the curves resemble ellipses but are not true conic sections). The sine functions for the negative and positive cases are zero along negative and positive lines through the origin. In both the cases, the sine function is also equal to zero on the dotted horizontal and vertical lines.

## 4.7 SIGNAL-TO-NOISE RATIO

The signal-to-noise ratio (SNR) of the FT spectra extracted using the JTI technique is derived in this section. In general, the SNR is a function of spatial frequency. The SNR at a particular frequency is determined by the magnitude of the source and reference spectra, the magnitude of the trigonometric modulation function, and the noise introduced during the modulation, optical FT, and detection processes. The detector noise is assumed to be independent of position on the spatial frequency plane. Its magnitude is determined from the dynamic range of the interferograms and the number of bits employed by the frame grabber to quantize them.

The objective of this section is to identify how the dynamic range of the detector array contributes to the SNR of the computed spectra. The development of a rigorous noise model and a complete SNR analysis is beyond the scope of this dissertation. Many simplifying assumptions have been made and various noise sources have been omitted in order to simplify the analysis and make the mathematical description tractable. Quantization noise introduced during the detection process is the only source of noise considered in this analysis. It is assumed that the quantization noise is spatially and temporally uncorrelated with itself and the detected signal. Speckle noise introduced by the SLM, which largely determines the overall SNR of a coherent imaging system and must be modeled as a multiplicative noise source, is not considered in this analysis. The assumptions described above, motivated primarily by mathematical tractability and the overall objective of the analysis, lead to a very simplified description of source spectra SNR.

In order to derive expressions for the SNR, assumptions must be made about the maximum intensity of the source and reference spectra. Specifically, the functional form of the reference spectrum and its maximum intensity must be known with respect to that of the source spectrum. A region of the zero-order spectrum, called the region of interest (ROI), is identified such that the dynamic range of its intensity distribution falls within that of the square-law detector. The reference spectrum is assumed to have uniform intensity (or amplitude) and phase in this ROI. The reference intensity which results in the most efficient use of the detector's dynamic range is determined and used to derive expressions for the optimum SNR.

### 4.7.1 Dynamic Range and Region of Interest

In this section, expressions for the SNR at an arbitrary spatial frequency are derived for the real and imaginary parts of the source image spectrum. As one might expect, the SNR increases with the number of bits used to quantize the interferograms and decreases with the interferogram's dynamic range. The dynamic range of a typical interferogram is much larger than that of most conventional detector arrays. Consider, for example, a $256\times256$ pixel binary image in which half of the pixels are active. Assume the dynamic range of the interferogram is given by the ratio of intensity produced by a single pixel to that due to the entire interferogram, observed at zero spatial frequency. The optical dynamic range is then $20\log(0.5\cdot256\cdot256)=90\,\text{dB}$ and its electrical dynamic range is twice that, or $180\,\text{dB}$. The electrical dynamic range of a typical commercial-grade CCD camera is about $50\ \text{dB}$. Obviously the CCD camera doesn't have enough dynamic range to accurately measure the entire zero-order intensity spectrum.

It is often the case that only a specific region on the spatial frequency plane is of particular interest. The trajectory estimation algorithm presented in the previous chapter required spatial frequency information at points on a circle

centered on the origin. In these cases, an ROI (region of interest), for example an annular region centered on the origin, can be identified on the spatial frequency plane such that the dynamic range of the spectrum in the ROI does not exceed that of the detector array. The detector is calibrated to *just saturate* at the maximum intensity in the ROI. Pixels on the detector array outside the ROI, in particular the large DC spike at the origin, will saturate the detector. The anti-blooming circuitry in the detector must accommodate the saturation and prevent charge from spilling into adjacent pixels in the ROI. Interferograms containing saturated pixels are not processed. In the unlikely event that a pixel inside the ROI becomes saturated, the particular interferogram is flagged, the CCD integration time is reduced, and the set of interferograms are measured again. The dynamic range of the intensity spectrum is effectively reduced by only measuring the intensity in a specific ROI which does not include the DC spike.

Selection of the ROI is application specific. The set of spatial frequencies at which the complex-valued FT is required determines how the ROI on the spatial frequency plane is defined. The reference image is then defined such that its FT will produce a desired intensity distribution in the ROI. In the general case, where the complex-valued FT is required at all spatial frequencies, the total spectrum must be assembled from a set of subspectra. The total spectrum is divided into a small set of subspectra and a suitable set of reference images are constructed. Each reference image is interfered with the source image using the JTI technique and the resulting subspectra are assembled to provide the total spectrum.

### 4.7.2 Source and Reference Spectra

The intensity distribution measured over the ROI is an interferogram produced by the source and reference images. The maximum intensity in the ROI depends on the maximum intensity of the individual spectra and their phase relationship. The maximum intensity determines the noise contribution that will be uniformly added to each pixel in the ROI. The interference term contains the real and imaginary

parts of the source image spectrum, however its amplitude is modulated by the magnitude of the reference spectrum. The reference spectrum must be chosen to maximize the magnitude of the interference term such that it fills the available dynamic range of the detector array. It is assumed that the reference spectrum has uniform phase and constant intensity equal to the maximum intensity of the source spectrum in the ROI. It is difficult to design a reference image whose spectrum satisfies these criteria over the entire zero-order spectrum; however, it is possible to design a reference image which satisfies these criteria in the ROI.

Assume that the reference spectrum has uniform amplitude and phase over some ROI on the spatial frequency plane. The ROI may include the entire zero-order spectrum, or a specific portion of it. The FT of the displaced source and reference images were presented in Equation 4-1 as the RHS of a FT pair. They are reproduced below with equivalent expressions written in magnitude and phase form,

$$\begin{aligned} S(\mathbf{k})\exp(i\,\mathbf{p}_s\cdot\mathbf{k}) &= \sqrt{I_s(\mathbf{k})}\exp(i\varphi_s'(\mathbf{k})) \\ R\exp(i\,\mathbf{p}_r\cdot\mathbf{k}) &= \sqrt{I_r}\exp(i\varphi_r'(\mathbf{k})) \end{aligned} \qquad (4\text{-}30)$$

where $I_s(\mathbf{k})$ and $I_r$ represent the intensity (or power) spectra of the displaced source and reference, and $\varphi_s'(\mathbf{k})$ and $\varphi_r'(\mathbf{k})$ represent their phase spectra. The total phase spectra is made up of two components: 1) the phase of the non-displaced source and reference spectra, $\varphi_s(\mathbf{k})$ and $\varphi_r$; and 2) the linear phase shifts, $\mathbf{p}_s\cdot\mathbf{k}$ and $\mathbf{p}_r\cdot\mathbf{k}$, due to the source-reference displacement. These two contributions sum to give $\varphi_s'(\mathbf{k})$ and $\varphi_r'(\mathbf{k})$, respectively. The reference spectrum is assumed uniform with constant phase in the ROI and so $R$, $I_r$, and $\varphi_r$ shown no $\mathbf{k}$ dependence.

### 4.7.3 Saturation Intensity

The maximum intensity in the region of interest depends on the source and reference intensity spectra and their phase relationship. The source intensity spectrum will have some maximum value, $I_s^{max}$ in the ROI, at a spatial frequency $\mathbf{k}_{max}$. The complex amplitude at the point of maximum intensity $\mathbf{k} = \mathbf{k}_{max}$ can then be represented in magnitude and phase form as,

$$S_{max} = \sqrt{I_s^{\max}} \exp[i\, \varphi_s'(\mathbf{k}_{\max})] \tag{4-31}$$

The intensity spectrum of the source-reference joint image is detected by a CCD camera or some other square-law detector. In order to make the most efficient use of the detector's dynamic range, the maximum signal intensity must *just saturate* the output of the camera.

The maximum intensity of the source spectrum $I_s^{max}$ occurs at $\mathbf{k} = \mathbf{k}_{max}$. The source-reference joint image will have its maximum intensity at the same spatial frequency if the source and reference complex amplitudes are in phase (the reference spectrum has uniform magnitude and phase). If the two spectra are not in phase, the maximum intensity may occur at some $\mathbf{k} \neq \mathbf{k}_{max}$. In general, the intensity at $\mathbf{k} = \mathbf{k}_{max}$ is given by,

$$I(\varphi') = |S_{max} + R|^2 = I_s^{max} + I_r + 2\sqrt{I_s^{max} I_r}\cos(\varphi') \tag{4-32}$$

and depends on the phase difference between the source and reference spectra at the point of maximum intensity, $\varphi' = \varphi_s'(\mathbf{k}_{max}) - \varphi_r'(\mathbf{k}_{max})$. When the source and reference spectra are exactly in phase, they add constructively and the intensity detected at $\mathbf{k} = \mathbf{k}_{max}$ is given by $I_{max} = I(0)$. If the source and reference spectra are out of phase, the maximum intensity may occur at some $\mathbf{k} = \mathbf{k}'_{max}$, however its value will be less than $I_{max}$ because $S(\mathbf{k}'_{max}) < S_{max}$ and the magnitude of the

reference is constant. Therefore, the maximum intensity of the source-reference joint spectra will always be less than or equal to $I_{max}$. In order to avoid possible saturation when the source and reference spectra are in phase, the camera must be set to *just saturate* when the incident intensity is equal to $I_{max}$.

### 4.7.4 Optimum Reference Intensity

The intensity detected at $\mathbf{k} = \mathbf{k}_{max}$, as described by Equation 4-32, can be considered as a sinusoidal information signal $2\sqrt{I_s^{max} I_r}\cos(\varphi')$ superimposed on an optical bias $I_s^{max} + I_r$, as illustrated in Figure 4-8. In order to make the most efficient use of the CCD camera's available dynamic range, the intensity of the reference should be chosen to maximize the depth of modulation,

$$\mathcal{V} = \frac{2\sqrt{I_s^{max} I_r}}{I_s^{max} + I_r} \tag{4-33}$$

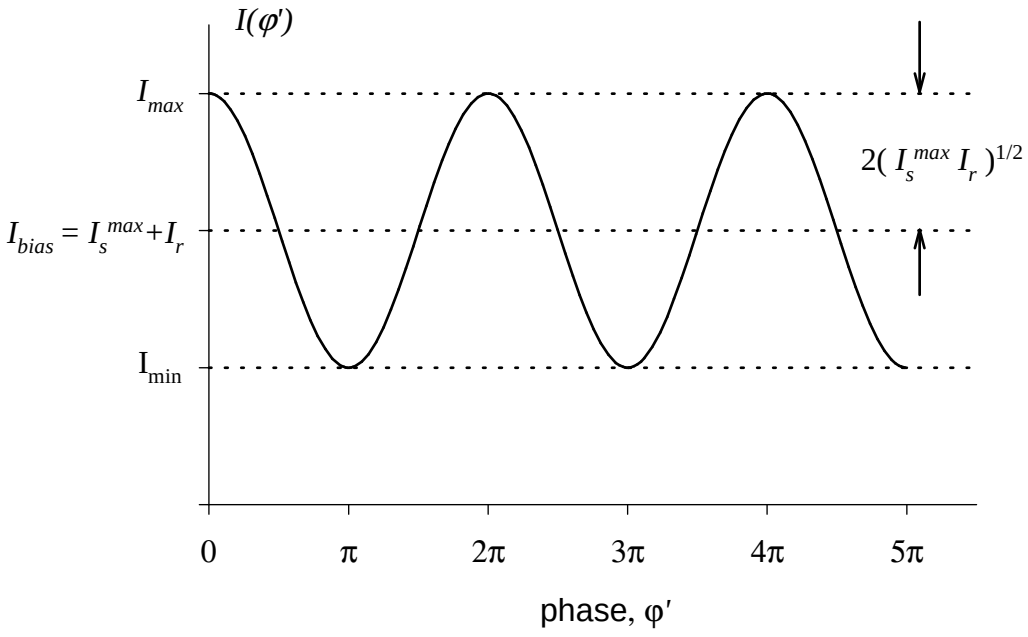



**Figure 4-8** Interference effect illustrated as a optical bias modulated by a cosinusoidal information signal.

of the signal. The depth of modulation is identical to the interference fringe visibility between two arbitrary coherent waves (Equation 2-14, Chapter 2). It was shown that the fringe visibility of the interferogram produced by two coherent waves is maximum when their intensities are equal (section 2.3.1). The depth of modulation is therefore maximum when the intensity of the reference is equal to the maximum intensity of the source, $I_s^{max} = I_r$. When the source and reference intensities are equal, $\mathcal{V} = 1$, and the detected intensity fills the entire range of available intensities between $0$ and $I_{max} = 4I_s^{max} = 4I_r$. In order to make the most efficient use of the camera's dynamic range, and avoid possible saturation when the source and reference are in phase, the intensity of the reference spectrum should be set to the maximum intensity of the source spectrum, $I_r = I_s^{max}$, and the camera should be set to just saturate at four times the reference intensity, $I_{max} = 4I_s^{max} = 4I_r$.

Figure 4-9 illustrates the normalized joint intensity detected at an arbitrary spatial frequency $\mathbf{k}_0$ when the saturation and reference intensities are optimized as described above. The source intensity is set equal to $I_s^{max}$, $\frac{1}{4}I_s^{max}$, $\frac{1}{16}I_s^{max}$, and $\frac{1}{64}I_s^{max}$, in Figures 4-9(a) through 4-9(d), respectively, and the phase varies from $0$ to $5\pi$ in each figure. The detected joint intensity is normalized with respect to the maximum source intensity. Figure 4-9(a) shows how the entire dynamic range of the camera is utilized as the source and reference spectra drift in and out of phase when $I_s = I_s^{max}$. As the intensity of the source contribution is reduced (Figures 4-9(b) through 4-9(d)), the bias intensity is reduced proportionately and the amplitude of the information signal decreases as its square root. The bias approaches $I_r$ as the source contribution approaches zero.

The reference spectrum is assumed to have a uniform spatial distribution in the ROI such that it has constant amplitude and phase. Any spatial or temporal variation in the power of the beam incident on the SLM will affect the power

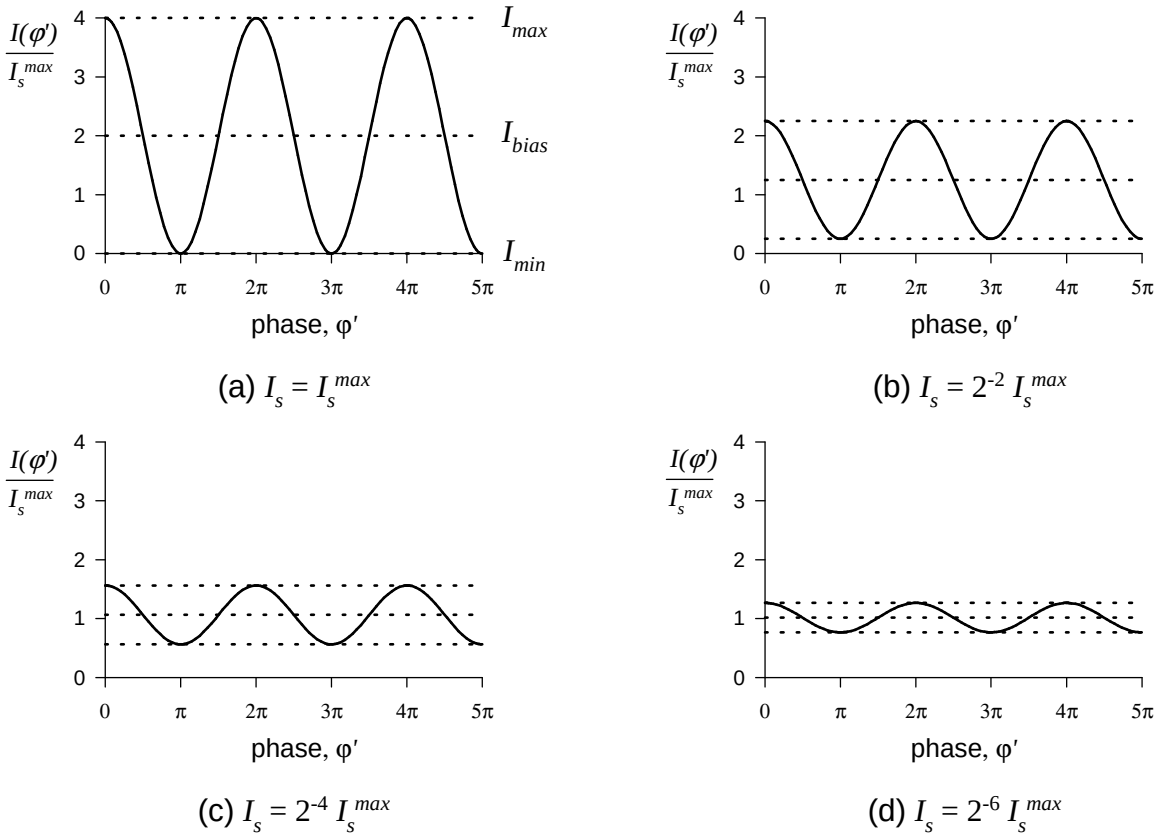



**Figure 4-9** Normalized intensity variation as a function of phase for different source intensities: (a) maximum intensity; (b) ¼ of maximum intensity; (c) 1⁄16 of maximum intensity; and (d) 1⁄64 of maximum intensity.

distribution in the intensity spectra and the interferograms. These variations will lead to additional noise in the computed source spectra.

### 4.7.5 Noise in Detection

The SNR of the square-law detector is determined by the thermal and shot noise added during detection and the number of bits used to quantize the analog video signal. Assume that quantization noise is much larger than both thermal and shot noise such that the camera is quantization-noise limited. This assumption is valid when eight bits are used to quantize the analog video signal. The detector is quantization-noise limited, however, the entire optical Fourier transform system is most likely speckle-noise limited.

Assume that the noise introduced by the quantization process can be modeled as an additive noise source (this assumption is not always valid, in particular, when the signal is very weak compared to the quantizer's dynamic range). The quantized intensity measured at a particular pixel during frame $n$ can be represented by,

$$\hat{I}(n) = I(n) + e(n) \tag{4-34}$$

where $I(n)$ is the sequence of exact intensities and $e(n)$ is the quantization error sequence. If a quantizer with $b$ bits of resolution is employed ($2^b$ possible levels of gray) and the analog intensity is quantized by rounding to the nearest gray level, the quantization error is bounded by,

$$-\frac{\Delta I}{2} < e(n) \leq \frac{\Delta I}{2} \tag{4-35}$$

where $\Delta I = 2^{-b}$ is the quantization "bin" width. The quantization error is not known for all $n$ and so a statistical model is usually employed [O&S75, page 413]. It is assumed that the error sequence $\{ e(n) \}$ is a sample sequence from a stationary random white-noise process whose probability distribution is uniform over the range described by Equation 4-35. Furthermore, it is assumed that the analog intensity sequence is normalized in order to avoid clipping distortion such that its maximum value is less than $1 - 2^{-b}$. Under these conditions, the quantization error has zero mean and variance, $\sigma^2 = \frac{1}{12} 2^{-2b}$.

The analog intensity sequence generally exceeds the dynamic range of the quantizer and must be reduced in amplitude to avoid clipping. Assume that the intensity sequence is normalized by dividing through by $I_{max}$, the maximum intensity in the ROI. In this way the entire dynamic range of the quantizer is utilized and there is no chance of any clipping distortion. If the variance in the normalized intensity sequence is $\sigma^2$, the variance in the original non-normalized

sequence $I(n)$ is $I_{max}^2 \sigma^2$. The mean-squared error in the intensity sequence is then given by,

$$E\left[e^2(n)\right] = I_{max}^2 \sigma^2 = I_{max}^2 \frac{2^{-2b}}{12} \quad (4\text{-}36)$$

where $E[\ ]$ denotes a temporal expectation or ensemble average.

The objective of the noise analysis presented in this section is to determine the temporal SNR of the intensity time sequence observed at a single arbitrary point $\mathbf{k}_0 = (k_{x,0}, k_{y,0})$ on the spatial frequency plane. The variation of the SNR with spatial frequency is not investigated. The temporal error sequence associated with some arbitrary point $\mathbf{k}_1$ is assumed to be uncorrelated with that associated with another point $\mathbf{k}_2$. This assumption is valid when the frequency-domain points are non-adjacent and mutually separated as in the case of the mixed-domain trajectory estimation algorithm (the points are separated and uniformly distributed around a circle on the spatial frequency plane). If the frequency domain points constituted a fully sampled region or sub-image within the ROI, a more general analysis would have to consider the correlation between the temporal error sequences realized at each of the points.

### 4.7.6 Imaginary Part SNR

Several algorithms were presented in Section 4.3 to extract the imaginary part of the source spectrum from two source-reference joint spectra designed to have opposite displacement. A more general algorithm is illustrated in Figure 4-10 as a block diagram. A source-reference joint spectrum with positive displacement phase, $S + \exp(i\phi)R$, is detected at some arbitrary spatial frequency by a single photodetector (PD) on the square-law detector array. Quantization noise from the error sequence $\{ e_1(n) \}$ is added to the detected intensity to yield $I_\phi + e_1$ and the result is stored temporarily. A source-reference joint spectrum with negative

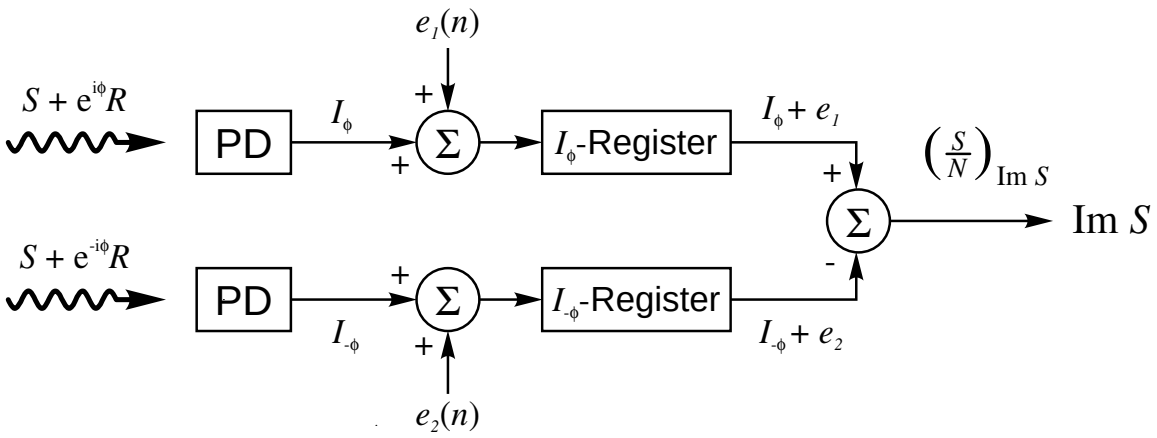



**Figure 4-10** Block diagram illustrating the extraction of the imaginary part of the source spectrum from two source-reference joint spectra with opposite displacement phase.

displacement phase, $S + \exp(-i\phi)R$, is detected at the same spatial frequency by the same photodetector at a later time. Quantization noise from the error sequence $\{e_2(n)\}$ is added to the detected intensity to yield $I_{-\phi} + e_2$ and the result is subtracted from the previously stored intensity. The difference between the two intensities represents one point of the modulated imaginary part of the source spectrum.

The expressions for the source-reference joint spectra with positive and negative displacement phase depend on the class of JTI employed. In general, they can be expressed in the follow form,

$$\begin{aligned} I_\phi &= I_s + I_r + 2\sqrt{I_r}\,[\,c(\mathbf{k}_0)\Re(S) + s(\mathbf{k}_0)\Im(S)\,] \\ I_{-\phi} &= I_s + I_r + 2\sqrt{I_r}\,[\,c(\mathbf{k}_0)\Re(S) - s(\mathbf{k}_0)\Im(S)\,] \end{aligned} \tag{4-37}$$

where $c(\mathbf{k}_0)$ and $s(\mathbf{k}_0)$ represent the cosine and sine modulation terms at an arbitrary spatial frequency $\mathbf{k}_0 = (k_{x,0}, k_{y,0})$, and the modulation terms which pre-multiply the individual source and reference spectra have been ignored. The modulation terms are functions of spatial frequency and their functional forms depend on the class of JTI. The reference spectrum is assumed to be real valued

in addition to having uniform intensity in the ROI (the reference spectrum is real valued when the reference image is even). The difference between the two source-reference joint spectra at the output of the block diagram illustrated in Figure 4-10 is then given by,

$$\left(I_{\phi}+e_1\right) \;-\; \left(I_{-\phi}+e_2\right) \;=\; 4\sqrt{I_r}\; s(\mathbf{k}_0)\,\Im(S) \;+\; e_1-e_2 \qquad (4\text{-}38)$$

The electrical SNR is calculated in the conventional manner as the mean electrical signal power over the mean electrical noise power. The squared signal and noise are averaged temporally over a large number of image frames. The SNR can be written directly from Equation 4-38,

$$\mathrm{SNR}_{\mathrm{Im}S} \;=\; \frac{E\left[\left(4\sqrt{I_r}\; s(\mathbf{k}_0)\,\Im(S)\right)^2\right]}{E\left[(e_1-e_2)^2\right]} \;=\; \frac{16\, s^2(\mathbf{k}_0)\, I_r E\left[\Im^2(S)\right]}{2\, I_{max}^2\, \frac{2^{-2b}}{12}} \qquad (4\text{-}39)$$

The optimal SNR is achieved when the depth of modulation (fringe visibility) is equal to unity. It was shown in section 4.7.4 that when $I_{\max}=4I_s^{\max}=4I_r$, the depth of modulation is maximum and there is no chance of saturation in the ROI when the source and reference spectra are in phase. If the optical processor is calibrated to satisfy these conditions, the optimal SNR can be written as,

$$\mathrm{SNR}_{\mathrm{Im}S} \;=\; 6\, s^2(\mathbf{k}_0)\, 2^{2b}\, \frac{E\left[\Im^2(S)\right]}{\left|S_{\max}\right|^2} \qquad (4\text{-}40)$$

The expression for SNR has three distinct components. Moving from left-to-right through Equation 4-40, it is clear that SNR is proportional to the square of the sinusoidal modulation function, dynamic range of the frame grabber, and the ratio of the mean-squared signal to the maximum source image power in the ROI. The sinusoidal modulation and mean-squared signal are functions of spatial frequency; the noise added by the frame grabber, due to its limited dynamic range, is independent of spatial frequency.

### 4.7.7 Real Part SNR

A general algorithm is illustrated in Figure 4-11 to extract the real part of the source spectrum from individual source and reference spectra and two source-reference joint spectra designed to have opposite displacement phase. The joint spectra are detected and quantization noise is added in exactly the same manner that was employed during the calculation of the imaginary part. The quantized joint-intensities are stored for further processing. The individual source spectrum $S$ is detected at the same spatial frequency by the same photodetector. Quantization noise from the error sequence $\{e_3(n)\}$ is added to the detected intensity to yield $I_s + e_3$ and the result is also stored. Similarly, the individual reference spectrum $R$ is detected at the same spatial frequency, quantization noise from $\{e_4(n)\}$ is added to yield $I_r + e_4$, and the result is stored. Finally, twice the sum of the individual intensities is subtracted from the sum of the joint

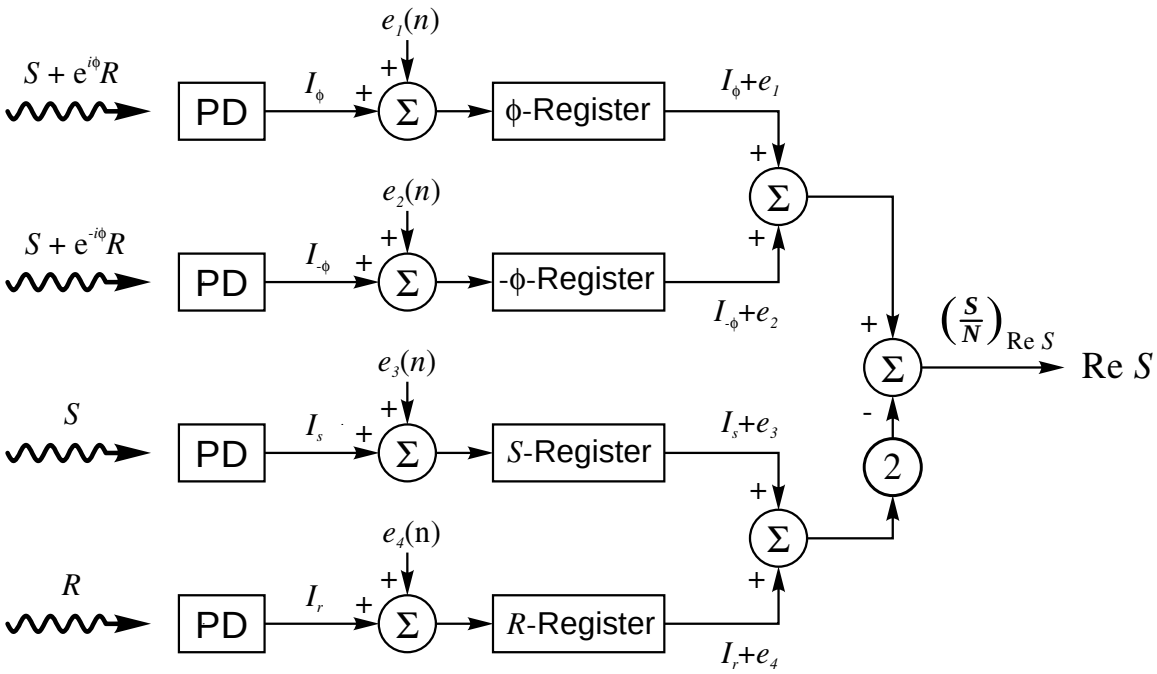



**Figure 4-11** Block diagram illustrating the extraction of the real part of the source spectrum from two source-reference joint spectra with opposite displacement phase and their individual spectra.

intensities to yield one point of the modulated imaginary part of the source spectrum.

General expressions for the source-reference joint spectra were presented in Equation 4-37, however, the modulation functions which pre-multiply $I_s$ and $I_r$ were ignored for clarity. It is assumed that the individual spectra, determined in the lower two stages of the block diagram of Figure 4-11, are modulated in the same fashion that $I_s$ and $I_r$ were in Equation 4-37 (this applies when the joint image is composed of more than two component images). The intensity at the output of the block diagram of Figure 4-11 is then given by,

$$\begin{aligned} \left(I_{\phi}+e_1\right) \;+\; \left(I_{-\phi}+e_2\right) \;-\; & 2\left[\left(I_s+e_3\right)+\left(I_r+e_4\right)\right] \\ = \; & 4\sqrt{I_r}\; c\left(\mathbf{k}_0\right)\Re\left(S\right) \;+\; \left[e_1+e_2-2\left(e_3+e_4\right)\right] \end{aligned} \tag{4-41}$$

As before, the electrical SNR is calculated as the mean electrical signal power over the mean electrical noise power. The SNR can be written directly from Equation 4-41,

$$\mathrm{SNR}_{\mathrm{Re}S} \;=\; \frac{E\left[\left(4\sqrt{I_r}\; c\left(\mathbf{k}_0\right)\Re\left(S\right)\right)^2\right]}{E\left[\left(e_1+e_2-2e_3-2e_4\right)^2\right]} \;=\; \frac{16\, c^2\left(\mathbf{k}_0\right) I_r E\left[\Re^2\left(S\right)\right]}{10\, I_{max}^2\, \frac{2^{-2b}}{12}} \tag{4-42}$$

The optimal SNR is achieved when the depth of modulation (fringe visibility) is equal to unity. Setting $I_{\max}=4I_s^{\max}=4I_r$, the optimal SNR can be written as,

$$\mathrm{SNR}_{\mathrm{Re}S} \;=\; \frac{6}{5}\, c^2\left(\mathbf{k}_0\right) 2^{2b}\, \frac{E\left[\Re^2\left(S\right)\right]}{\left|S_{\max}\right|^2} \tag{4-43}$$

The SNR of the real part is also proportional to the dynamic range of the frame grabber and to the ratio of the mean-squared signal to the maximum source image power in the ROI. The SNR of the real part is proportional to the square of the cosinusoidal modulation rather than to the square of the sinusoidal

modulation. This SNR is five time less than that of the imaginary part because of the additional quantization error introduced by the two individual spectra.

### 4.7.8 Maximizing the SNR

The dynamic range of the detector is fixed and is generally much less than that of the source spectra over the entire zero-order spatial frequency plane. The SNR is maximized by choosing a ROI such that the ratio of the mean squared signal to the maximum source image power, in the ROI, is large. This generally requires that the ROI exclude the zero frequency spike at the origin of the spatial frequency plane.

The SNR is reduced when the square of the modulation function is reduced from its maximum value. In fact, the SNR equals zero when the modulation function equals zero. The functional form of the modulation terms depends on the class of JTI employed. The modulation function zeros for the five classes of JTI are illustrated in Figures 4-4 through 4-7. If the spatial frequencies of interest are known in advance, the class of JTI can be picked to avoid zeros near the frequencies of interest. Alternatively, a weighted sum of two spectra, determined from two appropriate classes of JTI, can be calculated.

## 4.8 CHAPTER SUMMARY

A novel method of coherent detection suitable for an interferometric optical Fourier processor was presented in this chapter. The new technique, called JTI (joint transform interference), is a multi-point extension of the point-diffraction method employed in Chapter 3. It provides two significant advantages over the previous technique: 1) the power of the source and reference images can be easily matched in order to maximize the SNR of the FT spectra; and 2) both the real and imaginary parts of the complex-valued FT spectra are recovered.

The principle of JTI is simple. The interferogram produced by a source image and reference image, mutually displaced on the same SLM, is measured by a square-law detector in the back focal-plane of a Fourier transform lens. The real and imaginary parts of the FT spectra are extracted by adding and subtracting interferograms produced by two composite source-reference images with different mutual displacements. The real and imaginary parts of the source image FT are modulated, respectively, by cosinusoidal and sinusoidal functions.

Several classes of JTI, suitable for implementation on a binary SLM, were identified and presented. Composite or joint images were constructed from individual source, reference, and zero images. The composite images were placed into one of five classes, depending on the direction of the mutual displacement (either horizontal, vertical, or diagonal) and the number of source and reference images used to construct the composite (either two, three, or four components). The interferograms were described by expressions similar to the classical two-wave interference equation. Algorithms for extracting the real and imaginary parts of a source image FT from two interferograms with opposite displacement phase were presented for each JTI class.

The functional form of the interferogram modulation terms and their zeros were presented and discussed. An interferogram produced by a horizontally (vertically) displaced source-reference image is modulated by sinusoidal image

with vertical (horizontal) lines of constant phase. Similarly, an interferogram produced by a diagonally displaced source-reference image is modulated by sinusoidal image with diagonal lines of constant phase.

The SNR (signal-to-noise ratio) of the real and imaginary FT spectra extracted using the JTI technique was derived in the last section. It was shown that the intensity of the reference spectrum should be equal to the maximum intensity of the source spectrum, and that the CCD detector should be set to just saturate at four times the reference intensity, in order to make the most efficient use of the camera's dynamic range, and avoid possible saturation when the source and reference are in phase. The SNR was shown to be proportional to 1) the dynamic range of the frame grabber; and 2) the ratio of the mean-squared signal to the maximum source image power in the ROI (region of interest).

CHAPTER

# 5

# Multiple-Object Trajectory Estimation

A moving-object trajectory estimation system which estimated the speed and direction of single objects was demonstrated in Chapter 3. The system employed an optical Fourier processor and a PDI (point-diffraction interferometer) to determine the real part of the moving object's frequency domain representation. The point-diffraction method of coherent detection restricted the number of pixels in each image frame and lead to an ambiguity in the estimation of target direction. The JTI (joint transform interference) coherent detection technique developed in Chapter 4 solves these two problems. An improved multiple-object trajectory estimation system based on the JTI method of coherent detection is presented in this chapter. The new system employs a digital micromirror device (DMD) as its SLM (spatial light modulator).

## 5.1 INTRODUCTION

A hybrid optical-digital trajectory estimation system based on the JTI method of coherent detection is presented in this chapter. The original system, described previously in Chapter 3 and based on the point-diffraction method, was unable to accommodate multipixel image frames and could not determine the direction with which objects moved along the estimated trajectory. These problems are solved by the system presented in this chapter.

The problem of trajectory estimation, the advantages of spatiotemporal, mixed, and frequency-domain solutions, and the motivation behind a mixed-domain optical implementation, were discussed previously in Section 3.1. The specific type of moving-object trajectory estimation problem addressed in this dissertation is the detection and trajectory estimation of small, barely discernible, moving objects of unknown position and velocity. An optical implementation is of particular computational interest as typically more than ten high resolution image frames must be integrated to detect and track targets in a noisy environment. As discussed previously, the mixed-domain technique has several advantages over a purely spatiotemporal or frequency-domain approach.

An ideal moving point-object describes a line in space-time and a plane in 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 multiple objects is determined by estimating the slope of an equivalent number of frequency-space planes. An optical Fourier processor and the JTI method of coherent detection are employed in the trajectory estimation system to optically determine the complex-valued 2-D spatial FT (Fourier transform) of each frame in an image sequence. The sequence of complex-valued spectra are processed in the time domain using 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 using a time-domain linear-prediction method. The temporal frequencies estimated at a pre-selected set of spatial frequencies, constitute a set of frequency-triples which describe multiple planes in 3-D frequency space. The slope of each plane is easily determined from the set of frequency-triples to estimate the velocity components of each moving object.

A Texas Instruments digital micromirror device (DMD) is employed as the spatial light modulator (SLM) in the optical Fourier processor. The DMD is a $640\times480$ pixel array of individually addressable micromirrors, each of which can reflect light in one of two possible directions. The *on-state* blazing effect of the mirrors introduces a linear phase distribution over the plane of the pixel array which mutually displaces the (coherent) optical transfer function (OTF) and the set of interference maxima. This effect, combined with the source-spectrum replication introduced by the JTI zero-interleaving, results in a replicated spectrum which is not contaminated by contrast-ratio noise and coincides almost exactly with the maximum absolute value of the OTF.

The DMD pixel structure, switching time, and its far-field diffraction pattern are reviewed in Section 5.2. Several different classes of joint transform interference (JTI) were derived previously in Chapter 4. The weighted three-component JTI method is experimentally verified in Section 5.3. The remainder of the chapter is devoted to the application of the JTI method of coherent detection to the problem of trajectory estimation. The JTI method is employed in Section 5.4 with a single-pixel reference image to estimate the trajectories of single and multiple moving point-objects. In Section 5.5, a multipixel Bessel-function reference image is used to estimate the trajectory of multiple $3\times3$ pixel objects moving over a zero-intensity background, different synthetic backgrounds, and in the presence of Gaussian white noise. The trajectories of multiple single-pixel objects in white noise are estimated in the final section.

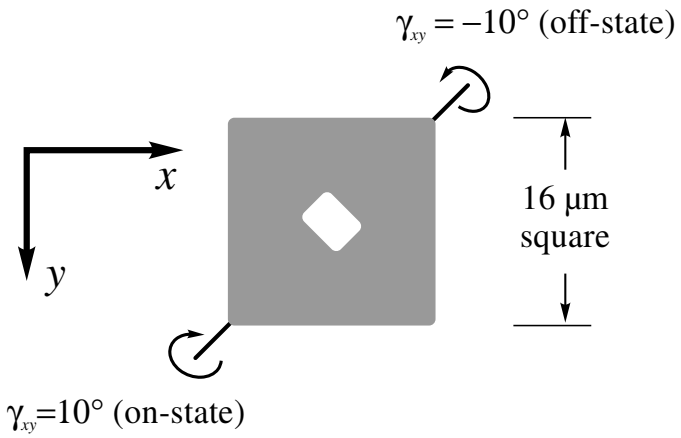
## 5.2 THE DIGITAL MICROMIRROR DEVICE (DMD)

The DMD (digital micromirror device) is a reflective SLM manufactured by Texas Instruments (TI) for digital projection and hardcopy applications [Hor95]. The device is a superstructure array of movable aluminum micromirrors monolithically integrated over a SRAM (static random access memory) circuit. Each micromirror in the superstructure is located over its own SRAM unit cell corresponding to one bit of memory. The mirrors are electrostatically bistable: a mirror in its *on-state* reflects incident light in some desired direction, towards a viewing screen for example, while a mirror in its *off-state* redirects incident light elsewhere. The mirror's state is toggled by changing the potential difference between it and one of two electrodes below the mirror. Each mirror can be individually addressed; the state of a particular mirror is switched by toggling the appropriate bit in SRAM. The micromirror superstructure array is constructed entirely of aluminum using standard semiconductor deposition and etching techniques. The SRAM is fabricated on a silicon substrate using standard CMOS (complementary metal oxide silicon) technology.

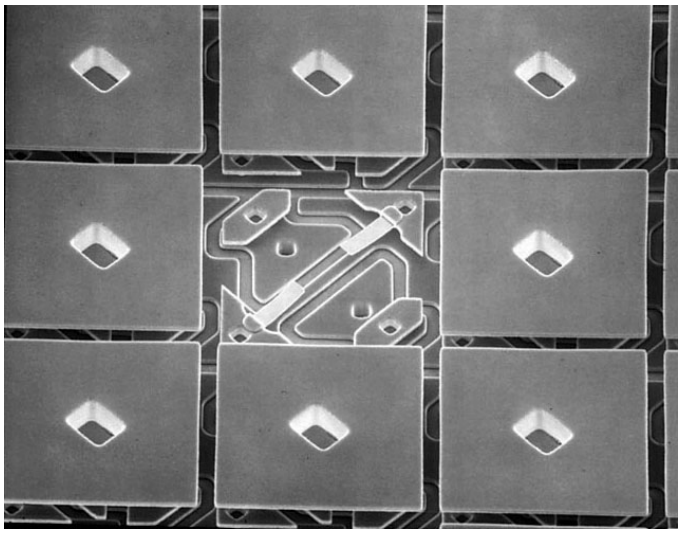
The DMD was designed primarily for projection display applications, however it has also been used for optical signal processing. The initial DMDs fabricated by TI were called deformable mirror (rather than digital micromirror) devices. These initial DMDs have been demonstrated by various researchers as amplitude modulators [Flo90], phase mostly modulators [Gre88, Flo90, Boy91], and full complex modulators [Flo91]. They have also been employed as SLMs in joint transform correlators [Flo89, Hud89, Kno89], VanderLugt correlators [Flo88, Hud89], and optical neural networks [Col89]. Kearney recently investigated employing a newer (digital micromirror) DMD as an optical mask in precision imaging and spectroscopy applications [Kea98]. MacAulay has proposed substituting the mechanical diaphragms found in conventional microscopes with several (digital micromirror) DMDs to provide dynamic control over illumination angle and intensity for applications in quantitative microscopy [MacA98].

### 5.2.1 Pixel Structure and Switching Time

The DMD is a $640 \times 480$ pixelated SLM. A micromirror at each pixel location selectively reflects incident light either into the pupil of a lens or towards an optical stop. The DMD pixel structure is illustrated in Figure 5-1. The square mirrors at each pixel have dimension $\Delta' = 16\,\mu\text{m}$ and center-to-center spacing




(a)



(b)

**Figure 5-1** DMD pixel structure: (a) switching principle of a single mirror; and (b) SEM photomicrograph of a $3 \times 3$ mirror array (SEM reproduced with permission from Texas Instruments [TI98]).

$\Delta = 17\ \mu\text{m}$ (there is $1\ \mu\text{m}$ of dead space between each pixel, the spatial duty cycle is $a = b = 0.94$, and the fill factor is approximately $90\ \%$). The switching principle is illustrated in Figure 5-1(a). Mirrors toggle between their two electromechanically bistable states by rotating about the diagonal axis shown in the figure. A mirror is off when $\gamma_{xy} = -10°$, on when $\gamma_{xy} = +10°$, and parked when $\gamma_{xy} = 0°$. An SEM (scanning electron microscope) photomicrograph of a $3 \times 3$ mirror array is illustrated in Figure 5-1(b). The center mirror has been removed to reveal the torsion hinge and addressing electrodes; the surrounding mirrors are in their parked state. The mechanical switching time (defined to be the time between a reset pulse and the instant when the mirrors have landed and settled to a level where they are electromechanically latched) is approximately $15\ \mu\text{s}$ [Hor95].

### 5.2.2 The DMD as a Blazed Diffraction Grating

The FT of a pixelated SLM was expressed in Section 2.3.1 as the product of its OTF (optical transfer function) and its DSFT (discrete space Fourier transform). It was shown that the maximum absolute value of the OTF and the zero spatial frequency occurred at the same point in the far-field diffraction pattern (see Figures 2-4 and 2-5). The maximum absolute value of the OTF and zero spatial frequency do not coexist at the same position in the DMD produced diffraction pattern: the position of the maximum absolute value of the OTF is shifted with respect to that of the DC spike.

The absolute value of the OTF (which is equivalent to the diffraction envelope) describes the distribution of maximum possible intensity in the DMD-produced far-field diffraction pattern. The absolute value of the OTF has a peak which is spatially shifted from the DC spike (zero-order interference maximum). This effect is sometimes introduced by design in blazed diffraction gratings by shaping the ridges of the grating in order to increase their efficiencies at higher interference orders [Hec87, pp. 424-427]. The same principle is employed in phased-array radar and sonar signal processing [Bro85]. In these systems, a

linear phase distribution is introduced across the plane of the antenna in order to steer the beam in a desired direction. In terms of the DMD, the pixels are blazed (or inclined) at an angle $\gamma$ and the light is made incident on the DMD at twice the blaze angle to introduce the phase shift. The blazing is built-in due to the inherent tilt of the mirrors. The linear phase distribution introduced by the mirror's tilt, and the corresponding frequency-domain shift, are quantified in the following sections.

*5.2.2.1 Linear Phase Distribution*

A series of optical wavefronts is illustrated in Figure 5-2 illuminating a row of *on-state* DMD pixels. The light is incident at an angle $\theta_i$ and reflected at an angle, $\theta_r = 2\gamma - \theta_i$, where $\gamma$ is the pixel blaze angle (projection of the mirror tilt on to either the $x$ or $y$ axis) and the angles of the incident and reflected beams are measured with respect to the DMD normal. The pixel blaze introduces a phase shift between adjacent pixels. It is clear from Figure 5-2 that the optical path difference between two rays, reflected by adjacent pixels, is given by $\overline{\mathrm{BC}} - \overline{\mathrm{CD}}$. An expression for $\overline{\mathrm{BC}}$ can be determined from triangle ABC, specifically, $\overline{\mathrm{BC}} = \Delta \sin(\theta_i)$, and similarly, an expression for $\overline{\mathrm{CD}}$ can be determined from triangle CDE, $\overline{\mathrm{CD}} = -\Delta \sin(\theta_r)$, where the negative sign on the second expression was used to be consistent with the chosen coordinate system. The phase difference between rays reflected by adjacent pixels is simply $2\pi\lambda^{-1}$ multiplied by the optical path difference,

$$\phi = \frac{2\pi\left(\overline{\mathrm{BC}} - \overline{\mathrm{CD}}\right)}{\lambda} = \frac{2\pi\Delta}{\lambda}\left[\sin(\theta_i) + \sin(2\gamma - \theta_i)\right] \qquad (5\text{-}1)$$

The preceding equation is a direct extension of the grating equation for oblique incidence [Bor75, p. 405]. The derivation was presented here to illustrate the underlying phase-shift effect more clearly. Equation 5-1 shows that the linear

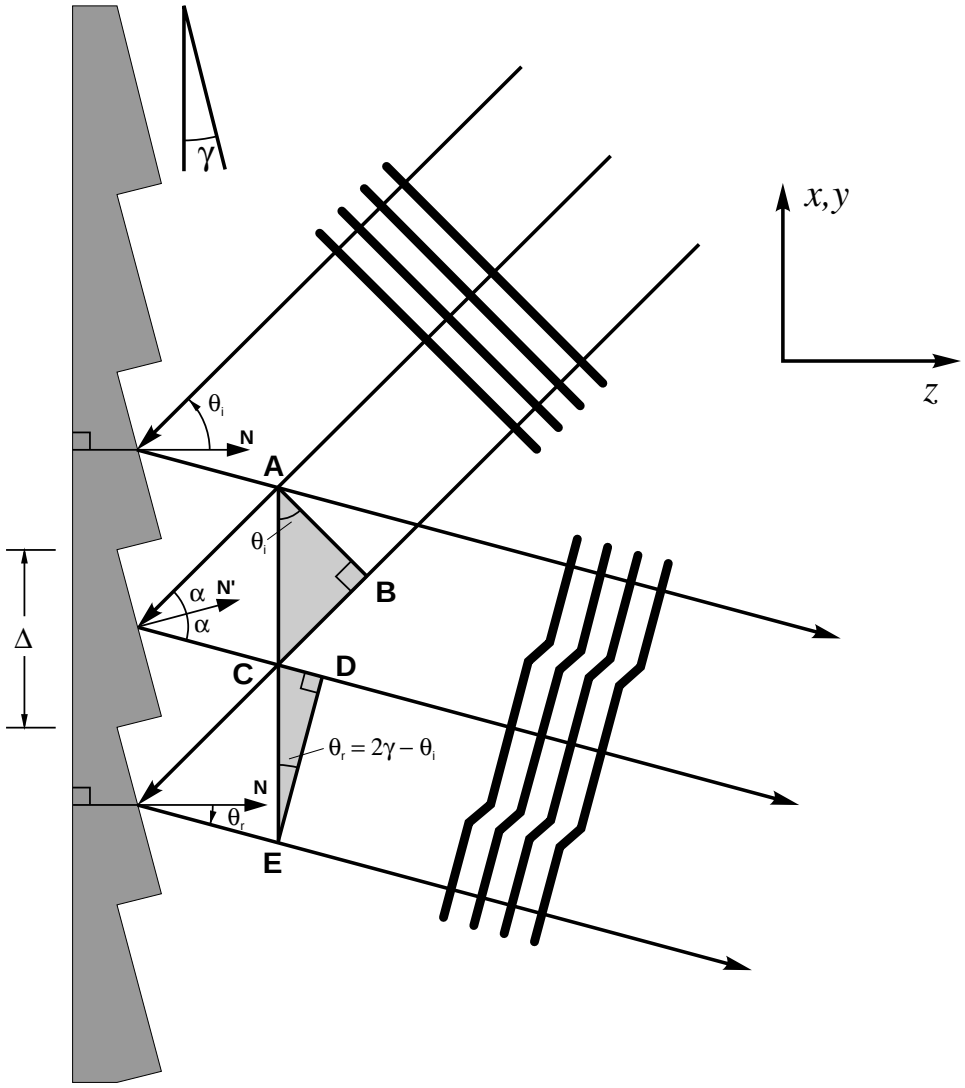



**Figure 5-2** A row of DMD on-state mirrors as a blazed diffraction grating.

phase distribution introduced by the DMD is a function of the angle of incidence and the pixel blaze angle.

For the case of minimum spherical distortion (the optical FT is valid only in the paraxial approximation), the DMD normal and the optical axis of the FT lens must be coincident, ie $\theta_r = 0$. Therefore, the angle of incidence must be twice the blaze angle, $\theta_i = 2\gamma$, and consequently the expression for the phase difference

simplifies to yield, $\phi = 2\pi\,\Delta\,\lambda^{-1}\sin(2\gamma)$. The phase difference under these conditions is easily calculated at the He-Ne laser wavelength using values for the DMD center-to-center pixel spacing and the blaze angle,

$$\phi \;=\; 2\pi\,\frac{17\ \mu\text{m}}{0.6328\ \mu\text{m}}\sin\!\left(\frac{2\cdot 10^{\circ}}{\sqrt{2}}\right) \;=\; 6.564\cdot 2\pi \quad \text{rad} \tag{5-2}$$

An *on-state* DMD mirror is tilted at angle $\gamma_{xy} = +10^{\circ}$ around the diagonal axis shown in Figure 5-1(a). The projection of this angle onto either the $x$ or $y$-axis, as required by Equation 5-2, was consequently expressed by $\gamma = \gamma_{xy}\big/\sqrt{2} \approx 7^{\circ}$.

*5.2.2.2 DMD Diffraction Pattern*

The Fourier transform of a pixelated SLM is the product of its (coherent) optical transfer function (OTF) and the discrete space Fourier transform (DSFT) of the displayed image sequence (see Section 2.3.1). The light which illuminates the DMD has a built-in linear phase distribution due to the oblique angle of incidence and the pixel blaze. If the image sequence written to the DMD is given by $f(m,n)$, the optical distribution immediately after the modulation is given by $\exp[i(\phi m + \phi n)]\,f(m,n)$. The linear phase distribution in the spatial domain causes a shift in the frequency domain,

$$\exp[i(m\phi + n\phi)]\,f(m,n) \quad\Leftrightarrow\quad F(k_x - \phi, k_y - \phi) \tag{5-3}$$

where the magnitude of the shift is given by Equation 5-2. The position of the OTF in the frequency domain is unchanged (its peak intensity is determined by the specular reflection from the mirrors). The phase difference between DMD pixels effectively slides the periodic spectrum $F(k_x, k_y)$ under the OTF by an amount $\phi = 6.564\cdot 2\pi$ rad, in both the $k_x$ and $k_y$ directions, as shown in Figure 5-3. The interference maximum corresponding to the zero-order zero frequency of $f(m,n)$ and the maximum absolute value of the OTF do not coincide as they did

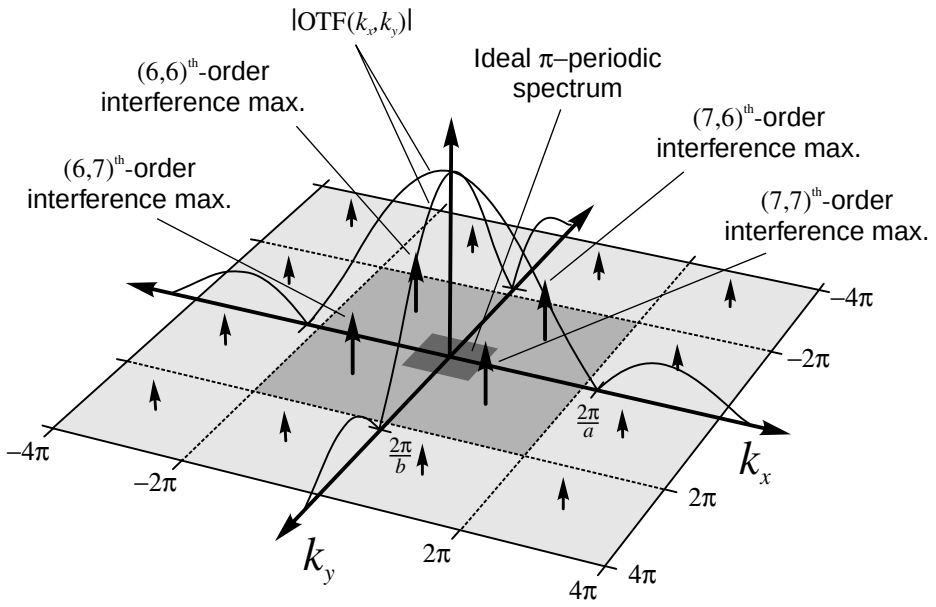



**Figure 5-3** DMD far-field diffraction pattern illustrating the absolute value of its OTF (diffraction envelope) and replicated DSFT spectra (2π-periodic interference maxima).

in Figure 2-4. As shown in Figure 5-3, the absolute value of the OTF is maximum at a point approximately half way between the four DC spikes (interference maxima) belonging to the $(6,6)^{\text{th}}$, $(6,7)^{\text{th}}$, $(7,6)^{\text{th}}$, and $(7,7)^{\text{th}}$-order spectra. The optical energy is concentrated around the Nyquist frequency $(\pm\pi,\pm\pi)$ rather than a DC spike (interference maxima). This phenomenon has also been observed and described by Kearney [Kea98].

*5.2.2.3 Advantage for JTI*

The finite contrast of all SLMs generally causes the region of the spatial frequency plane surrounding a DC spike to be very noisy. Non-modulated light reflected (or diffracted) by the DMD due to its dead zone between pixels and the mirror support posts (see Figure 5-1) results in a finite contrast ratio. The non-modulated light is Fourier transformed into the DC region as contrast-ratio noise. The pixel blazing on the DMD shifts the DC spikes out of the region of concentrated intensity and effectively reduces the noise energy. The

implementation of the JTI algorithm on a binary SLM results in the compression of spatial frequency space by a factor of two due to the upsampling and stretching of the space domain (see Section 4.3). This frequency domain compression produces additional $\pi$-periodic (replicated and compressed) spectra mid-way between each interference maxima. These new spectra are not contaminated by the contrast-ratio noise associated with the spectra centered on each interference maxima. In addition, and most importantly, the new noise-free spectrum corresponding to the fictitious $(6.5,6.5)^{th}$-order interference maximum, shown in Figure 5-3 as the ideal $\pi$-periodic spectrum, is approximately at the same position as the maximum absolute value of the OTF. This spectrum is free of contrast-ratio noise at DC and receives the maximum diffracted intensity.

## 5.3 EXPERIMENTAL VERIFICATION OF JTI

Several JTI (joint transform interference) algorithms were proposed in the previous chapter. The purpose of this section is to experimentally verified one of those algorithms and show that it provides the correct real and imaginary parts of a source image FT. Although calculating the entire FT spectrum of an image using the JTI technique may be computationally inefficient, the experiment is presented to illustrate, at least qualitatively, that the technique correctly provides the complex-valued spectra of an image, something which is not as obvious if only selected points on the spatial frequency plane are determined.

Five different classes of JTI were defined in Section 4.3. The classes depended on the mutual source-reference displacement (horizontal, vertical, or diagonal) and the number of source and reference image components. In this section, the three-component JTI algorithm (see Section 4.3.3) is experimentally verified using single-pixel source and reference images. Trivial images were employed for two reasons: 1) the FT of a single-pixel reference is constant and therefore the calculated spectra do not require normalization by the reference

spectrum; and 2) the source and reference power spectra are matched (equal) which guarantees good fringe visibility.

### 5.3.1 Demodulation and Spectral Weighting Functions

The algorithm for three-component JTI is described by Equations 4-18 and 4-19. The equations express the (modulated) imaginary and real parts of a source image spectrum in terms of several joint and individual intensity spectra. The spatial frequencies at which the modulated spectra equal zero, referred to as the modulation function zeros, were illustrated in Figure 4-6. The real part of the modulated zero-order spectrum equals zero at its four corners; the imaginary part equals zero on a positive diagonal through the origin for positive three-component JTI, and similarly, equals zero on the negative diagonal for negative JTI. In order to calculate a source image spectrum with correct imaginary values on both diagonals, the positive and negative modulated spectra are calculated and their weighted sum is calculated on a point-by-point basis to yield an imaginary spectrum with only one indeterminate point at the origin. The positive and negative modulated spectra are combined in an equally weighted sum to yield an average real spectrum with a indeterminate point in each corner. The weighting functions are chosen to demodulate the positive and negative modulated-spectra, and at the same time, weight their sum depending on the position of the modulation zeros.

#### *5.3.1.1 Real Part*

An expression for the real part of the modulated source-spectrum using negative three-component (double source and single reference) JTI was presented on the first line of Equation 4-19. This expression is repeated below with its companion for positive three-component JTI,

$$\begin{aligned} \alpha_{neg}(k_x,k_y) &= I_{\substack{0\,s\\ s\,r}} + I_{\substack{r\,s\\ s\,0}} - 2I_{\substack{0\,s\\ s\,0}} - 2I_r \\ \alpha_{pos}(k_x,k_y) &= I_{\substack{s\,r\\ 0\,s}} + I_{\substack{s\,0\\ r\,s}} - 2I_{\substack{s\,0\\ 0\,s}} - 2I_r \end{aligned} \tag{5-4}$$

The functions $\alpha_{neg}(k_x,k_y)$ and $\alpha_{pos}(k_x,k_y)$ represent the cosine-modulated real part of the source-image spectrum, one calculated using the negative algorithm, and the other calculated using the positive algorithm ($R^* = 1$ for a single-pixel reference image). The modulated spectra are demodulated and averaged in an equally weighted sum on a point-by-point basis,

$$\Re\{S(2k_x,2k_y)\} = w_1\,\alpha_{neg}(k_x,k_y) + w_2\,\alpha_{pos}(k_x,k_y) \tag{5-5}$$

to yield the real part of the source image spectrum. In the special case of three-component JTI, the positive and negative weighting functions are equal,

$$w_1 = w_2 = \frac{1}{8c_4} = \frac{1}{8(\cos(k_x)+\cos(k_y))} \tag{5-6}$$

and the weighting is actually an average of the two demodulated spectra. The identical weighting functions were selected by inspection to demodulate each spectrum by dividing out the cosine modulation,

$$\begin{aligned} & w_1\,\alpha_{neg}(k_x,k_y) + w_2\,\alpha_{pos}(k_x,k_y) \\ &= \frac{1}{8c_4} 4c_4\,\Re\{S(2k_x,2k_y)\} + \frac{1}{8c_4} 4c_4\,\Re\{S(2k_x,2k_y)\} \\ &= \Re\{S(2k_x,2k_y)\} \end{aligned} \tag{5-7}$$

### *5.3.1.2 Imaginary Part*

An expression for the imaginary part of the modulated source-spectrum was presented on the first line of Equation 4-18. This expression is repeated below with its companion for positive three-component JTI,

$$\begin{aligned} \beta_{neg}(k_x,k_y) &= I_{\substack{0\,s\\ s\,r}} - I_{\substack{r\,s\\ s\,0}} \\ \beta_{pos}(k_x,k_y) &= I_{\substack{s\,r\\ 0\,s}} - I_{\substack{s\,0\\ r\,s}} \end{aligned} \tag{5-8}$$

Both these expressions represent the sine-modulated imaginary part of the source-image spectrum, the former calculated via the negative JTI algorithm, and the latter calculated using the positive algorithm. The modulated spectra are demodulated and averaged in a weighted sum on a point-by-point basis,

$$\Im\left(S\left(2k_x,2k_y\right)\right) = w_3\,\beta_{neg}(k_x,k_y) + w_4\,\beta_{pos}(k_x,k_y) \tag{5-9}$$

to yield the imaginary part of the source image spectrum. The weighting functions,

$$\begin{aligned} w_3 &= \frac{s_4}{4\left(s_3^2+s_4^2\right)} = \frac{\sin(k_x)+\sin(k_y)}{8\left(\sin^2(k_x)+\sin^2(k_y)\right)} \\ w_4 &= \frac{s_3}{4\left(s_3^2+s_4^2\right)} = \frac{\sin(k_x)-\sin(k_y)}{8\left(\sin^2(k_x)+\sin^2(k_y)\right)} \end{aligned} \tag{5-10}$$

are selected by inspection to demodulate and weight the spectra. Substitution of the weighting functions from Equation 5-10 and the JTI spectra from Equation 4-18 into Equation 5-9 illustrates how the modulated spectra are weighted and demodulated to yield the imaginary part of the spectrum,

$$\begin{aligned} & w_3\beta_{neg}(k_x,k_y) + w_4\beta_{pos}(k_x,k_y) \\ &= \frac{s_4}{4\left(s_3^2+s_4^2\right)}4s_4\Im\left(S\left(2k_x,2k_y\right)\right) + \frac{s_3}{4\left(s_3^2+s_4^2\right)}4s_3\Im\left(S\left(2k_x,2k_y\right)\right) \\ &= \Im\left(S\left(2k_x,2k_y\right)\right)\left[\frac{s_4^2}{s_3^2+s_4^2} + \frac{s_3^2}{s_3^2+s_4^2}\right] \\ &= \Im\left(S\left(2k_x,2k_y\right)\right) \end{aligned} \tag{5-11}$$

The weighting functions in Equation 5-10 were selected by inspection and other variations are possible. The 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 ($w_3 = 0$) is assigned to $\beta_{neg}(k_x, k_y)$ along the negative diagonal, $k_x = -k_y$, where its demodulated value is indeterminate (zero over zero), and a very small weight is assigned at points near the diagonal. Similarly, zero weight ($w_4 = 0$) is assigned to $\beta_{pos}(k_x, k_y)$ on the positive diagonal, $k_x = k_y$, where its demodulated value is indeterminate.

### 5.3.2 Computational Efficiency

When a large number of spatial frequency-domain points is required, the technique of JTI followed by point-by-point interferogram processing is computationally inefficient compared to 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 required 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, they can be computed directly or by employing a recursively pruned radix-$(2 \times 2)$ 2-D FFT algorithm [Knu93b]. The number of operations required for the direct computation of a single point is $O(N^2)$ while 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 is required and $O(N^2 \log_2 N)$ for all the points (neglecting the frequency 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.

### 5.3.3 Experimental Setup and Procedure

A block diagram of the experimental setup is illustrated in Figure 5-4. A He-Ne laser beam is passed through a spatial filter and beam expander assembly, composed of objective lens L1 (10X objective, $f_1$=14.8 mm), pinhole aperture PH ($\phi$=50 μm diameter), and collimating lens L2 (5X objective, $f_2$=25 mm), to produce an expanded and filtered beam approximately $2w_1$=890 μm in diameter. The expanded beam is redirected by mirrors M1 and M2 and made incident on the DMD at $2\gamma_{xy} = 20°$ to its normal ($2\gamma_{xy}/\sqrt{2} \approx 14°$ projected onto the $x$ and $y$ axes) such that it is reflected normally when the DMD mirrors are in their *on-state*. The modulated light distribution in the plane of the DMD is Fourier transformed by lens FTL ($f$=125 mm) and its resulting intensity spectrum is measured by a CCD camera located in the back focal plane of the lens. The same basic setup is employed by all experiments presented in this chapter.

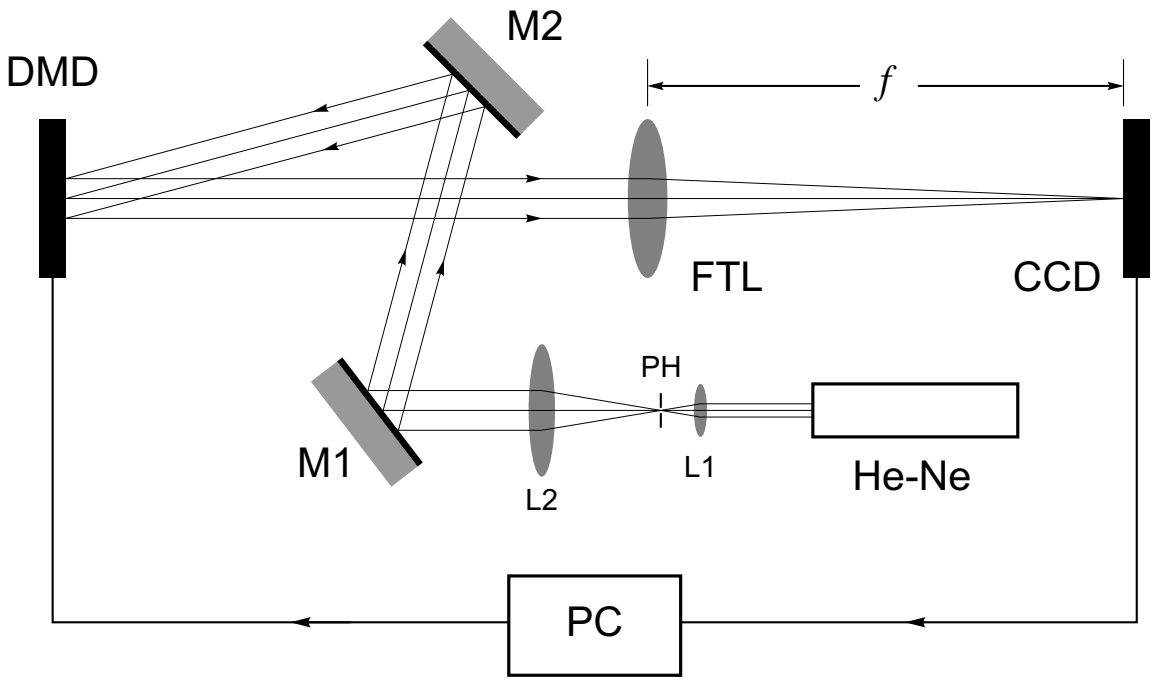



**Figure 5-4** Experimental setup employed to verify the JTI method of coherent detection and for the trajectory estimation experiments.

The three-component JTI technique was experimentally verified using several single-pixel source images. The reference image was the same for each source image and consisted of a single *on-pixel* at the origin. The sequence of seven composite images required by the three-component algorithm were constructed from each source image and the single-pixel reference image. Each composite image in the sequence was optically Fourier transformed using the apparatus illustrated in Figure 5-4 and the corresponding sequence of intensity spectra were stored for further processing. The CCD camera measured the intensity of the ideal $\pi$-periodic (replicated and compressed) spectrum, located mid-way

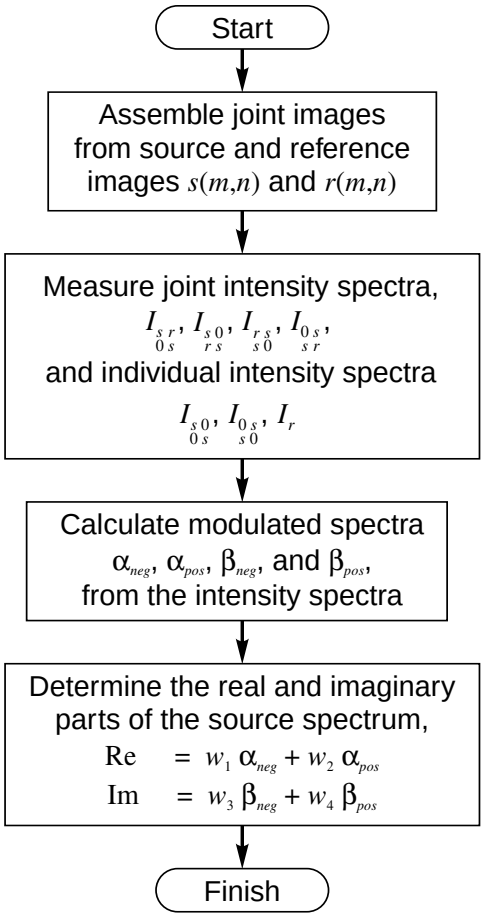



**Figure 5-5** Experimental method and post-processing procedure employed to verify the weighted three-component JTI algorithm.

between the $6^{\text{th}}$ and $7^{\text{th}}$-order interference maxima of the Fraunhofer diffraction pattern, as shown in Figure 5-3. The real and imaginary parts of the source image spectrum were calculated using Equations 5-5 and 5-8 respectively. The experimental method and post-processing procedure are shown in Figure 5-5.

### 5.3.4 Results and Discussion

The real and imaginary parts of a source image consisting of a single on-pixel at $(m,n)=(3,3)$ is illustrated in Figure 5-6. The source image and its DSFT pair are described by,

$$s(m,n)=\begin{cases}1 & m=n=3\\ 0 & \text{otherwise}\end{cases} \quad \Leftrightarrow \quad S(\tilde{k}_x,\tilde{k}_y)=\cos\left[3(\tilde{k}_x+\tilde{k}_y)\right]+i\sin\left[3(\tilde{k}_x+\tilde{k}_y)\right] \quad (5\text{-}12)$$

where $(\tilde{k}_x,\tilde{k}_y)$ represents the compressed frequency space $(2k_x,2k_y)$. The real and imaginary parts of the spectrum are shown in parts (a) and (b) of Figure 5-6, respectively. The real and imaginary amplitudes along the $k_x$ axis are illustrated in parts (c) and (d), and similarly, the real and imaginary amplitudes along the $k_y$ axis are shown in parts (e) and (f). The real and imaginary parts exhibit the quadrature phase relationship (quarter cycle or $\frac{\pi}{2}$ phase shift) described by Equation 5-12 and there are exactly three sinusoidal cycles per $2\pi$ rad along each direction. The effects of the four indeterminate points, due to the zeros of the cosine modulation-functions, can be seen in the four corners of the real spectrum. Similarly, the effect of the single indeterminate point, due to the sine modulation, can be identified at the origin of the imaginary spectrum.

The complex spectrum of a second test image is illustrated in Figure 5-7. The source image and its DSFT pair are described by,

$$s(m,n)=\begin{cases}1 & m=n=10\\ 0 & \text{otherwise}\end{cases} \quad \Leftrightarrow \quad S(\tilde{k}_x,\tilde{k}_y)=\cos\left[10(\tilde{k}_x+\tilde{k}_y)\right]+i\sin\left[10(\tilde{k}_x+\tilde{k}_y)\right] \quad (5\text{-}13)$$

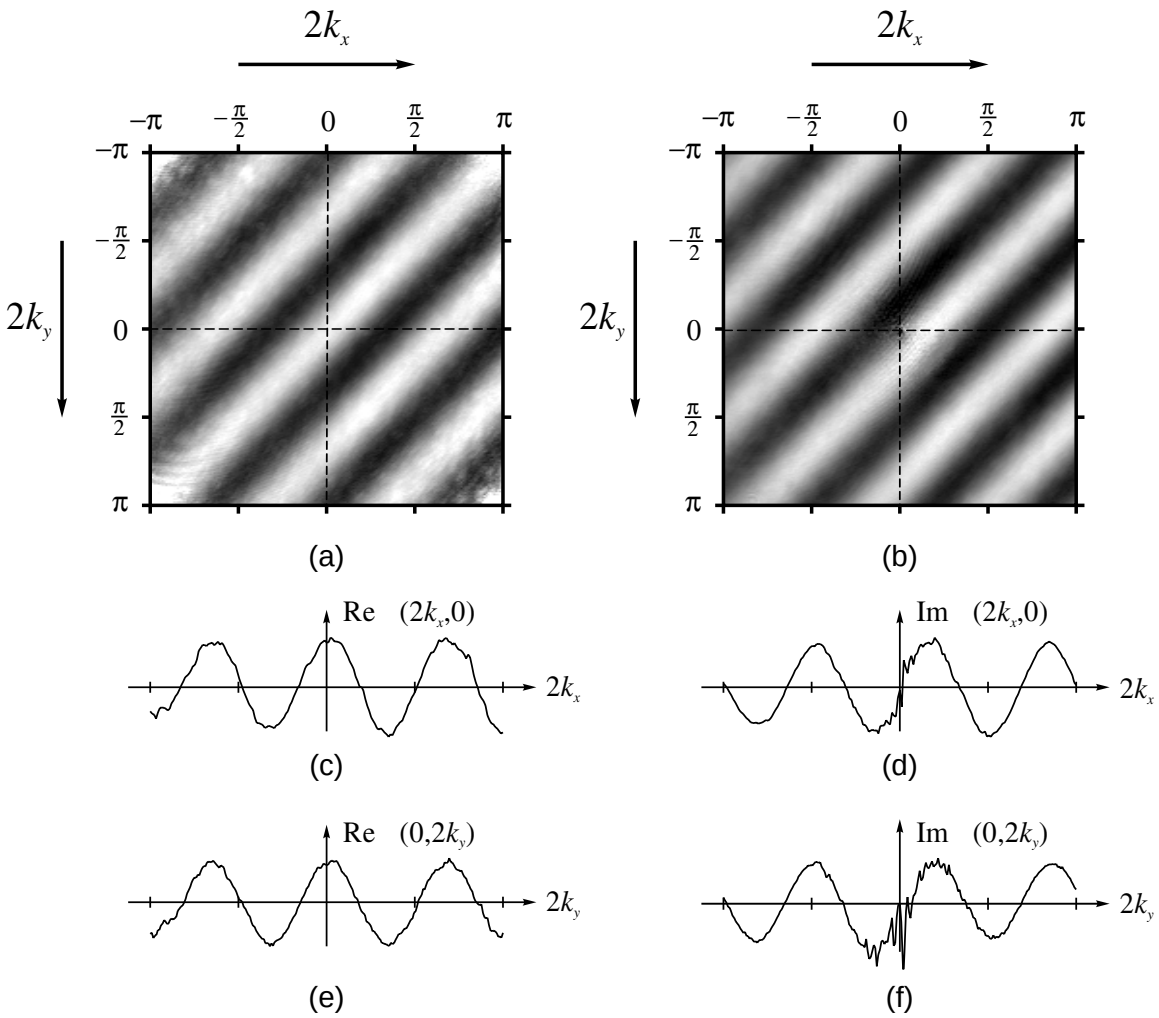



**Figure 5-6** Spectrum of single-pixel $(m,n)=(3,3)$ source image: (a) real part; (b) imaginary part; (c) real and (d) imaginary parts along $k_x$ axis; and (e) real and (f) imaginary parts along $k_y$ axis.

As before, the real and imaginary parts of the spectrum are shown in parts (a) and (b) of the figure, and the amplitude of the real and imaginary parts along the $k_x$ and $k_y$ axes are illustrated in parts (c) and (d), and parts (e) and (f). There are exactly ten sinusoidal cycles per $2\pi$ rad along each direction and the real and imaginary parts exhibit the required quadrature phase relationship. As before, the indeterminate points affect the accuracy of the spectra in the four corners of the real spectrum and at the origin of the imaginary spectrum.

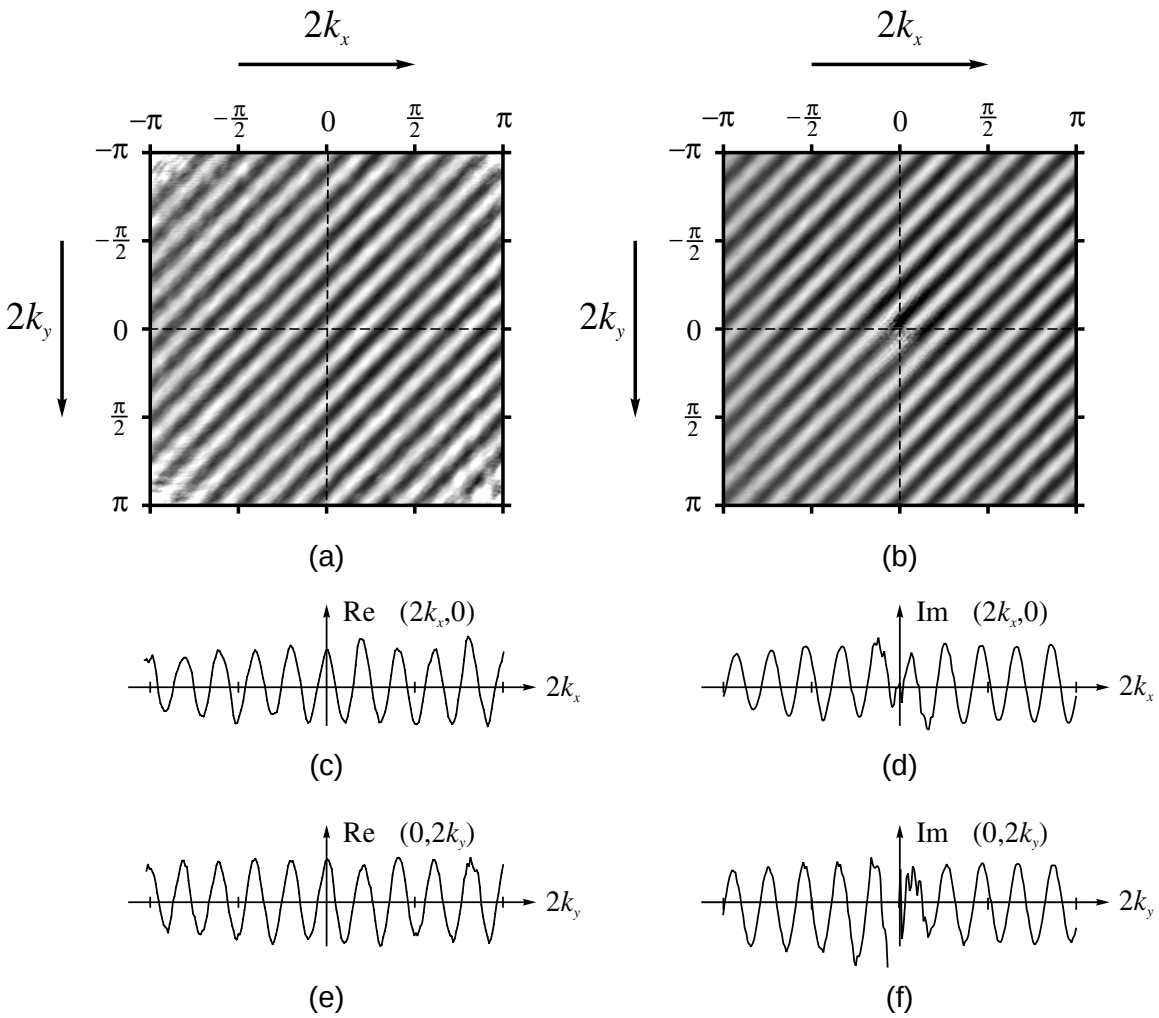



**Figure 5-7** Spectrum of single-pixel $(m,n)=(10,10)$ source image: (a) real part; (b) imaginary part; (c) real; and (d) imaginary parts along $k_x$ axis; and (e) real and (f) imaginary parts along $k_y$ axis.

## 5.4 TRAJECTORY ESTIMATION WITH A SINGLE-PIXEL REFERENCE

The mixed-domain method of multiple-object trajectory estimation was reviewed in Section 3.2. A hybrid optical-digital signal processing system for single-object trajectory estimation was described and its operation experimentally verified in Sections 3.3 through 3.5. The system employed an optical Fourier processor equipped with a point-diffraction interferometer to compute the speed and

direction of moving objects. The system successfully estimated the trajectory of several different point-objects, however, it was unable to determine the direction with which the objects moved along their trajectory. In addition, the system could not accommodate multiple objects, large objects, a stationary background, or random noise.

An optical Fourier processor and the JTI method of coherent detection are employed in this section to demonstrate an improved optical-digital trajectory estimation system. The new system overcomes the disadvantages associated with the previous architecture based on the point-diffraction method. The system described in the present section (Section 5.4) employs the JTI method and a single-pixel reference image to estimate the trajectory of single and multiple moving point-objects. The system in the following section (Section 5.5) employs the same JTI technique and a multipixel reference image to estimate the trajectory of multiple large (non-point) objects moving in the presence of background imagery and random noise.

### 5.4.1 Single Point-Object Trajectory Estimation

Consider first the trajectory estimation of a single point-object, with maximum contrast, moving on a black noiseless background. This trivial experiment was performed in Chapter 3 using the point-diffraction based Fourier processor and is duplicated here using the JTI technique.

A set of image sequences, which parallel those used in the point-diffraction experiment, were constructed from a series of $320 \times 240$ pixel images. Each sequence is $65$ frames in duration and describes the motion of a single-pixel object moving on a linear trajectory. Objects in the new sequence were assigned an ID letter which corresponded to its initial position in the original point-diffraction experiment (object A in the new experiment moves from initial position A to final position A', and was identified as trajectory A-A' in the previous experiment). A primed ID letter represents an object moving in the opposite

**Table 5-1** Actual Trajectories for Objects A, B, C, D, and I through P

| ID | Velocity, speed, and Direction | | | | Initial Position | | Final Position | |
|---|---|---|---|---|---|---|---|---|
| | $v_x$ [ppf] | $v_y$ [ppf] | $v$ [ppf] | $\phi$ [deg.] | $m_1$ | $n_1$ | $m_2$ | $n_2$ |
| A | 0.50 | -0.25 | 0.56 | -26.6 | -11 | 13 | 21 | -3 |
| B | 1.00 | -0.50 | 1.12 | -26.6 | -27 | 21 | 37 | -11 |
| C | 2.00 | -1.00 | 2.24 | -26.6 | -59 | 37 | 69 | -27 |
| D | 4.00 | -2.00 | 4.47 | -26.6 | -123 | 69 | 133 | -59 |
| I | 1.00 | 0.00 | 1.00 | 0.0 | -32 | 5 | 32 | 5 |
| J | 1.00 | 1.00 | 1.41 | 45.0 | -37 | -27 | 27 | 37 |
| K | 0.00 | 1.00 | 1.00 | 90.0 | -5 | -32 | -5 | 32 |
| L | -1.00 | 1.00 | 1.41 | 135.0 | 27 | -37 | -37 | 27 |
| M | -1.00 | 0.00 | 1.00 | 180.0 | 32 | -5 | -32 | -5 |
| N | -1.00 | -1.00 | 1.41 | -135.0 | 37 | 27 | -27 | -37 |
| O | 0.00 | -1.00 | 1.00 | -90.0 | 5 | 32 | 5 | -32 |
| P | 1.00 | -1.00 | 1.41 | -45.0 | -27 | 37 | 37 | -27 |

direction along the same path (the initial and final positions are swapped). The actual velocity, speed, direction, and initial and final positions, of the 16 objects employed in the experiment, are summarized in Table 5-1 (primed objects are not included).

*5.4.1.1 Experimental*

A block diagram of the optical setup was illustrated previously in Figure 5-4. The present experiment employs the same setup, however the focal length of lens L2 and the diameter of the pinhole have been changed to produce a larger beam ($f_1$=14.8 mm, $f_2$=125 mm, $\phi$=50 μm, and $2w_1$=4.34 mm). The diameter of the beam incident on the DMD was increased such that its power on a four-pixel cell is approximately equal to that which was incident upon a single-pixel during the point-diffraction experiment.

A sequence of images describing the linear motion of an object are written to the DMD. The complex-valued FT of each image frame is calculated using the JTI algorithms described by Equations 4-11 and 4-12. Modulated spectra (the

cosine-modulated real part and the sine-modulated imaginary part) are calculated with both the horizontal and vertical JTI algorithms. The two modulated spectra are added in a weighted sum, on a point-by-point basis, in order to avoid the indeterminate points associated with the zeros of the modulation functions on the coordinate axes (see Section 5.3.1), and at the same time, demodulate the spectrum. The complex FT of an image frame is therefore determined by,

$$S(2k_x, 2k_y) = w_1\alpha_{vert} + w_2\alpha_{horz} + i(w_3\beta_{vert} + w_4\beta_{horz}) \qquad (5\text{-}14)$$

where the modulated spectra calculated via both the vertical and horizontal JTI algorithms are given by,

$$\begin{aligned} \alpha_{vert} &= I_{\substack{s\,r\\0\,0}} + I_{\substack{r\,s\\0\,0}} - 2I_s - 2I_r, \qquad \beta_{vert} = I_{\substack{s\,r\\0\,0}} - I_{\substack{r\,s\\0\,0}}, \\ \alpha_{horz} &= I_{\substack{s\,0\\r\,0}} + I_{\substack{r\,0\\s\,0}} - 2I_s - 2I_r, \qquad \beta_{horz} = I_{\substack{s\,0\\r\,0}} - I_{\substack{r\,0\\s\,0}} \end{aligned} \qquad (5\text{-}15)$$

and the trigonometric weighting functions are,

$$\begin{aligned} w_1 &= \frac{\cos(k_x)}{4\left(\cos^2(k_x)+\cos^2(k_y)\right)}, \qquad w_2 = \frac{\cos(k_y)}{4\left(\cos^2(k_x)+\cos^2(k_y)\right)} \\ w_3 &= \frac{\sin(k_x)}{4\left(\sin^2(k_x)+\sin^2(k_y)\right)}, \qquad w_4 = \frac{\sin(k_y)}{4\left(\sin^2(k_x)+\sin^2(k_y)\right)} \end{aligned} \qquad (5\text{-}16)$$

Inspection of Equation 5-15 reveals that a total of six intensity spectra are required to calculate the full complex-valued FT of a single image (the real part of the FT in the four corners of the calculated spectrum and the imaginary part at the origin are indeterminate). The six intensity spectra were measured sequentially at 1/6 video frame rates during the experiments, however, a parallel implementation would be more attractive on an actual real-time system.

An image sequence describing the motion of a moving object was written to the DMD and the set of six intensity spectra were measured for each frame in the sequence. The images written to the DMD were joint (or composite) images

composed of one source image from the animated sequence and a reference image. The reference image consisted of a single pixel at the origin. The complex-valued FT of each image frame was calculated using Equation 5-14 at a set of $T = 16$ points $(k_x, k_y)$, uniformly spaced around a circle of radius $k_r^{circle}$ on the spatial frequency plane (the circle is referred to as a *frequency-domain observation circle*). The radii of the circles used in each trajectory experiment were chosen to match those used in the point-diffraction experiments. The sequence of FTs was post-processed off-line using MATLAB to estimate the trajectory of each object. A flow-chart of the JTI and trajectory estimation post-processing procedure is illustrated in Figure 5-8.

The trajectory estimation procedure illustrated in Figure 5-8 is identical to that of the point-diffraction experiment with one exception: the data vector employed in the new procedure is complex-valued. In the point-diffraction experiment, the modified forward-backward linear prediction (FBLP) algorithm was employed to estimate the temporal frequency of the sinusoid contained in a real-valued data-vector $\boldsymbol{u}_t$. One of two possible directions for the object was arbitrarily chosen and the signs of the temporal frequency estimates at each spatial frequency were picked such that the set of frequency-domain triples $(k_x, k_y, \omega)$ described a plane in 3-D frequency space. The data vector in the present experiment is complex-valued; the JTI technique provides the full complex-valued spectrum. The temporal frequencies estimated by the FBLP algorithm are complete with their correct sign and the direction ambiguity is therefore resolved. The overall complexity of the algorithm is also reduced because the sign of each frequency estimate does not have to be determined once the object's direction has been chosen (this is particularly important for multiple objects where each object has its own plane in 3-D frequency space).

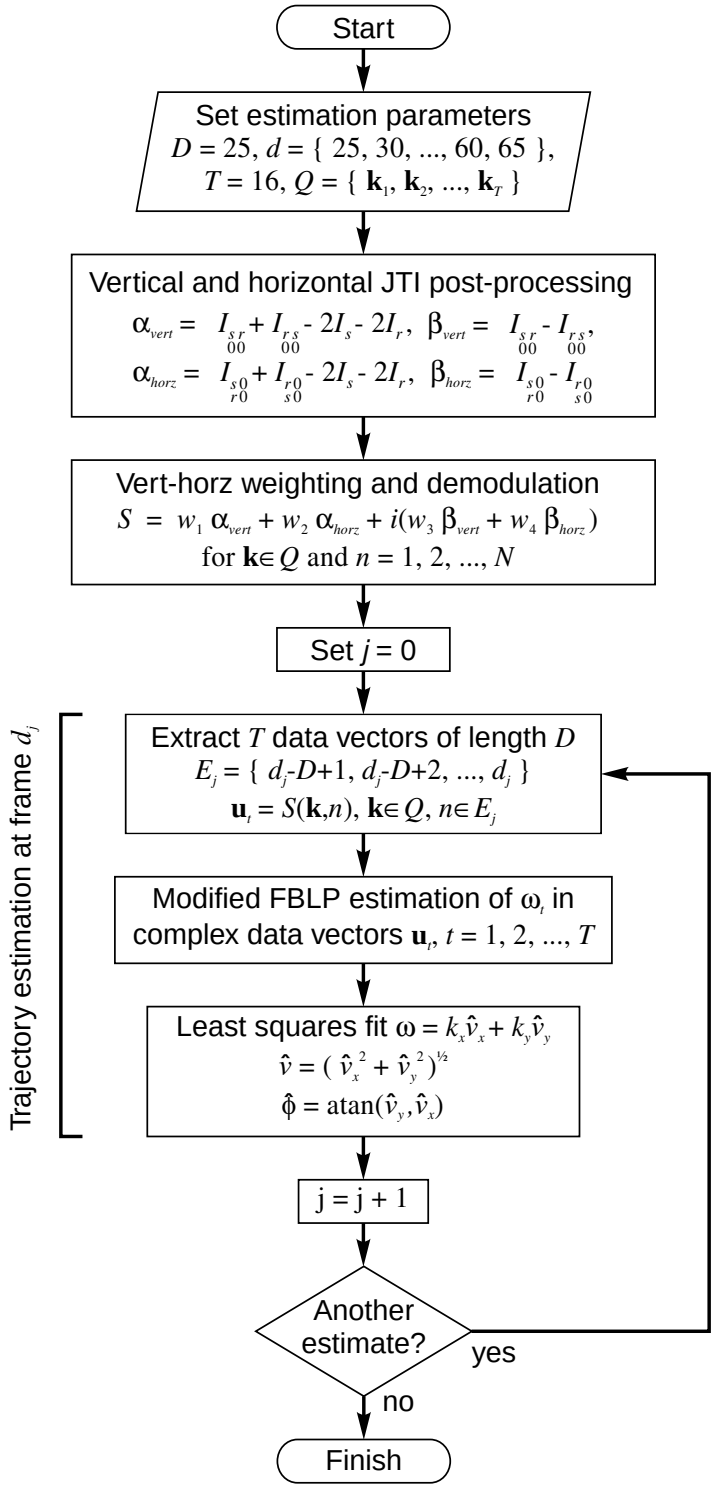



**Figure 5-8** JTI and trajectory estimation post-processing procedure.

*5.4.1.2 Results*

The velocity components of 16 point-object trajectories were estimated in two sets of experiments. During the first set of experiments, targets with different speeds were moved along a straight line at -27° to the $x$ axis. During the second set, targets with a velocity of either $1$ or $\sqrt{2}$ ppf (pixels per frame) were moved along paths at $0$, $\pm 45$, $\pm 90$, $\pm 135$, and $180°$ to the $x$ axis. The velocity components were estimated every $5^{\text{th}}$ frame starting with the $25^{\text{th}}$ frame and the observation interval was $25$ frames.

The trajectories of objects A, B, C, and D, and their primed counterparts (primed objects move along the same path but in the opposite direction), were estimated in the first set of experiments. The slowest objects (A and A') moved at $0.56$ ppf along the shortest trajectory ($36$ pixels) and the fastest objects (D and D') moved at $4.47$ ppf along the longest trajectory ($282$ pixels). Speed estimates at every $5^{\text{th}}$ frame, starting with the $25^{\text{th}}$ frame, are shown in Figure 5-9. The actual

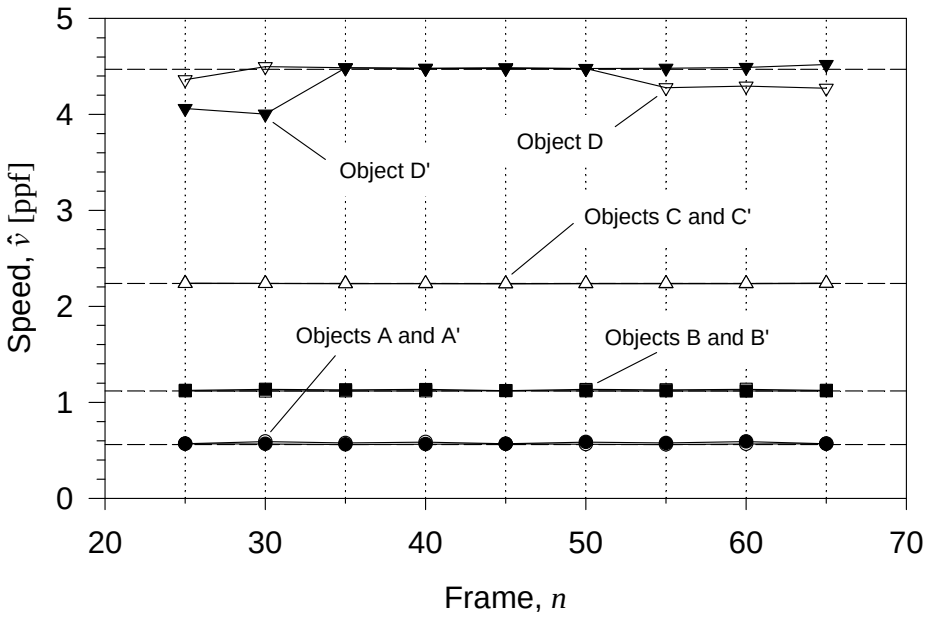



**Figure 5-9** Speed estimates for objects A, B, C, and D. Primed objects move along the same path in the opposite direction.

**Table 5-2** Actual and Estimated Trajectories for Objects A, B, C, and D

| ID | Radius | Speed Estimate | | | Direction Estimate | | |
|---|---|---|---|---|---|---|---|
| | $k_r^{circle}$ [rads] | $v$ [ppf] | $\hat{v}$ [ppf] | % Error | $\phi$ [deg] | $\hat{\phi}$ [deg] | Abs. Err. |
| A | $0.75\pi$ | 0.56 | $0.57 \pm 0.01$ | 2.28 | -26.6 | $-26.5 \pm 1.2$ | 0.0 |
| B | $0.57\pi$ | 1.12 | $1.13 \pm 0.01$ | 0.58 | | $-27.0 \pm 0.6$ | 0.5 |
| C | $0.28\pi$ | 2.24 | $2.24 \pm 0.00$ | 0.07 | | $-26.7 \pm 0.1$ | 0.1 |
| D | $0.19\pi$ | 4.47 | $4.40 \pm 0.10$ | 1.53 | | $-24.7 \pm 3.4$ | 1.8 |
| A’ | $0.75\pi$ | 0.56 | $0.57 \pm 0.01$ | 2.27 | 153.4 | $153.5 \pm 1.2$ | 0.0 |
| B’ | $0.57\pi$ | 1.12 | $1.13 \pm 0.01$ | 0.59 | | $153.0 \pm 0.6$ | 0.5 |
| C’ | $0.28\pi$ | 2.24 | $2.24 \pm 0.00$ | 0.05 | | $153.4 \pm 0.1$ | 0.1 |
| D’ | $0.19\pi$ | 4.47 | $4.39 \pm 0.20$ | 1.92 | | $151.6 \pm 3.4$ | 1.8 |

speed and direction, the average speed and direction estimated over the duration of motion (average of $9$ estimates), and the estimation errors, are summarized in Table 5-2 for each object. The frequency domain observation radius $k_r^{circle}$ employed in each experiment is also shown in the table.

The trajectories of objects I, J, K, L, M, N, O, and P, were estimated in the second set of experiments. Trajectories parallel to coordinate axes were $64$ pixels long and diagonal trajectories were $91$ pixels long. Direction estimates at every $5^{\text{th}}$ frame starting with the $25^{\text{th}}$ frame are shown in Figure 5-10. The actual speed and direction, the average speed and direction estimated over the duration of motion (average of $9$ estimates), and the estimation errors are summarized in Table 5-3 for each trajectory. The radius of the frequency domain observation circle was $k_r^{circle} = 0.57\pi$ rad and is equivalent to that used in the point-diffraction experiment.

*5.4.1.3 Discussion*

The average object speed was correctly estimated to within $\pm 3$ % of its actual value during the first set of experiments. The maximum estimation error is slightly lower than that observed during the same point-diffraction experiment where the

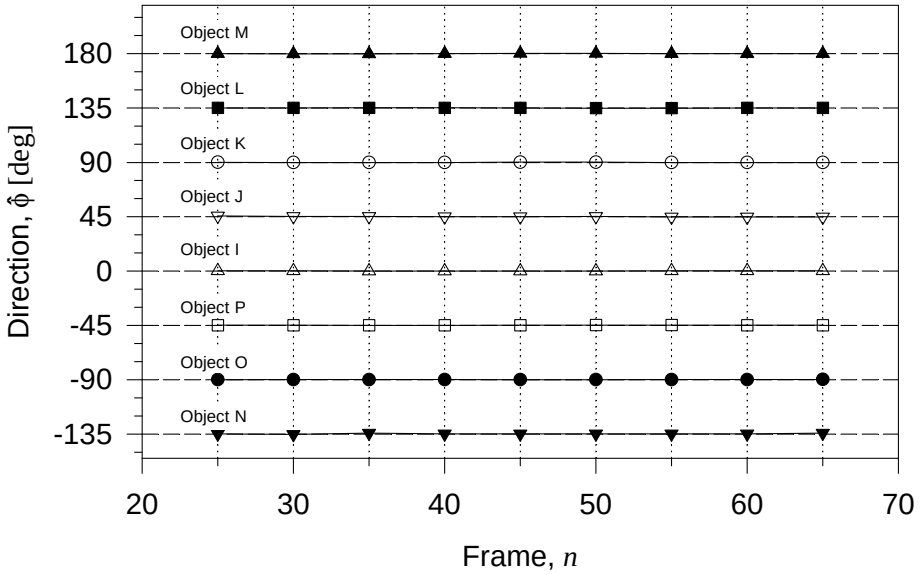



**Figure 5-10** Direction estimates for objects I through P.

average speed was estimated to within $\pm 4$ % of its actual value. The new system was able to track and consistently estimate the speed of the two fastest objects (objects D and D') whereas the previous system was unable to accurately estimate the speed of the same object. The speed estimates were consistently too low for object D at frames $55$, $60$, and $65$ (end of its trajectory), and for object D' at frames $25$ and $30$ (beginning its trajectory). This effect is most likely caused by insufficient illumination in the region around the end of trajectory D (beginning of trajectory D') . The new system was also able to estimate the direction to within $\pm 2°$ in which the objects moved along the line (the previous system could not detect the direction with which the objects moved along the line).

The average object direction was correctly estimated to within $\pm 0.2°$ of its actual value during the second set of experiments. The maximum estimation error is lower than that observed during the same point-diffraction experiment where the average direction was estimated to within $\pm 2.2°$ of its actual value. Most importantly, the new system is able to estimate direction on $360°$ because the full complex-valued FT is computed. The previous system could only

**Table 5-3** Actual and Estimated Trajectories for Objects I through P

| ID | Speed Estimate | | | Direction Estimate | | |
|---|---|---|---|---|---|---|
| | $v$ [ppf] | $\hat{v}$ [ppf] | % Error | $\phi$ [deg] | $\hat{\phi}$ [deg] | Abs. Err. |
| I | 1.000 | $1.001 \pm 0.001$ | 0.07 | 0.0 | $0.0 \pm 0.1$ | 0.0 |
| J | 1.414 | $1.415 \pm 0.000$ | 0.05 | 45.0 | $45.1 \pm 0.3$ | 0.1 |
| K | 1.000 | $1.000 \pm 0.000$ | 0.00 | 90.0 | $90.0 \pm 0.1$ | 0.0 |
| L | 1.414 | $1.415 \pm 0.001$ | 0.03 | 135.0 | $135.1 \pm 0.1$ | 0.1 |
| M | 1.000 | $1.001 \pm 0.001$ | 0.12 | 180.0 | $180.0 \pm 0.1$ | 0.0 |
| N | 1.414 | $1.415 \pm 0.000$ | 0.05 | -135.0 | $-134.8 \pm 0.3$ | 0.2 |
| O | 1.000 | $1.000 \pm 0.000$ | 0.05 | -90.0 | $-90.0 \pm 0.1$ | 0.0 |
| P | 1.414 | $1.415 \pm 0.001$ | 0.06 | 45.0 | $-44.9 \pm 0.1$ | 0.1 |

estimate direction up to a reversal of $180°$ because only the real part of the FT was detected.

### 5.4.2 Multiple Point-Object Trajectory Estimation

The JTI-based trajectory estimation system is extended to multiple objects in the this section. The experimental setup and the method of coherent detection (weighted horizontal-vertical JTI with a single-pixel reference image) are identical to those used in the previous section, however, the trajectory estimation algorithm is extended to accommodate multiple moving objects.

#### *5.4.2.1 Multiple-Object Trajectory Estimation*

An ideal object moving in space-time describes a plane in 3-D frequency space with slope (along the spatial frequency axes) equal to the object's velocity. Two ideal objects describe two planes; the slope of each plane equals the velocity of its respective object (Figure 3-2 shows the 3-D frequency representation of two objects moving at the same speed in opposite directions). In general, each ideal moving object in space-time will describe its own plane in frequency space. The frequency-domain representation of a moving object is a complex-valued function

of three variables; the square of its absolute value, at some point $(k_x, k_y, \omega)$, describes the energy associated with that particular frequency triple. A plane is said to exist in frequency space if the energy distribution is such that points on the plane have high energy (infinite energy as described by δ-function for an ideal moving object) and all other points have low energy (zero energy in the ideal case).

The mixed-domain trajectory estimation algorithm determines the trajectory of an object by estimating the temporal frequency of a mixed-domain signal $S(k_x, k_y, n)$ using the modified FBLP algorithm. The signal at any point $(k_x, k_y)$ describes a complex-valued sinusoid with temporal frequency proportional to the object's velocity. A single component sinusoid is detected for a single moving object and multiple sinusoids for multiple objects. In general, the number of component sinusoids detected in $S(k_x, k_y, n)$ is equal to the number of moving objects. Consider the case of $P$ moving objects. The modified FBLP algorithm is employed to estimate the temporal frequencies of $P$ component sinusoids at each of the $T$ observation points on the spatial frequency plane. This provides a set of $T \cdot P$ frequency-domain triples $(k_x, k_y, \omega)$, each of which lies on one of $P$ possible planes through the origin. The set of triples is searched and divided into $P$ subsets such that each triple in the subset belongs to the same plane. The subset of triples belonging to each plane is used to characterize its slope and yield the velocities of the $P$ objects.

The trajectory estimation algorithm proposed by Knudsen employed a systematic brute-force method to separate the set of $T \cdot P$ triples into subsets of 3-tuples belonging to different planes [Knu92a]. The algorithm presented here employs the Hough transform [Hou62, Bal82] to identify planes in 3-D frequency space. The Hough transformation maps triples $(k_x, k_y, \omega)$ from 3-D frequency

space into lines in $v_x$-$v_y$ parameter space. An ideal moving object is described in 3-D frequency space by a plane through the origin,

$$v_x k_x + v_y k_y - \omega = 0 \tag{5-17}$$

and is parameterized by the object's velocity components. Each frequency-domain triple on the plane describes a line,

$$v_y = -\frac{k_x}{k_y} v_x + \frac{\omega}{k_y} \tag{5-18}$$

in Hough parameter space. The slope of the line depends on the position of the observation point on the spatial frequency plane and the $v_y$ intercept depends on (and is proportional to) the temporal frequency estimated at that point. Each triple in 3-D frequency space describes a different line in 2-D parameter space. Triples that lie on a plane parameterized by an object with velocity $\mathbf{v}_1 = \left[v_{x,1}\ v_{y,1}\right]^{\mathsf{T}}$ map to a family of lines which intersect at the point $\left(v_{x,1}, v_{y,1}\right)$ in parameter space. The set of triples corresponding to each frequency-space plane maps to its own family of intersecting lines in parameter space. An object's velocity components are determined by identifying the point of intersection in $v_x$-$v_y$ parameter space corresponding to its family of lines.

One can not tell with absolute certainty if a particular line in parameter space belongs to a given set of intersecting lines. Under ideal noiseless conditions, all lines which belong to a particular object will intersect perfectly at the same point on the $v_x$-$v_y$ plane and lines that do not include this point must therefore belong to a different object (a line which belongs to one object, however, can include the point of intersection of a second object, and vice versa). This parameter-space identification problem is analogous to determining whether a particular frequency triple belongs to a given plane in 3-D frequency space. In the presence of noise, the $v_x$-$v_y$ plane is searched for the most probably points of intersection.

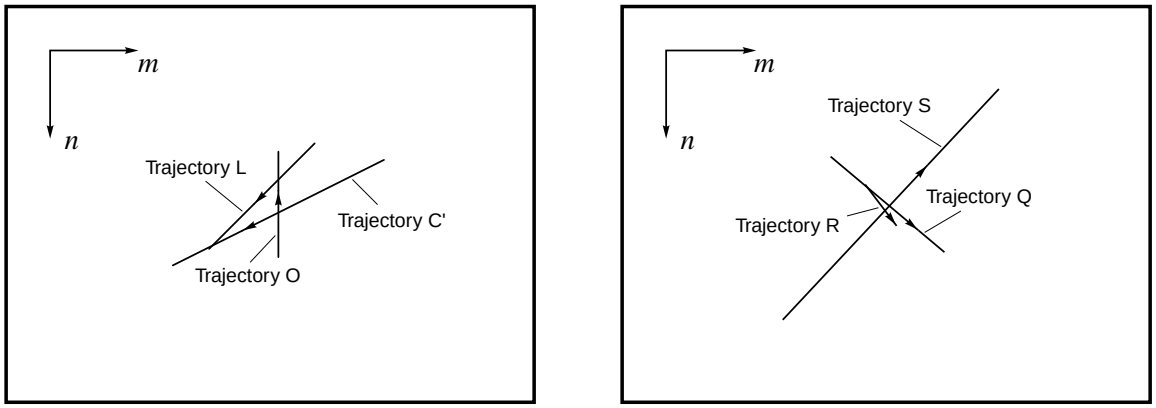
The Hough transform algorithm is well documented [Bal82]. A line given by Equation 5-18 is *drawn* in a discrete parameter-space voting-array for each frequency-domain triple. The elements in the array at which all the lines belonging to a particular plane intersect, will receive the most votes. The array is searched for the elements with the most votes and the indices of these elements correspond to the velocity components of the different objects. The Hough transform is a more robust estimator of the velocity parameters than the least-squares fit employed in the single-object algorithm because the outliers are ignored.

*5.4.2.2 Experimental*

The trajectories of three point-objects moving simultaneously over a zero intensity background were estimated in two sets of experiments. Two different $65$ frame animated sequences, referred to as image sequences 1 and 2, and each describing the linear motion of three single-pixel objects, were constructed from a sequence of $320\times240$ pixel images. The velocity, speed and direction, and initial and final positions, of the three objects moving in the two different image sequences, are shown in Table 5-4. The object trajectories are also illustrated in

**Table 5-4** Actual Trajectories for Image Sequences 1 and 2

| ID | Velocity, speed, and Direction | | | | Initial Position | | Final Position | |
|---|---|---|---|---|---|---|---|---|
| | $v_x$ [ppf] | $v_y$ [ppf] | $v$ [ppf] | $\phi$ [deg.] | $m_1$ | $n_1$ | $m_2$ | $n_2$ |
| Image Sequence 1 | | | | | | | | |
| O | 0.00 | -1.00 | 1.00 | -90.0 | 5 | 32 | 5 | -32 |
| L | -1.00 | 1.00 | 1.41 | 135.0 | 27 | -37 | -37 | 27 |
| C’ | -2.00 | 1.00 | 2.24 | 153.4 | 69 | -27 | -59 | 37 |
| Image Sequence 2 | | | | | | | | |
| R | 0.30 | 0.39 | 0.49 | 52.8 | -10 | -12 | 9 | 13 |
| Q | 1.08 | 0.91 | 1.41 | 40.0 | -31 | -29 | 38 | 29 |
| S | 2.05 | -2.19 | 3.00 | -46.9 | -60 | 70 | 71 | -70 |




(a) (b)

**Figure 5-11** Actual trajectories for: (a) image sequence 1 (objects O, L and C'); and (b) image sequence 2 (objects R, Q, and S).

Figure 5-11. Image sequence 1 contained objects O, L, and C', which move on simple trajectories with unit pixel displacement per frame and well separated directions. Image sequence 2 contained objects R, Q, and S, which moved on slightly more complicated trajectories (objects R and Q move in almost the same direction).

The post-processing trajectory estimation algorithm is extended to accommodate three moving objects. The extension is illustrated in Figure 5-12 and replaces the segment of the original algorithm labeled "trajectory estimation at frame $d_j$", shown in Figure 5-8. The post-processing algorithm assumes that there are exactly three moving objects in the image sequence and the modified FBLP algorithm is instructed to estimate the temporal frequency of three sinusoids in each data vector. The absolute value of temporal frequencies, estimated at diametrically opposed spatial frequency points, were averaged in order to: 1) reduce the number of frequency-domain triples; and 2) increase the accuracy of the temporal frequency estimates. At some point on the frequency-domain observation circle, the temporal frequency has a sign which is opposite to that of the temporal frequency at a diametrically opposed point,

$$\omega(k_x, k_y) \;=\; k_x v_x + k_x v_y \;=\; -\omega(-k_x, -k_y) \tag{5-19}$$

because temporal frequency is an odd function of spatial frequency. The temporal frequencies around the first half of the observation circle were combined with those around the second half in a simple average,

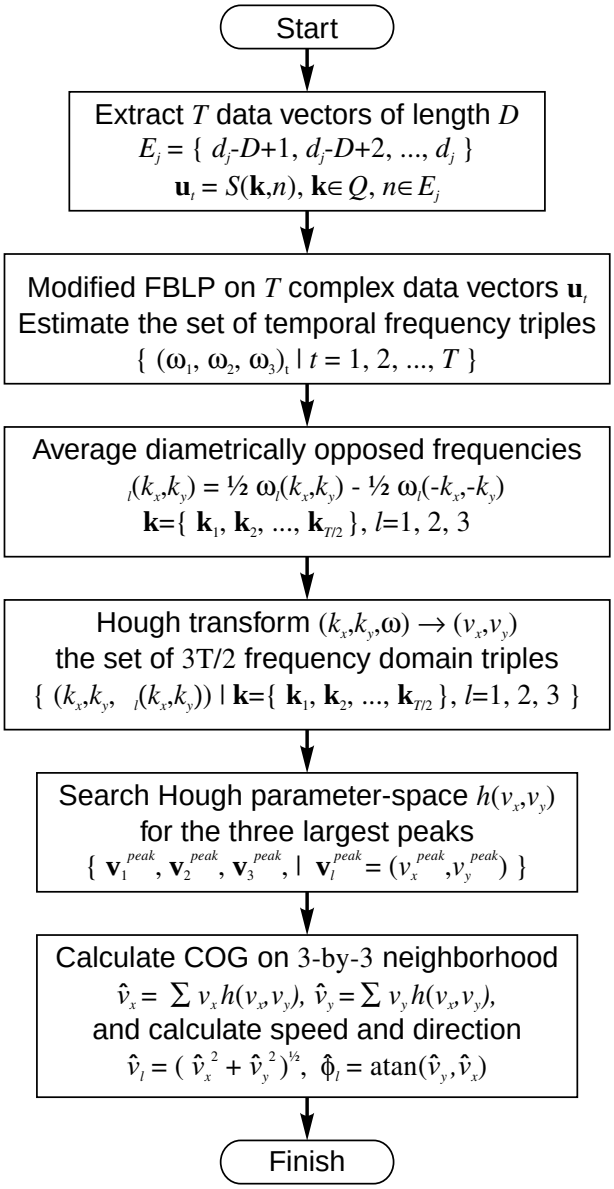



**Figure 5-12** Modified trajectory estimation algorithm extended to accommodate three moving objects.

$$\overline{\omega}(k_x,k_y) \;=\; \tfrac{1}{2}\omega(k_x,k_y) \;-\; \tfrac{1}{2}\omega(-k_x,-k_y) \qquad (5\text{-}20)$$

as indicated by the algorithm of Figure 5-12. The Hough-transform parameter-space array had dimensions $65\times65$ and its voting elements were linearly distributed between $v_x^{min}=v_y^{min}=-2.5\,\text{ppf}$ and $v_x^{max}=v_y^{max}=2.5\,\text{ppf}$, resulting in a velocity quantization width of $\Delta v=\Delta v_x=\Delta v_y=0.077\,\text{ppf}$. The three objects were identified in parameter space as the three maximum-valued elements each belonging to a different peak. The velocity components of each object were then calculated as the center-of-gravity (COG) on a $3\times3$ neighborhood centered on each maximum.

*5.4.2.3 Results*

The velocity components of the $3$ objects in each animated sequence were estimated at every $5^{\text{th}}$ frame starting with the $25^{\text{th}}$ frame (the duration of the observation interval was $D=25$ frames). The radius of the frequency-domain

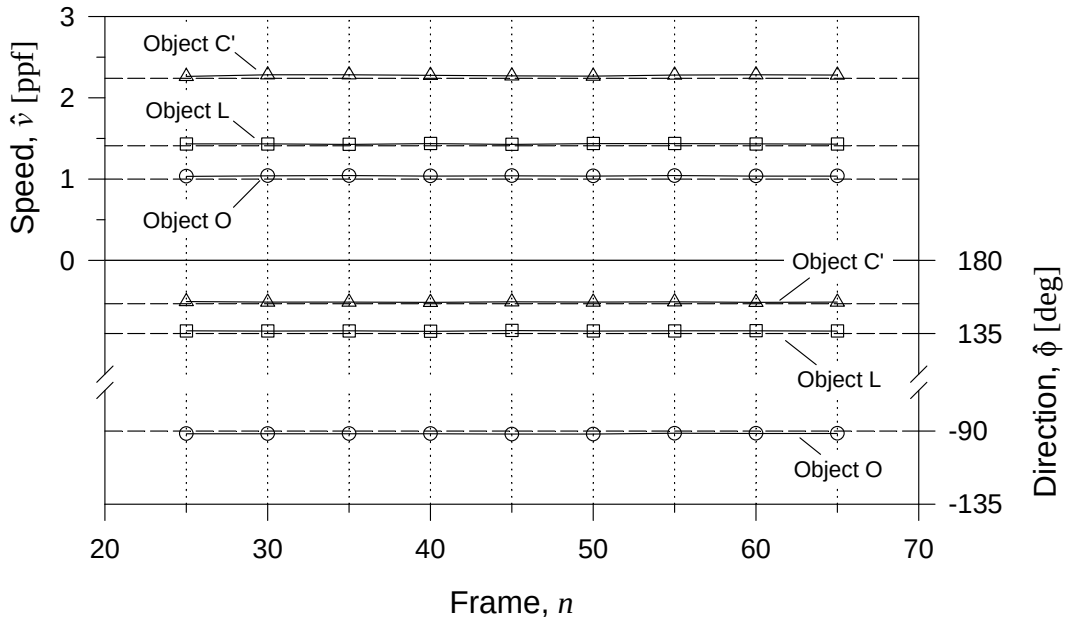



**Figure 5-13** Speed and direction estimates for the three objects O, L, and C' of image sequence 1.

observation circle was set to $k_r^{circle} = \frac{\pi}{3}\,\text{rad}$ (the *a priori* maximum speed was assumed to be $v^{max} = 3\,\text{ppf}$). The speed and direction estimates for the first and second sequences are shown respectively in Figures 5-13 and 5-14. The actual speed and direction, the average speed and direction estimated over the duration of motion (average of $9$ estimates), and the estimation errors, for the three objects in each sequence, are summarized in Table 5-5.

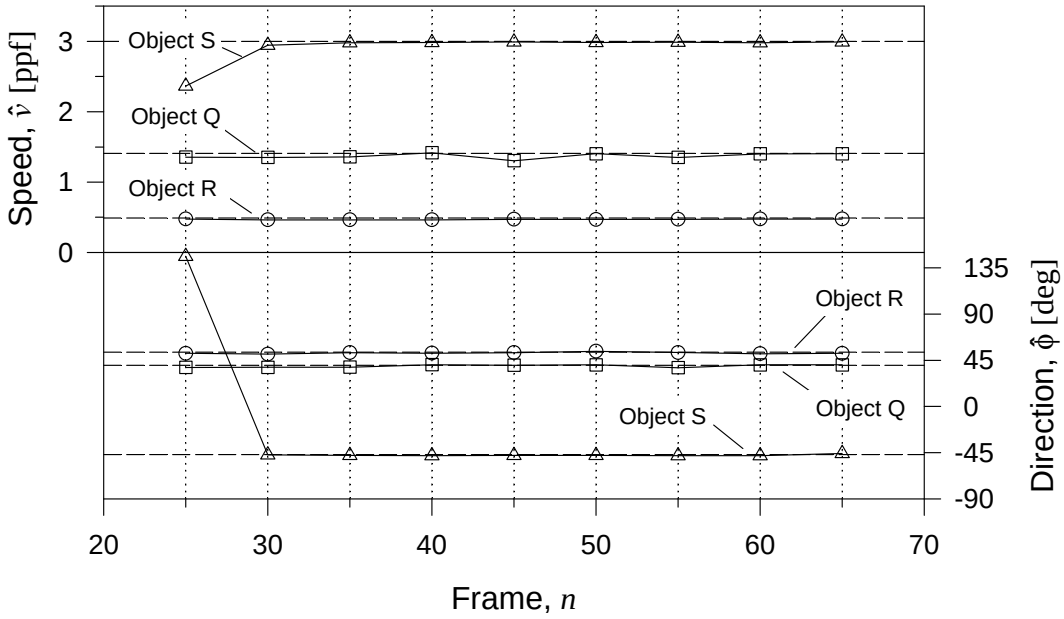



**Figure 5-14** Speed and direction estimates for the three objects R, Q, and S of image sequence 2.

*5.4.2.4 Discussion*

Objects O, L, and C' in the first sequence 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 $\pm 2$ % of its actual value (expressed as a percentage of $v^{max}$) and the average direction of each object was correctly estimated to within $\pm 2°$. Objects R, Q, and S in the second sequence were successfully tracked and their trajectories were correctly

estimated at each estimation frame except for object S frame 25. The average speed and direction for each object were correctly estimated to within $\pm 2$ % and $\pm 2°$ of their actual values (speed and direction for object S at frame 25 were not included in the average).

**Table 5-5** Estimated Trajectories for Image Sequences 1 and 2

| ID | Speed Estimate | | | Direction Estimate | | |
|---|---|---|---|---|---|---|
| | $v$ [ppf] | $\hat{v}$ [ppf] | % Error | $\phi$ [deg] | $\hat{\phi}$ [deg] | Abs. Err. |
| Image Sequence 1 | | | | | | |
| O | 1.00 | $1.039 \pm 0.003$ | 1.3 | -90.0 | $-91.7 \pm 0.2$ | 1.7 |
| L | 1.41 | $1.433 \pm 0.003$ | 0.6 | 135.0 | $136.5 \pm 0.1$ | 1.5 |
| C' | 2.24 | $2.277 \pm 0.007$ | 1.4 | 153.4 | $154.3 \pm 0.1$ | 0.9 |
| Image Sequence 2 | | | | | | |
| R | 0.49 | $0.468 \pm 0.005$ | 0.8 | 52.8 | $52.0 \pm 0.1$ | 0.7 |
| Q | 1.41 | $1.371 \pm 0.018$ | 1.2 | 40.0 | $39.4 \pm 1.9$ | 0.7 |
| S | 3.00 | $2.981 \pm 0.004^{\dagger}$ | 0.5 | -46.9 | $-47.5 \pm 0.2^{\dagger}$ | 0.5 |

$^{\dagger}$ Estimate at frame 25 removed from average and standard deviation calculation

## 5.5 TRAJECTORY ESTIMATION WITH A MULTIPIXEL REFERENCE

The trajectory estimation system presented in the previous section employed the JTI method of coherent detection and a single-pixel reference image. The source images were limited to either one or three pixels in order to maximize the fringe visibility of the joint transform interferograms. The system presented in this section employs a multipixel reference to estimate the trajectory of multiple large (non-point) objects moving in the presence of background imagery and random noise.

The multipixel system described here is identical to that presented in Section 5.4.2 except: 1) a multipixel rather than a single-pixel reference is employed; and 2) the beam incident on the DMD is enlarged. The optical setup was illustrated previously in Figure 5-4. The beam expander in the new experimental setup

consists of objective lens L1 ($10\text{X}$ objective, $f_1$=8.3 mm), collimating lens L2 (corrected triplet, $f_2$=125 mm), and pinhole PH ($\phi$=25 μm), and provides an incident beam with diameter $2w_1$=7.74 mm. The multipixel system employs the weighted horizontal-vertical JTI post-processing algorithm (Figure 5-8) and the multiple-object trajectory estimation algorithm (Figure 5-12). The *a priori* maximum speed was set to $v^{max} = 3\,\text{ppf}$. This choice also determines the radius of the frequency-domain observation circle, $k_r^{circle} = \frac{\pi}{3}\,\text{rad}$. Trajectories were estimated at every $5^{\text{th}}$ frame starting with the $25^{\text{th}}$ frame and the duration of the observation interval was $D = 25$ frames.

The trajectory estimation system described was employed to determine the trajectory of multiple moving objects in four sets of experiments: 1) $3\times3$ pixel objects on a zero intensity background, 2) $3\times3$ objects on a simple artificial background; 3) $3\times3$ objects in additive noise; and 4) single-pixel objects in additive noise. Each experiment employed an image sequence which described the simultaneous motion of three identical objects moving on their own trajectories. The multipixel reference images are binarized (or if you prefer thresholded) Bessel functions, constructed to provide uniform amplitude and phase in the frequency domain, around an annular region of interest, which coincides with the observation points.

### 5.5.1 Bessel Function Reference Image

The intensity of the source and the reference spectra must be matched in order to produce an interferogram with good visibility and provide complex-valued source spectra with good SNR (see Section 4.7). An ideal frequency-domain reference has: 1) a constant intensity distribution equal to the maximum intensity of the source spectrum; and 2) a uniform phase distribution. In theory, a δ-function reference image provides an 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 trajectory estimation algorithm does not require the FT of each image frame at every point on the spatial frequency plane – the algorithm requires the FT at $T$ uniformly spaced points on an observation circle of radius $k_r^{circle}$. An annular region of interest is defined which includes the observation circle. A reference image is constructed such that the frequency-domain reference has uniform phase and intensity matched to that of the source in the region of interest. The frequency-domain reference (and its corresponding spatial-domain reference image) should have circular symmetry. The zero-order Hankel transform (a 1-D Fourier transform with a zero-order Bessel-function kernel) of a zero-order Bessel function is a δ-function ring of radius $k_r^{circle}$. The Hankel transform-pair is given by,

$$k_r^{circle}\ J_0\left(k_r^{circle} r\right) \quad \Leftrightarrow \quad 2\pi\ \delta\left(k_r - k_r^{circle}\right) \tag{5-21}$$

where $r = \sqrt{m^2 + n^2}$ and $k_r = \sqrt{k_x^2 + k_y^2}$ are radial coordinates in the spatial and frequency domains. 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 constructing the appropriate spatial-domain Bessel image. The finite amount of optical power reflected (or transmitted) by the pixels of the reference image is mapped into the δ-function ring on the spatial frequency plane. The intensity of the ring can be matched to that of the source spectra by modifying the amplitude of the Bessel image on a gray-scale SLM or by changing the threshold employed to binarize the reference on a binary SLM such as the DMD.

The algorithm used to generate the binarized Bessel images is illustrated in Figure 5-15. All reference images were constructed with a design frequency equal to that of the frequency-domain observation circle. The Bessel images were binarized using a threshold value such that the resulting image was composed of $L$ on-pixels. The FT of a binarized and discretely-sampled reference image is not an exact δ-function ring, however, the energy of the reference is still concentrated in a region about $k_r = k_r^{circle}$ on the spatial frequency plane. The intensity of the δ-function ring can be modified by choosing a reference image with an appropriate number $L$ of on-pixels. A series of Bessel function reference-images, constructed using the algorithm with $L = 5, 17, 41, 117, 229,$ and $449$ pixels, and $k_r^{circle} = \frac{\pi}{3}\,\text{rad}$, are illustrated in Figure 5-16. The reference images shown in the figure are $51 \times 51$ pixels in size.

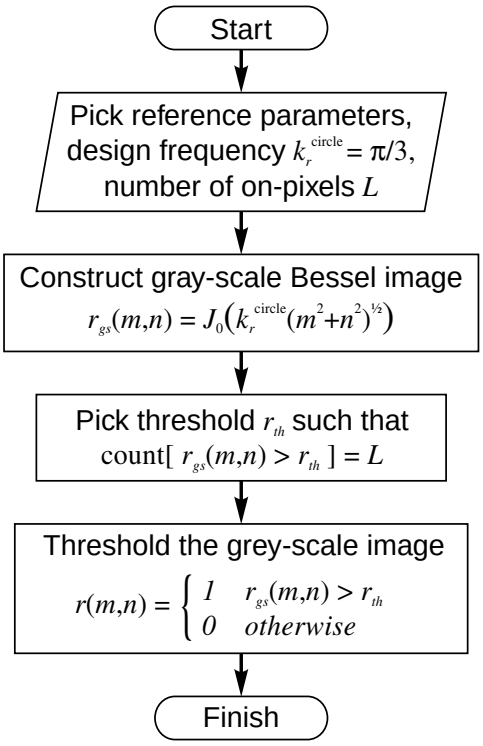



**Figure 5-15** Algorithm to generate a binarized Bessel-function reference image.

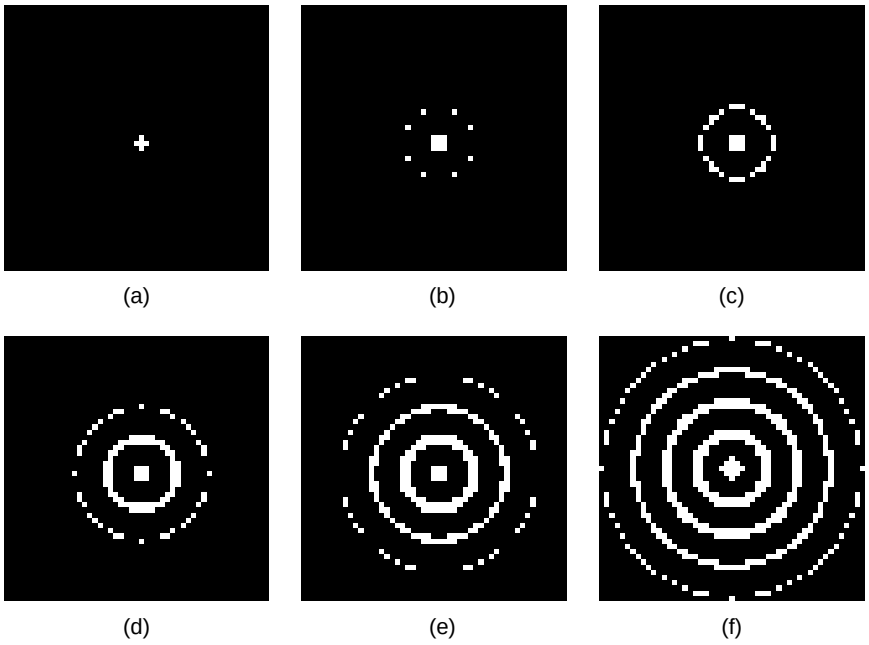



**Figure 5-16** Bessel reference images: (a) $L$=5 pixels; (b) $L$=17 pixels; (c) $L$=41 pixels; (d) $L$=117 pixels; (e) $L$=229 pixels; and (f) $L$=449 pixels.

### 5.5.2 Multiple $3\times3$ Pixel Objects

#### *5.5.2.1 Experimental*

The trajectories of three $3\times3$ pixel objects were estimated in the first of four multipixel reference experiments. Two different $65$ frame animated sequences, referred to as image sequences 3 and 4, and each describing the linear motion of three $3\times3$ pixel objects, were constructed from a sequence of $128\times128$ pixel images. Image sequence 3 contained objects M, J, and T, which moved on simple trajectories with unit pixel displacement per frame and well separated directions. Image sequence 4 contained objects U, V, and M, which moved on slightly more complicated trajectories. The velocity, speed and direction, and initial and final positions, of the three objects in each sequence, are shown in Table 5-6. The trajectory experiment was run several times employing different Bessel-function reference images. A total of 14 different reference images were

**Table 5-6** Actual Trajectories for Image Sequences 3 and 4

| ID | Velocity, speed, and Direction | | | | Initial Position | | Final Position | |
|---|---|---|---|---|---|---|---|---|
| | $v_x$ [ppf] | $v_y$ [ppf] | $v$ [ppf] | $\phi$ [deg.] | $m_1$ | $n_1$ | $m_2$ | $n_2$ |
| Image Sequence 3 | | | | | | | | |
| 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 |
| Image Sequence 4 | | | | | | | | |
| U | 0.30 | 0.39 | 0.49 | 52.8 | -6 | -8 | 13 | 17 |
| V | 1.08 | 0.91 | 1.41 | 40.0 | -29 | -29 | 40 | 29 |
| W | 1.86 | -1.98 | 2.72 | -46.9 | -59 | 63 | 60 | -64 |

tested and the number of on-pixels in each image was logarithmically distributed between $L = 5$ pixels and $L = 449$ pixels.

*5.5.2.2 Results and Discussion*

The average fringe visibility (depth of modulation) was determined for each Bessel-function reference image tested with image sequence 3. The average

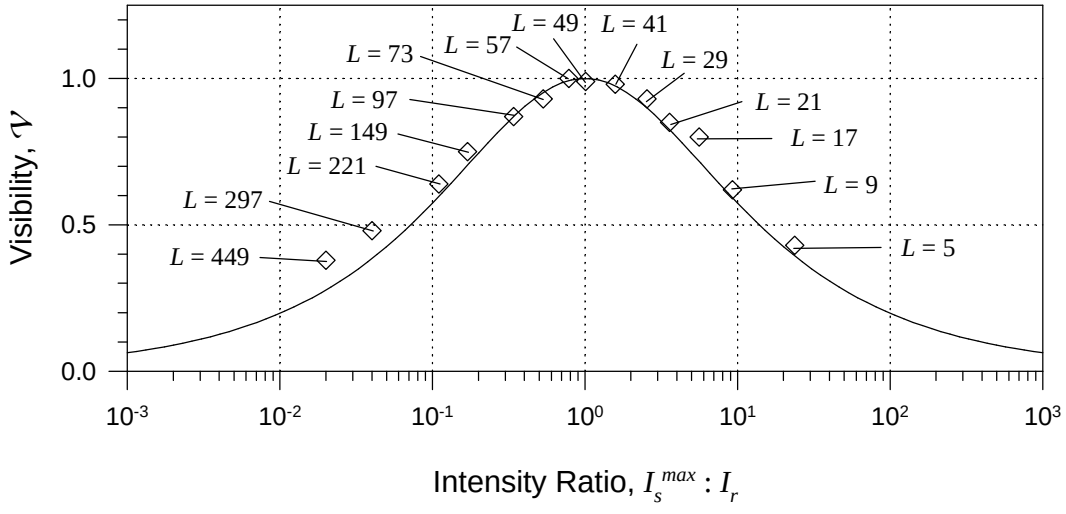



**Figure 5-17** Average fringe visibility versus the ratio of maximum source intensity to reference intensity for various Bessel reference images.

fringe visibility was determined by calculating the fringe visibility (Equation 4-33) at each point on the frequency-domain observation circle, for each frame in the image sequence, and then computing their average. The average fringe visibility versus the ratio of maximum source intensity $I_s^{max}$ to reference intensity $I_r$ is plotted in Figure 5-17. The theoretical visibility curve described by Equation 4-33 is also shown. The reference images with $L$=41, 49, and 57 on-pixels, with respective average fringe visibilities $\mathcal{V}$=0.98, 0.99, and 1.00, best match the source image sequence. Reference images with twice (and half) as many on-pixels provide a fringe visibility close to unity ($\mathcal{V}$=0.87 for $L$=97 and $\mathcal{V}$=0.85 for $L$=21).

The trajectory estimation algorithm successfully tracked and correctly estimated the velocity components of all three objects, in both image sequences, to within $\pm 2\Delta v = \pm 0.154$ ppf (where $\Delta v$ is the Hough transform parameter-space quantization width), for all reference images, except $L$=5 and $L$=449 (the velocity components of object T were correctly estimated at only 8 of the 9 frames in the two exceptional cases).

**Table 5-7** Estimated Trajectories for Image Sequences 3 and 4

| ID | Speed Estimate | | | Direction Estimate | | |
|---|---|---|---|---|---|---|
| | $v$ [ppf] | $\hat{v}$ [ppf] | % Error | $\phi$ [deg] | $\hat{\phi}$ [deg] | Abs. Err. |
| Image Sequence 3 | | | | | | |
| M | 1.00 | $1.034 \pm 0.005$ | 1.1 | 180.0 | $-178.7 \pm 0.3$ | 1.3 |
| J | 1.41 | $1.404 \pm 0.004$ | 0.3 | 45.0 | $45.2 \pm 0.2$ | 0.2 |
| T | 2.24 | $2.186 \pm 0.009$ | 1.7 | -26.6 | $-28.3 \pm 0.5$ | 1.7 |
| Image Sequence 4 | | | | | | |
| U | 0.49 | $0.468 \pm 0.004$ | 0.8 | 52.8 | $52.0 \pm 0.6$ | 0.8 |
| V | 1.41 | $1.390 \pm 0.037$ | 0.6 | 40.0 | $40.2 \pm 0.8$ | 0.2 |
| W | 2.72 | $2.697 \pm 0.027$ | 0.7 | -46.9 | $-47.5 \pm 1.2$ | 0.6 |

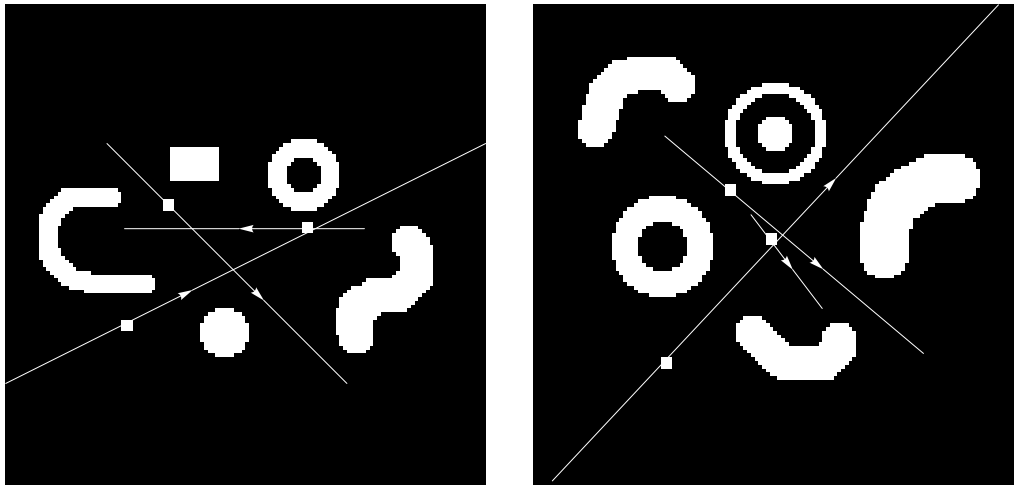
The average speed and direction estimated using the $L$=57 Bessel image are shown in Table 5-7. The table summarizes the actual speed and direction, the average speed and direction estimated over the duration of motion (average of 9 estimates), and the estimation errors, for each object. Objects M, J, and T in image sequence 3, and objects U, V, and W in image sequence 4, 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 $\pm 2$ % of its actual value and the average direction of each object was correctly estimated to within $\pm 2°$.

Knudsen and Bruton presented similar experimental results for three objects moving simultaneously on linear trajectories [Knu92a]. In their experiment, the three moving objects were square $3\times 3$ pixel targets of intensity 255 moving over a background of zero intensity (objects of maximum contrast in a noiseless environment). The three objects moved on trajectories similar to those employed in this work (object 1: $v_1 = \sqrt{2}$ ppf , $\phi_1 = 40°$; object 2: $v_2 = \frac{1}{2}$ ppf , $\phi_2 = 53°$, and object 3: $v_3 = 3$ ppf , $\phi_3 = -47°$). The trajectories were estimated every 5$^{th}$ frame starting with the 25$^{th}$ frame. In the Knudsen and Bruton experiment, the maximum absolute velocity error was 0.06 ppf (2% of the maximum velocity) and the maximum absolute direction error was 1.8°. These results compare favorably with those produced by the hybrid optical-digital system. However, in the Knudsen and Bruton experiment, the third object left the image sequence after the 61$^{st}$ frame and their system was able to detect the loss of the third object and correctly estimated the trajectory of the remaining two objects. Three moving objects were known *a priori* in the hybrid optical-digital trajectory estimation system presented in this dissertation.

### 5.5.3 Multiple $3\times3$ Pixel Objects on a Stationary Background

#### *5.5.3.1 Experimental*

A synthetic background was added to animated sequences 3 and 4 and the trajectories of the $3\times3$ pixel objects were estimated using the trajectory estimation system. The two background images, which remained the same for each frame in the sequence, are illustrated in Figure 5-18. The figure shows frame $n=17$ for the two animated sequences. The objects can be identified in each frame as the $3\times3$ pixel squares; the trajectories are illustrated by the narrow lines indicating the object's direction. The background of sequence 3 is composed of five simple objects adding a total of $1212$ on-pixels to each frame of the sequence; the background of sequence 4 is composed of six larger objects adding a total of $2008$ on-pixels to each frame.



(a) (b)

**Figure 5-18** Synthetic backgrounds and object trajectories: (a) image sequence 3 (objects M, J, and T) with background; and (b) image sequence 4 (objects U, V, and M) with background. The $3\times3$ pixel objects are shown at frame $n$=17.

*5.5.3.2 Results and Discussion*

The $L$=257 and $L$=205 reference images provided the best results for animated sequences 3 and 4, respectively. The actual speed and direction, the average speed and direction estimated over the duration of motion, and the estimation errors, for the three objects in each sequence, are summarized in Table 5-8. 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 $\pm 2$ % of its actual value and the average direction of each object was correctly estimated to within $\pm 2°$. The addition of a stationary background to the animated sequences did not appear to make any difference to the accuracy of the trajectory estimation system. It is curious that the trajectories of the sequence with the most background pixels (image sequence 4), were more accurately estimated using the reference image with the least pixels. The reason for this is that the objects in image sequence 4 were larger than those in image sequence 3, and consequently, image sequence 4 had more power near the origin of frequency space, and less power on the frequency-domain

**Table 5-8** Estimated Trajectories for Image Sequences 3 and 4 with a Background

| ID | Speed Estimate | | | Direction Estimate | | |
|---|---|---|---|---|---|---|
| | $v$ [ppf] | $\hat{v}$ [ppf] | % Error | $\phi$ [deg] | $\hat{\phi}$ [deg] | Abs. Err. |
| Image Sequence 3 | | | | | | |
| M | 1.00 | $1.030 \pm 0.004$ | 1.0 | 180.0 | $-178.2 \pm 0.4$ | 1.8 |
| J | 1.41 | $1.401 \pm 0.003$ | 0.5 | 45.0 | $45.2 \pm 0.1$ | 0.2 |
| T | 2.24 | $2.199 \pm 0.016$ | 1.2 | -26.6 | $-27.4 \pm 1.3$ | 0.9 |
| Image Sequence 4 | | | | | | |
| U | 0.49 | $0.467 \pm 0.004$ | 0.8 | 52.8 | $52.0 \pm 0.4$ | 0.7 |
| V | 1.41 | $1.394 \pm 0.023$ | 0.5 | 40.0 | $40.1 \pm 0.9$ | 0.0 |
| W | 2.72 | $2.692 \pm 0.028$ | 0.9 | -46.9 | $-47.1 \pm 1.0$ | 0.2 |

observation circle, than image sequence 3. Sequence 4 therefore required a Bessel image with fewer pixels than sequence 3 in order to be intensity matched on the frequency-domain observation circle.

### 5.5.4 Multiple $3\times3$ Pixel Objects with Additive Noise

#### *5.5.4.1 Experimental*

A synthetic white-noise image was added to each frame of image sequence numbers 3 and 4. The trajectories of the three $3\times3$-pixel objects were estimated using the hybrid optical-digital trajectory estimation system and then using 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 using a 2-D FFT algorithm with 64 bits of floating point accuracy. Each $128\times128$-point FFT required approximately 1 million floating point operations using the 2-D FFT algorithm in MATLAB version 5.2 and consumed approximately $70\ \text{msec}$ of CPU time on an Intel $166\ \text{MHz}$ Pentium processor.

The white-noise sequence was constructed by thresholding $65$ different realizations of a $128\times128$ pixel i.i.d. (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 target signal-to-noise ratio is defined as the target signal power divided by the noise power,

$$\text{SNR} = \frac{\sum_{\text{targets}} s(m,n)^2}{\sum_{\text{noise}} s(m,n)^2} = \frac{\text{number of target pixels}}{\text{number of noise pixels}} \tag{5-22}$$

and is equal to the ratio of the number of target pixels to the number of noise pixels in each frame because both target and noise pixels have the same intensity (the DMD is used as a binary SLM). Figure 5-19 shows frame $n=17$ of image sequence 3 with four levels of additive white noise. The frames illustrated

in parts (a) through (d) have SNR = 11.3, 0.0, -6.0, and –10.5 dB, which correspond to 2, 27, 107, and 303 noise pixels, respectively (there are $3 \cdot 3^2 = 27$ target pixels per frame). The objects can be identified in each image frame as the $3 \times 3$ pixel squares; the trajectories are illustrated by the narrow lines indicating the object's direction.

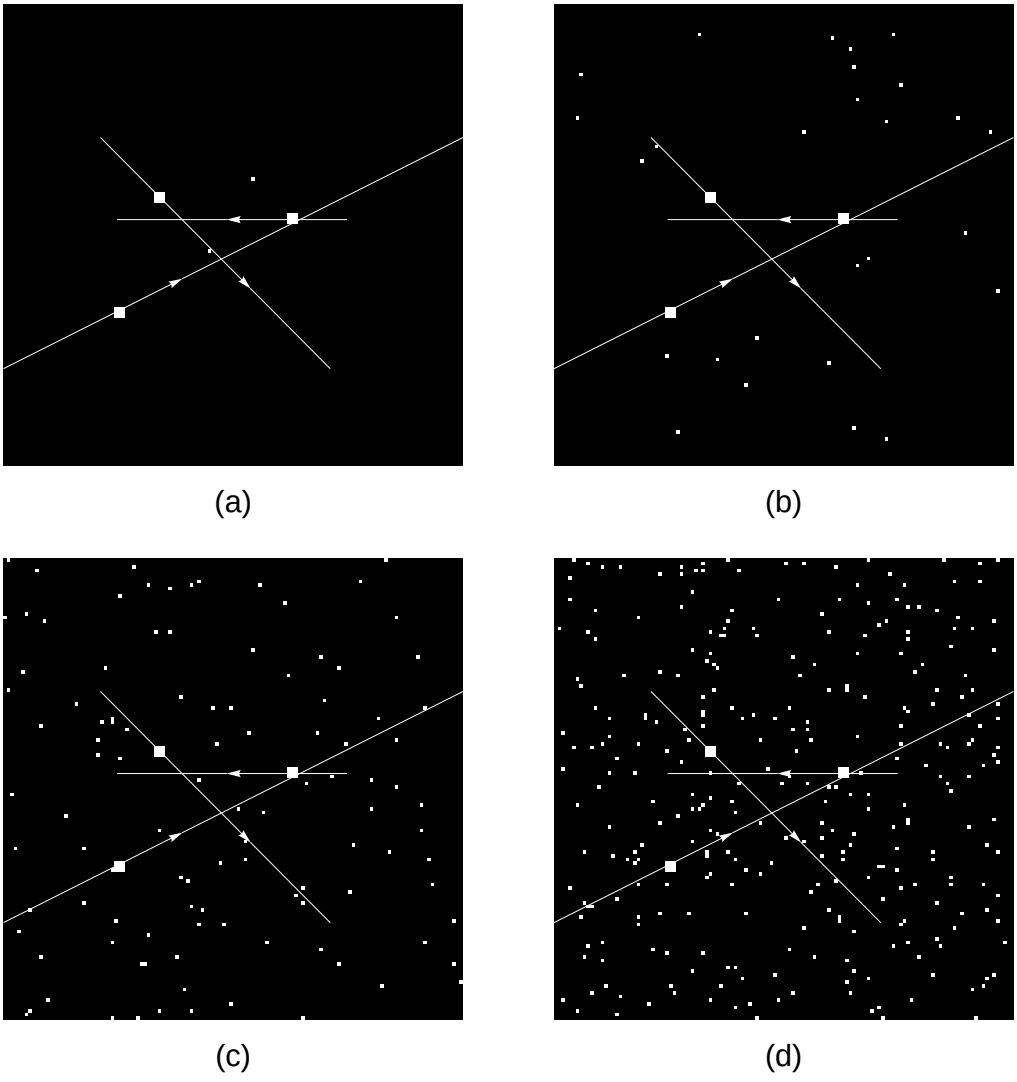



**Figure 5-19** Image sequence 3 with four levels of noise: (a) SNR = 11.3 dB (2 noise pixels); (b) SNR=0.0 dB (27 noise pixels); (c) SNR=-6.0 dB (107 noise pixels); and (d) SNR=–10.5 dB (303 noise pixels). The objects are shown at frame *n*=17.

*5.5.4.2 Results and Discussion*

The trajectory of each object in image sequences 3 and 4 was estimated at 9 different frames over the duration of motion (every 5$^{th}$ frame starting with the 25$^{th}$ frame) at four different noise levels. Tables 5-9 and 5-10 summarize the number of frames out of 9, for image sequences 3 and 4, where the velocity components of the three objects were correctly estimated. An object was correctly estimated if both its velocity components were estimated to within two times the quantization error, $\pm 2\Delta v = \pm 0.154$ ppf, introduced by the Hough parameter-space voting matrix.

At a noise level corresponding to a target SNR of $11.3$ dB, the three objects in both image sequences were correctly estimated by the hybrid optical-digital trajectory estimation system at all 9 estimation frames (total of 27 objects correctly estimated in each sequence). The number of correctly estimated objects decreased monotonically with SNR. Further investigation at other noise levels indicated that the trajectories were generally estimated correctly for an SNR greater than $-3$ to $0$ dB (depending on the trajectory) and that the number of correctly estimated objects decreased linearly with SNR below this threshold. When the noise power was 10 times the target power (SNR=-10.5 dB), 1 out of 27 (object M at frame $55$) and 0 out of 27 possible objects were correctly identified, during image sequences 3 and 4, respectively. The performance for the two

**Table 5-9** Number of Frames Correctly Estimated in Image Sequence 3 for $3\times3$ pixel objects (Total of 9 Estimation Frames)

| SNR [dB] | # of noise pixels | Optical-Digital System | | | | All-Digital System | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | M | J | T | Tot. | M | J | T | Tot. |
| 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 |

**Table 5-10** Number of Frames Correctly Estimated in Image Sequence 4 for $3\times3$ pixel objects (Total of 9 Estimation Frames)

| SNR [dB] | # of noise pixels | Optical-Digital System | | | | All-Digital System | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | U | V | W | Tot. | U | V | W | Tot. |
| 11.3 | 2 | 9 | 9 | 9 | 27 | 9 | 9 | 9 | 27 |
| 0.0 | 27 | 9 | 8 | 7 | 24 | 9 | 9 | 9 | 27 |
| -6.0 | 107 | 5 | 2 | 3 | 10 | 4 | 3 | 3 | 10 |
| -10.5 | 303 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 2 |

noisiest image sequences (-6 dB and -10dB) could be improved substantially by preprocessing the sequences using a simple median filter to improve their SNR.

The performance of the all-digital trajectory estimation system was generally better than that of the hybrid optical-digital system. Both systems performed equally poorly at the lowest SNRs. The pixel aperture of the CCD camera was simulated in the all-digital processor by calculating 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.

### 5.5.5 Multiple Single-Pixel Objects with Additive Noise

#### *5.5.5.1 Experimental*

The experiment described in Section 5.5.4 was repeated with single-pixel objects instead of $3\times3$ pixel objects. The trajectories of sequence 3 objects M, J, and T, and sequence 4 objects U, V, and W, were unchanged, however, the 3-by-3 objects in each frame were replaced by single pixels. This experiment is a more realistic test of the mixed-domain trajectory estimation algorithm which was originally proposed for the detection and trajectory estimation of small, barely discernible, multiple moving objects of unknown position and velocity. Because the target and noise pixels in this experiment have the same shape and intensity, it is impossible to distinguish between the two in a single frame without knowing the trajectories *a priori*. The image sequence *must* be integrated, or rather observed, over several frames in order to estimate the trajectories.

#### *5.5.5.2 Results and Discussion*

The object trajectories in image sequences 3 and 4 were again estimated at 9 different frames over the duration of motion (every 5$^{th}$ frame starting with the 25$^{th}$ frame) at four different noise levels. Tables 5-11 and 5-12 summarize the number of frames out of 9, for image sequences 3 and 4, where the velocity components of the three objects were correctly estimated.

**Table 5-11** Number of Frames Correctly Estimated in Image Sequence 3 for single-pixel objects (Total of 9 Estimation Frames)

| SNR [dB] | # of noise pixels | Optical-Digital System | | | | All-Digital System | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | M | J | T | Tot. | M | J | T | Tot. |
| ∞ | 0 | 9 | 9 | 9 | 27 | 9 | 9 | 9 | 27 |
| 0.0 | 3 | 6 | 2 | 3 | 11 | 7 | 4 | 4 | 15 |
| -3.0 | 6 | 3 | 2 | 1 | 6 | 2 | 0 | 0 | 2 |
| -6.0 | 12 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |

When no noise was added (infinite SNR), the three objects in image sequence 3 were correctly estimated by the hybrid optical-digital trajectory estimation system at all 9 estimation frames (total of 27 objects correctly estimated). Image sequence 4 objects U and V were correctly estimated by the hybrid optical-digital system 9 times out of 9, however, object W was correctly estimated at only 7 of the 9 estimation frames (total of 25 objects correctly estimated). The number of correctly estimated objects decreased monotonically with SNR. Further investigation at other noise levels indicated that the trajectories were generally estimated correctly for an SNR greater than 2 to 4 dB (depending on the trajectory) and that the number of correctly estimated objects decreased linearly with SNR below this threshold. When there were twice as many noise pixels as target pixels (SNR=-3.0 dB), 6 and 3 out of the 27 possible objects were correctly identified, during image sequences 3 and 4, respectively. No objects were correctly identified in either sequence when the noise level was 4 times the target signal level (SNR=-6.0 dB).

Once again, the performance of the all-digital trajectory estimation system was generally better than that of the hybrid optical-digital system. Both systems performed equally poorly at the lowest SNRs. As before, the FT at a particular spatial frequency was calculated as an average of four spectral values centered on the desired point. When the FT at each spatial frequency was computed as a

**Table 5-12** Number of Frames Correctly Estimated in Image Sequence 4 for single-pixel objects (Total of 9 Estimation Frames)

| SNR [dB] | # of noise pixels | Optical-Digital System | | | | All-Digital System | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | U | V | W | Tot. | U | V | W | Tot. |
| ∞ | 0 | 9 | 9 | 7 | 25 | 9 | 9 | 9 | 27 |
| 0.0 | 3 | 7 | 5 | 2 | 14 | 4 | 5 | 6 | 15 |
| -3.0 | 6 | 1 | 2 | 0 | 3 | 4 | 0 | 0 | 4 |
| -6.0 | 12 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |

single sampled point rather than an average of four points, the hybrid optical-digital system usually outperformed the all-digital system. It is speculated that if the FT at a particular frequency was calculated as the average of a densely sampled region, the all-digital system would always perform at least as well as the hybrid optical-digital system.

Both trajectory estimation systems were able to consistently track and correctly estimate the trajectories of single-pixel objects when the SNR was greater than about $3$ dB. The hybrid optical-digital system required the measurement of 6 interferograms and consumed $16 \cdot 13 = 208$ floating-point operations to determine the FT at the $T = 16$ desired spatial frequencies of each image frame. The all-digital system consumed approximately 1 million floating point operations to calculate the FT of each image (additional computational savings may be gained by using a pruned FFT algorithm). The FFT algorithm consumes almost 4 orders of magnitudes more floating-point operations for a $128 \times 128$ pixel image than the optical Fourier processor. The computational advantage of the optical Fourier processor is even more significant for larger images. Time domain trajectory estimation systems [Blo91] have demonstrated tracking and trajectory estimation of 10 moving point-objects when the SNR is less than -$20$ dB. Time-domain trajectory estimation techniques appear to be more robust to noise than the mixed-domain methods investigated in this dissertation.

## 5.6 CHAPTER SUMMARY

The JTI method of coherent detection was experimentally verified. The spectra of two single-pixel source images were determined using the weighted three-component JTI algorithm and a single-pixel reference image. The real and imaginary parts of the source spectra were correctly calculated as 2-D sinusoidal functions with period determined by the source pixel's displacement from the origin. The real and imaginary parts exhibited the required quadrature phase relationship. The source spectra were calculated as a weighed sum of positively and negatively modulated spectra. The weighted sum avoided the indeterminate points associated with the modulation function zeros (except at the four corners of the real spectrum and the origin of the imaginary spectrum), and demodulated the JTI spectra, at the same time.

The *on-state* tilt of the DMD mirrors introduced a blazing effect which produced a linear phase shift across the plane of the device. The phase shift caused a mutual displacement between the (coherent) optical transfer function (OTF) and the set of replicated spectra (interference maxima). This effect, combined with the source-spectrum replication introduced by the JTI zero-interleaving, resulted in a spectrum, uncontaminated by contrast-ratio noise, between the $6^{th}$ and $7^{th}$ interference maxima. The zero spatial frequency of this spectrum coincided almost exactly with the maximum absolute value of the OTF.

The full trajectory of single and multipixel objects moving through a synthetic stationary background and in the presence of white noise were successfully estimated using the improved JTI based optical Fourier processor. The point-diffraction experiment presented in Chapter 3 was repeated using the improved system. The object trajectories were estimated with greater accuracy and more importantly the direction ambiguity was successfully resolved. A Hough transform technique was employed to characterize the parameters of the object's velocity planes in 3-D frequency space. The trajectories of three single-pixel objects were

correctly estimated using this technique and a single-pixel JTI reference. Multi-pixel JTI reference images were constructed by thresholding a 2-D Bessel function. The binarized Bessel images were constructed to produce a $\delta$-function ring which coincided with the frequency-domain observation circle and an intensity on the circle equal to the maximum source spectra intensity. The trajectories of three $3\times3$ pixel objects moving on two different synthetic backgrounds were successfully estimated using various Bessel reference images. The trajectories of three $3\times3$ pixel and three single-pixel objects were correctly identified and successfully estimated in several experiments where the target SNR ratio was decreased by adding white Gaussian noise to the image sequences. The system was consistently able to track and correctly estimate the trajectories of $3\times3$ and single-pixel objects when the SNRs were greater than $-3$ dB and $3$ dB, respectively.

CHAPTER

# 6

# Conclusions and Suggestions for Further Work

Two interferometric optical Fourier transform (FT) processors were constructed and experimentally verified. One processor used a conventional point-diffraction interferometer (PDI) to detect the real part of the complex-valued optical distribution in the back focal-plane of a FT lens. The second processor employed a novel phase-shifting method of coherent detection, called joint transform interference (JTI), to measure both the real and imaginary parts of the optical FT. The new JTI technique is insensitive to vibration and temperature fluctuations due to its simple common-path design. The JTI source and reference images are easily matched in order to guarantee good fringe visibility.

Both optical processors were demonstrated in a trajectory estimation system designed to estimate the speed and direction of objects moving in an image sequence. The PDI based system successfully estimated the speed and direction (up to a reversal of 180°) of single point-objects of maximum contrast moving over a black background. The JTI based system successfully estimated the trajectory of multiple point-objects and large $3\times3$ pixel objects moving over a synthetic background and in the presence of Gaussian white noise.

The objectives, achievements, and contributions made in this dissertation are summarized in Section 6.1. Suggestions for further work are discussed in Section 6.2. The suggestions for further work are divided into those which apply to the JTI method of coherent detection and those which apply to its application to moving-object trajectory estimation. Other applications for interferometric optical Fourier processors are discussed in Section 6.3.

## 6.1 OBJECTIVES, ACHIEVEMENTS, AND CONTRIBUTIONS

There were two primary objectives in this dissertation. The first was to develop an indirect interferometric means of detecting and measuring the complex-valued optical field mapped into the back focal-plane of a Fourier transform lens. The second was to construct an optical Fourier transformer, based on the method of coherent detection, and employ it to demonstrate a moving-object trajectory estimation system using the Knudsen and Bruton mixed-domain algorithm. The summary of contributions presented in the following paragraphs demonstrate how these objectives were achieved.

A hybrid optical-digital trajectory estimation system based on the mixed-domain method of trajectory estimation was constructed and demonstrated using a point-diffraction based optical Fourier processor. The mixed-domain technique was developed and demonstrated in software by Knudsen and Bruton [Knu92a, Knu92b]. A hardware implementation was proposed in this dissertation and demonstrated using an optical Fourier processor and point-diffraction interferometer (PDI). The optical processing was performed at 1/3 video frame rates and the balance of the trajectory estimation algorithm was performed off-line using MATLAB. The system was limited to the trajectory estimation of a single-pixel object of maximum contrast moving over a zero intensity background. The system correctly estimated the path and speed of the moving object, however, the direction with which the object moved along the path was indeterminate. The addition of a single target, background, or noise pixel to the

image sequence reduced the visibility of the fringe patterns to a point where the trajectory of the moving object could not be accurately estimated.

A novel interferometric technique called *joint-transform interference (JTI)* was developed to recover the complex-amplitude spectrum in an optical Fourier processor. The FT of a source image is interfered with the FT of a reference image in a common-path interferometer. Unlike the PDI, both the real and imaginary parts of the complex-valued spectrum are determined, and in addition, the source and reference images are easily matched in order 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 to 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.

A spectrum which is unaffected by contrast-ratio noise was identified in the far-field diffraction pattern of the Texas Instruments DMD (digital micromirror device) and used to demonstrate the JTI-based optical Fourier processor. The spectrum was observed when an image was displayed on the DMD using only every second pixel and is a direct result of the tilted mirrors. The spectrum of an image, which is expanded by inserting zeros between valid pixels and then displayed on the DMD, is compressed by a factor of two in the horizontal and vertical directions. One of the replicated and compressed spectra is located in a region of the diffraction pattern unaffected by contrast-ratio noise and centered on the DMD Nyquist frequency. The linear phase distribution introduced by the titled mirrors at each DMD pixel shifts the position of the diffraction envelope (region of maximum intensity) to the Nyquist frequency. The spectrum centered at the DMD Nyquist frequency has a high diffraction efficiency and is much less noisy than the spectrum at DC.

The novel JTI technique was experimentally verified and applied to the problem of moving-object trajectory estimation. A hybrid optical-digital trajectory estimation system, based on the mixed-domain method of trajectory estimation, was constructed and demonstrated using 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 using MATLAB. Unlike the PDI-based optical processor described earlier, the JTI-based processor was able to estimate the trajectory (speed and direction over 360°) of multiple objects moving over both stationary and white noise backgrounds. The system was able to estimate the trajectories of $3\times3$ and single-pixel objects when the target-to-noise ratios were greater than $-3$ dB and $3$ dB, respectively. The trajectories estimated using the optical Fourier processor compared favorably with those calculated using a digital FFT. 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 via the FFT required more that three orders of magnitude more floating-point operations. It was known *a priori* that exactly three moving objects were present in each image sequence. The current implementation is limited because it does not detect the number of moving objects.

The trajectory estimation algorithm proposed by Knudsen and Bruton was modified by substituting the systematic search for 3-tuples with a Hough transform technique. The Hough transform search technique described here is mathematically tractable and easy to implement, however, the computational advantage (or disadvantage) compared to that of the original Knudsen and Bruton algorithm was not assessed because it was outside the scope of the dissertation. An ideal moving object is described in 3-D frequency space by a plane through the origin. Its trajectory is estimated by finding a set of frequency-domain triples which lie on the plane. The original algorithm searched a large set

of frequency-domain triples for 3-tuples (sets of three frequency triples) which described the same object. In the approach described in this dissertation, the large set of frequency-domain triples is Hough transformed into a velocity parameter space, in order to reduce the size of the search space.

## 6.2 SUGGESTIONS FOR FURTHER WORK

The suggestions for further work are divided into those which apply to the JTI method of coherent detection and those which apply to its application to moving-object trajectory estimation.

### 6.2.1 Joint Transform Interference

#### *6.2.1.1 Parallel Computation of JTI Interferograms*

The interferograms required by the JTI algorithms to construct the complex-valued FT of a single source-image are easily computed in parallel. The optical processor presented in Chapter 5 computed six interferograms in a serial fashion for each complex-valued source-image FT. This serial process can be efficiently implemented in parallel as a $2\times3$ array of optical Fourier transformers, each of which shares the same SLM and detector array.

The spectrum determined from a horizontally displayed source-reference composite image is modulated by a vertical sinusoidal pattern. Similarly, the spectrum determined from a vertically displayed source-reference composite is modulated horizontally. The horizontal and vertical spectra are combined in a weighted sum to demodulate the spectra. The pixel-by-pixel weighting assigned to each spectrum is proportional to the amplitude of the modulation at each pixel such that parts of the spectra which are modulated to zero (or close to zero) are not included in the sum (or given little weight). The architecture shown in Figure 6-1 illustrates how the six JTI interferograms can be calculated and weighted in parallel.

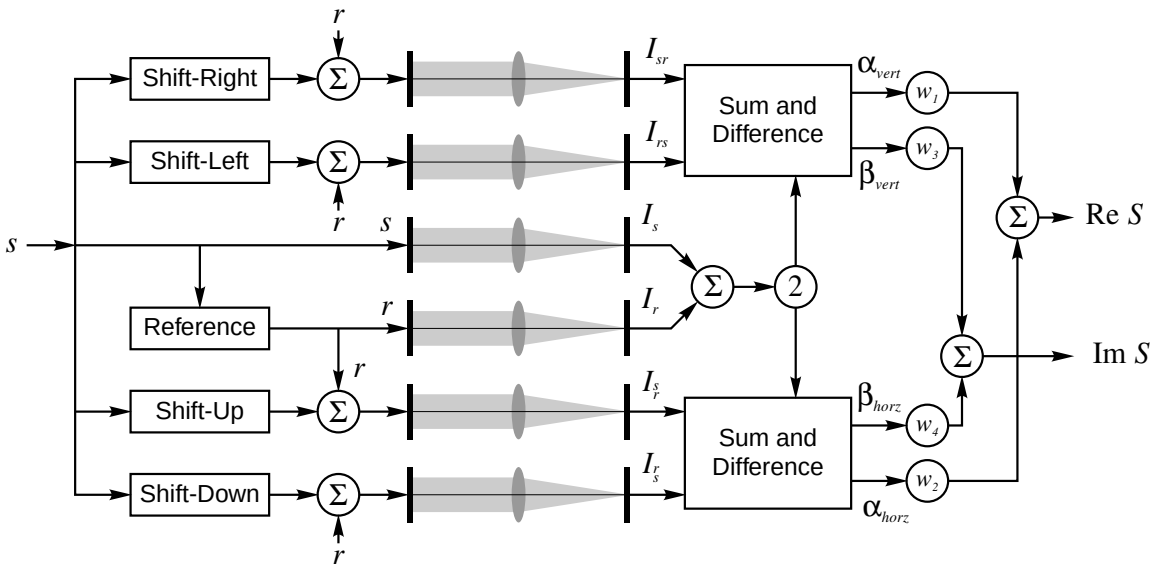



**Figure 6-1** Parallel interferometric optical Fourier transform processor.

The block diagram illustrated in Figure 6-1 is a parallel implementation of the weighted horizontal-vertical JTI algorithm described by Equations 5-12 through 5-14 and presented in Chapter 5. 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 adding the unit-shifted source images to the reference. Each of the four composite images, and the individual source and reference images, are written to the SLM of a different optical Fourier processor. The six interferograms are combined, according to Equations 5-12 through 5-14, to yield the complex-valued source spectrum. The post-processing is simple pixel-by-pixel addition and subtraction followed by pixel-by-pixel weighting.

The parallel processor could be implemented as a $2\times3$ array of optical Fourier processors. A single a high-resolution SLM could be employed to display all six composite images in parallel. A $2\times3$ lenslet array of Fourier transform lenses, or a suitable holographic optical element, would optically FT all six images, and a high-resolution detector would measure the six interferograms.

*6.2.1.2 CMOS Image Sensor Array with Smart Pixels*

The experiments presented in this dissertation employed a CCD (charge coupled device) image sensor in the back focal-plane of the Fourier transform lens. A smart-pixel CMOS (complementary metal-oxide-silicon) sensor offers three major advantages over a CCD sensor: 1) higher dynamic range; 2) the ability to integrate some or all of the post-processing electronics on the same silicon chip; and 3) the ability to individually address pixels.

The electrical dynamic range of a typical commercial-grade CCD sensor is about $50$ dB compared to $120$ dB for currently available CMOS devices. In the simple example presented in Section 4.7.1, the electrical dynamic range of a typical interferogram, produced by a $256\times256$ pixel image, was estimated to be $180$ dB. C-Cam Technologies (Heverlee, Belgium) produces a CMOS image sensor with $120$ dB of electrical dynamic range based on technology developed at IMEC (Inter-university Micro Electronics Center) [Ric95]. Experimental sensors with $140$ dB of dynamic range have also been demonstrated [Sch97]. The huge dynamic range is achieved by employing a non-linear resistor, generally a MOSFET (metal-oxide-silicon field-effect transistor) in weak inversion, in series with a pixel photodiode, to generate a logarithmic photocurrent response to incident light intensity [Ric95, Fos97].

A CMOS image sensor also offers the opportunity to integrate much of the down-stream electronics, specifically the analog-to-digital converter (ADC) and the JTI post-processing electronics, on the image sensor chip (the CMOS process is incompatible with the CCD process). Monolithic integration of the ADC circuitry with the CMOS image sensor, yielding the so called camera-on-a-chip [Fos97], represents the present state-of-the-art. Smart pixel arrays, in addition to providing a logarithmic response at each pixel in the array, can also provide other simple processing functions [Sch97]. The JTI post processing involves addition, subtraction, and a weighted sum, at each pixel, to calculate the real and

imaginary parts of an arbitrary FT spectrum. These simple operations could implemented in analog VLSI and monolithically integrated on the same silicon chip as the image sensor.

A CMOS image sensor enables parallel and random access to any pixel in its array. This is advantageous to many image processing applications which are based on processing neighboring pixel or a small number of isolated pixels. The trajectory estimation algorithm, which required the complex-valued FT at a set of pixels uniformly spaced on a frequency-space circle, could take advantage of the individually addressable pixels of a CMOS image sensor.

Although CMOS image sensors can accommodate signals of very large dynamic range, they can also produce large quantities of fixed-pattern noise (FPN). A CMOS sensor might be able to cover over six orders of magnitude in incident light level, however, its SNR can be quite small (typically $45$ dB) due primarily to temporal noise and small voltage swings [Fos97]. FPN in logarithmic CMOS sensors has been corrected to acceptable levels for normal imaging applications using both on and off-chip signal processing [Ric95]. Any noise sources added by a CMOS imager, in addition to those already present, would degrade the SNR and overall performance of the PDI and JTI-based interferometric optical Fourier processors.

#### *6.2.1.3 Reference Images*

The functional form of the JTI reference image received very little attention in this dissertation and requires further investigation. An ideal reference has uniform magnitude and phase in the frequency domain. A $\delta$-function reference image provides an ideal frequency-domain reference, however, it is not realizable on a conventional SLM (a single pixel approximates the spatial extent of a $\delta$-function however its amplitude is finite).

The references employed during much of the theoretical and experimental work were designed to be ideal in some region of interest (ROI). The SNRs for

the real and imaginary parts of the complex-valued JTI spectrum were derived in Section 4.7 assuming the reference was ideal in a specific ROI. The weighted three-component JTI algorithm was experimentally verified in Section 5.3 using a single-pixel approximation to a δ-function reference. A Bessel-image reference was employed in the JTI-based optical processor of Chapter 5 to provide a constant frequency domain reference in an annular ROI.

It is conceivable that a small number of reference images, each of which is ideal in its own ROI, could be employed to calculate an entire spectrum. Such a procedure would be similar to the demodulation of horizontally and vertically modulated spectra in a weighted sum as described by Equations 5-12 through 5-14. Can two reference images, each of which are ideal in a different ROI, be employed independently of one another to compute the source FT in that region, and then combined to produce the entire FT on the whole spectrum? How may reference images would be required? There are many questions which must be answered. Is there a generic source image that would work well for most source images or must the reference be matched to source?

The FT spectrum of a random image has constant magnitude (and random phase) and its suitability as a reference should be investigated. Smartt and Steel suggested an extended aperture, such as a random pinhole array or line source, rather than a single diffracting point, for their point-diffraction interferometer [SMA74]. They pointed out the similarities between such a device and a scatter plate interferometer [Bur53, Sco69, Hec87 pp. 378-381]. Their extended aperture PDI employed two identical scattering plates, one as an extended source, and the other, in a conjugate image plane, as the PDI mask. Application of this technique to an interferometric Fourier transformer deserves further investigation.

#### *6.2.1.4 Gray-Scale SLM*

The JTI technique is equally applicable to both gray-scale and binary SLM implementations. Its application to currently available binary SLMs necessitated

an interleaved approach; five classes of JTI, specific to a binary implementation, were presented in Section 4.3. On a gray-scale SLM, the source and reference images can be mutually shifted, summed and then written directly to the SLM, without row (and/or column) interleaving. A gray-scale implementation is therefore more efficient because it utilizes the entire space-bandwidth product of the SLM. One bit of gray-scale quantization must be forfeit, however, in order to avoid saturating the source-reference sum.

#### *6.2.1.5 Application to Synthetic Aperture Radar*

Application of the JTI method of coherent detection to synthetic aperture radar (SAR) optical processors should be investigated. The temporal radar signal from a chirp-dechirp SAR can be processed as a 2-D Fourier transform when it is spatially distributed as a 2-D polar formatted signal [Ale90, Ale95]. In general, a SAR processor takes a complex-valued input and provides a complex-valued output. In some optical SAR processors, the real and imaginary parts of the raw input signal are processed independently [Han94]. In others, the real and imaginary parts are processed together by modulating the amplitude and phase of a spatial carrier with the complex-valued input [Ale90, Ale95]. Modern all-electronic real-time SAR processors can consume as much as $1\ \mathrm{kW}$ when they are run at the throughput rates demanded by current SAR applications. Real-time optical SAR processors, which consume orders of magnitude less power, would be very attractive for air-borne or space-borne installations where large reserves of power are not available.

### 6.2.2 Moving-Object Trajectory Estimation

#### *6.2.2.1 Real-World Image Sequence*

The JTI based trajectory estimation system was evaluated using a synthetic image sequence. The moving objects were either single pixels or $3\times3$ pixel blocks, the background was composed of various geometric shapes, and the

noise was realized from a Gaussian white noise-process. The system was designed to estimate the trajectory of small, barely discernible, moving objects of unknown position and velocity; its performance should be characterized using an actual real-world image sequence. Blostein and Huang tested their trajectory estimation system on a variety of image sequences, including a road scene containing an approaching runner, another with a moving car, and a night-sky sequence (provided by the Lawrence Livermore Laboratory) containing moving satellites, meteors, and other celestial bodies [Blo91]. Choi *et al.* embedded synthetic point-objects and Gaussian white noise into a background sequence provided by the Naval Research Laboratory [Cho95, Cho97].

#### *6.2.2.2 Gray-Scale SLM*

The point-spread function (PSF), mapped into the image plane of an optical imaging system by a point-object, is a continuous intensity function and can not be accurately reconstructed using a binary SLM. In this work, the object was represented by a single binary pixel; the PSF is more correctly represented by an array of gray-scale pixels. A single-pixel representation leads to jerky motion for non-integer pixel displacements per fame and this sometimes causes trouble for the time-domain component of the mixed-domain algorithm. The jerky motion adds other temporal frequencies to the data vectors which can confuse the FBLP algorithm into thinking there are additional frequency components. The additional components make it difficult to perform target detection, which is implemented in the Knudsen and Bruton algorithm by estimating the number of sinusoidal components in the mixed-domain time sequences. An optical Fourier processor equipped with a gray-scale SLM would more accurately reconstruct the PSF of a moving object, increase the accuracy of the estimates for targets moving with non-integer displacements per frame, and allow for the detection of an arbitrary number of targets.

#### *6.2.2.3 Target Detection*

A gray-scale-capable trajectory estimation system could detect the number of moving objects in an image sequence in addition to estimating their trajectory. The number of moving objects was known *a priori* by the system presented in Chapters 5. The moving objects were tracked, however the number of objects was not detected for the reasons discussed in the previous section. An optical Fourier processor equipped with a gray-scale SLM could detect the number of moving targets. Each moving object produces a sinusoid in the mixed-domain time sequence with a different temporal frequency. Detecting the number of sinusoids in a noisy time series is not a trivial problem. Several techniques have been proposed, most notably, the minimum descriptor length (MDL) criterion [Wax85], which is used in the original Knudsen and Bruton trajectory estimation algorithm, and the information-theoretic criterion (AIC) [Aka74].

#### *6.2.2.4 Non-linear Trajectories*

A mixed-domain method for the trajectory estimation of objects moving on non-linear trajectories has been proposed and demonstrated by Knudsen and Bruton [Knu92b]. The frequency-domain processing is identical to that of the constant-velocity (linear) algorithm, however, a chirp parameterization method is employed in the time domain rather a spectral estimation method. Ideal constant-velocity (linear) objects produce pure sinusoids in the mixed domain while constant-acceleration (non-linear) objects produced chirped sinusoids. The initial frequency (proportional to the object's initial velocity) and the frequency rate (proportional to acceleration) are estimated using a chirp-parameter estimation algorithm. The Knudsen and Bruton non-linear trajectory estimation method employs the chirp estimation algorithm recently proposed by Djuric and Kay [Dju90].

## 6.3 OTHER APPLICATIONS

The Fourier transform is critically important to many signal and image processing algorithms because it is often more computationally efficient to solve problems in the frequency domain rather than the time or space domains. In order to realize these savings, fast and efficient methods, such as the FFT (fast Fourier transform) and the optical Fourier transform, are necessary to transform the input data sets to and from the frequency domain. Whether or not an optical Fourier processor is appropriate for a particular application depends on several factors including computational throughput, accuracy, dynamic range, power consumption, and space constraints. The intelligent approach to finding other applications is to take inventory of what the optical Fourier processor does well and match this with what is required by different applications. The optical Fourier transformer is an analog processor; it should be applied to problems where an analog solution is most appropriate.

Analog optical processors are ideally suited to those problems where a high degree of accuracy is not required, and a digital electronic solution would be computationally prohibitive, consume too much power, or occupy too much space. The optical Fourier transform is particularly well suited to image and other 2-D signal processing problems because of its inherent 2-D operation. The impressive computational ability of the optical Fourier transform has been emphasized in the analog optical processing literature and was one of the major motivations for the present research. Another motivation for analog optical processors, one which has not been emphasized to the same degree in the literature and deserves special recognition, is their reduced power consumption (neglecting the power consumed by the light source and additional supporting electronics).

Almost 10 years ago in 1990, Mead pointed out that a state-of-the-art digital super-computer, constructed from the ultimate silicon technology that could be

envisioned at that time, would require $10\ \mathrm{MW}$ of power to process information at the rate at which it is processed by a single human brain (the brain processes $10^{16}$ operations per second and expends $10^{-16}\ \mathrm{J}$ of energy per operation for a total power dissipation of $1\ \mathrm{W}$; a state-of-the-art microprocessor expends $10^{-9}\ \mathrm{J}$ per operation) [Mea90]. The human brain is seven orders of magnitude more energy efficient than a computationally-equivalent implementation in digital electronics because it is fundamentally an analog processor.

The brain very efficiently exploits the analog processing capabilities of simple biological processors connected in a network to perform extremely complex tasks. Similarly, the optical Fourier processor pragmatically exploits the natural law of diffraction to produce a very computationally intensive calculation. The type of analog processing in these two computing machines is very different, however, both machines take advantage of a natural computational primitive, synapses in the case of the nervous system and diffraction in the case of the optical Fourier processor, to perform the processing with great efficiency. The moving-object trajectory estimation systems demonstrated in this dissertation were in fact based on models of human visual-motion sensing and psychophysics. These types of machine vision problems seem particularly appropriate because the input is optical in nature and the optical Fourier transform is inherently a 2-D process.

The author feels that the JTI-based interferometric optical Fourier transformer developed in this dissertation may have commercial potential in machine vision applications, particularly visual-motion sensing and trajectory estimation for autonomous robots. It may also have commercial value for space-borne synthetic aperture radar (SAR), portable seismic processing, and acoustic imaging applications, where high computational throughput and low power consumption are required. The author and his supervisor are presently working in conjunction with NU-TECH (Nova Universities Technology Inc.) to evaluate the commercial potential of the novel interferometric optical Fourier processor.

# References

[US5740276] G.J. Tomko and A. Stoianov, “Hybrid optical-digital method and apparatus for fingerprint verification,” Unites States Patent assigned to Mytec Technologies Inc., Don Mills, Ontario, April 1998.

[US5761330] A. Stoianov, C. Soutar, and G.J. Tomko, “Hybrid optical-digital method and apparatus for fingerprint verification,” Unites States Patent assigned to Mytec Technologies, Inc., Don Mills, Ontario, June 1998.

[Abr83] E. Abraham, C.T. Seaton, and S.D. Smith, “The Optical Computer,” *Scientific American*, vol. 248(2), pp. 85-93, 1983.

[Ade85] E.H. Adelson and J.R. Bergen, “Spatio-temporal energy models for the perception of motion,” *J. Opt. Soc. Am. A*, vol. 2(2), pp. 284-299, 1985.

[Agg88] J.K. Aggarwal and N. Nandhakumar, “On the computation of motion from sequences of images – a review,” *Proc. of the IEEE*, vol. 76(8), pp. 917-935, Aug. 1988.

[Ahm98] F. Ahmed, M.S. Alam, and M.A. Karim, “Multireference joint transform correlator for distortion estimation,” *Optical Engineering*, vol. 37(1), pp. 38-42, January 1998.

[Aka74] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control*, vol. AC-19(6), pp. 716-723, December 1974.

[Ale90] C.C. Aleksoff and N.S. Subotic, “Compact real-time interferometric Fourier transform processors,” *Proc. of SPIE*, vol. 1347, pp. 427-440, 1990.

[Ale95] C.C. Aleksoff and N.S. Subotic, “Optical processor for SAR image formation: performance trades,” *Proc. of SPIE*, vol. 2489, pp. 142-157, 1995.

[Ama90] J. Amako and T. Sonehara, "Computer-Generated Hologram Using TFT Active Matrix Liquid Crystal Spatial Light Modulator (TFT-LCSLM)," *Japanese Journal of Applied Physics*, vol. 29(8), pp. L1533-L1535, 1990.

[Arsn84] H.H. Arsenault, Y.-N. Hsu, and K. Chalasinska-Macukow, “Rotational-invariant pattern recognition,” *Optical Engineering*, vol. 23, pp. 705-709, 1984.

[Bal82] D.H. Ballard and C.M. Brown, Computer Vision, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[Ban91] P.P. Banerjee and T.-C. Poon, Principles of Applied Optics, Richard D. Irwin Inc., and Aksen Associates, Inc., Homewood, IL, 1991.

[Bel86] T.E. Bell, “Optical computing: A field in flux,” *IEEE Spectrum*, vol. 23(8), pp. 34-57, August 1986.

[Ber97] A. Bergeron, H.H. Arsenault, M. Doucet, L. Veilleux, and D. Gingras, “Optodigital neural network classifier,” *Optical Engineering*, vol. 36(11), pp. 3134-3139, November 1997.

[Bha86] B. Bhanu, “Automatic target recognition: State of the art survey,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22(4), pp. 364-379, July 1986.

[Blo91] S.D. Blostein and T.S. Huang, “Detecting small, moving objects in image sequences using sequential hypothesis testing,” *IEEE Trans. Signal Processing*, vol. 39(7), pp. 1611-1629, July 1991.

[Boo96] S.A. Boothroyd, A. David, S. Chang, P. Palacharla, and J. Chrostowski, “Nonlinear joint transform correlator with a multiple quantum well photorefractive device,” 1996 International Topical Meeting on Optical Computing, Sendai, Japan, April 21-26.

[Bor75] M. Born and E. Wolf, Principles of Optics, Fifth Edition, Pergamon Press, Oxford, England, 1975.

[Boy91] R. M. Boysel, "A 128x128 frame-addressed deformable mirror spatial light modulator," *Optical Engineering*, vol. 30(9), pp. 1422-1427, 1991.

[Bro85] E. Brookner, “Phased-array radars,” *Scientific American*, vol. 252(2), pp. 94-102, February 1985.

[Brow69] W.M. Brown and J.L. Porcello, “An introduction to synthetic-aperture radar,” *IEEE Spectrum*, vol. 6(9), pp. 52-62, Sept. 1969.

[Bru85] L.T. Bruton and N.R. Bartley, “Three-dimensional Image Processing Using the Concept of Network Resonance,” *IEEE Transactions on Circuits and Systems*, vol. CAS-32(7), pp. 664-672, July 1985.

[Bru86] L.T. Bruton and N.R. Bartley, “The enhancement and tracking of moving objects in digital images using adaptive three-dimensional recursive filters,” *IEEE Trans. on Circuits and Systems*, vol. CAS-33(6), pp. 604-612, June 1986.

[Bur53] J.M. Burch, “Scatter fringes of equal thickness,” *Nature*, vol. 171, pp. 889-890, May 1953.

[Car69] H.R. Carleton, W.T. Maloney, and G. Meltz, “Collinear Heterodyning in Optical Processors,” *Proc. of the IEEE*, vol. 57, pp. 769-775, May 1969.

[Cas81] D.P. Casasent, “Pattern recognition: a review,” *IEEE Spectrum*, vol. 18(3), pp. 28-33, March 1981.

[Cas94a] D.P. Casasent and J. Smokelin, “Neural net design of macro Gabor wavelet filters for distortion-invariant object detection in clutter,” *Optical Engineering*, vol. 33, pp. 2264-2271, 1994.

[Cas94b] D.P. Casasent, “General-purpose optical pattern recognition image processors,” *Proc. of the IEEE*, vol. 82(11), pp. 1724-1734, Nov. 1994.

[Cho95] J.H. Choi and S.A. Rajala, “Three-dimensional location estimation of trajectories of point targets using a projection based transformation method,” *Optical Engineering*, vol.34(3), pp. 933-939, March 1995.

[Cho97] J.H. Choi, J.W. Jang, S.P Lee, and H.S. Kwak, “Multiple moving object estimation in image sequences of a natural scene,” *Optical Engineering*, vol. 36(8), pp. 2176-2183, August 1997.

[Cla92] N. Clark, K.J. Powell, and M.K. Giles, "Using Liquid Crystal Devices as Input and Filter SLMs," *Proc. of SPIE*, vol. 1704, pp. 237-247, 1992.

[Col89] D.R. Collins, J.B. Sampsell, L.J. Hornbeck, J.M. Florence, P.A Penz, and M.T. Gately, "Deformable mirror device spatial light modulators and their applicability to optical neural networks," *Applied Optics*, vol. 28(22), pp. 4900-4907, 1989.

[Cow83] A.E. Cowart, W.E. Snyder, and W.H. Ruedger, “The detection of unresolved targets using the Hough transform,” *Computer Vision, Graphics, and Image Processing*, vol. 21, pp. 222-238, 1983.

[Dav90] J.A. David, E.A. Merrill, D.M. Cottrell, and R.M. Bunch, “Effects of sampling and binarization in the output of the joint Fourier transform correlator,” *Optical Engineering*, vol. 29, pp. 1094-1100, 1990.

[Des98] S. Deschênes, Y. Sheng, and P.C. Chevrette, “Three-dimensional object recognition from two-dimensional images using wavelet transforms and neural networks,” *Optical Engineering*, vol. 37(3), pp. 763-770, March 1998.

[Dju90] P.M. Djurić and S.M. Kay, “Parameter estimation of chirp signals,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38(12), pp. 2118-2126, Dec. 1990.

[Dou96] R. Dou and M.K. Giles, “Simple technique for measuring the phase property of a twisted nematic liquid crystal television," *Optical Engineering*, vol. 35(3), pp. 808-812, 1996.

[Eva93] B.L. Evans, L.J. Karam, K.A. West, and J.H. McClellan, “Learning signals and systems with Mathematica,” *IEEE Transactions on Education*, vol. 36, pp. 72-78, February 1993.

[Fei88] D.G. Feitelson, Optical Computing – A Survey for Computer Scientists, MIT Press, Cambridge, MA, 1988.

[Fen91] D. Feng, H. Zhao, and S. Xia, “Amplitude-modulated JTC for improving correlation discrimination,” *Optics Communications*, vol. 86, pp. 260-264, 1991.

[Flo88] J.M. Florence and R.O. Gale, "Coherent optical correlator using a deformable mirror device spatial light modulator in the Fourier plane," *Applied Optics*, vol. 27(11), pp. 2091-2093, 1988.

[Flo89] J.M. Florence, "Joint-transform correlator systems using deformable-mirror spatial light modulators," *Optics Letters*, vol. 14(7), pp. 341-343, 1989.

[Flo90] J.M. Florence, T.-H Lin, W.-R. Wu, and R.D Juday, "Improved DMD configurations for image correlation," *Proc. of SPIE*, vol. 1296, pp. 101-109, 1990.

[Flo91] J.M. Florence and R.D. Juday, "Full complex spatial filtering with a phase mostly DMD," *Proc. of SPIE*, vol. 1558, pp. 487-498, 1991.

[Fos97] E.R. Fossum, “CMOS image sensors: electronic camera-on-a-chip,” *IEEE Transactions on Electron Devices*, vol. 44(10), pp. 1689-1698, Oct. 1997.

[Fry90] E. Freysz et al, “Optical wavelet transform of fractal aggregates,” *Physical Review Letters*, vol. 64(7), pp. 745-748, 1990.

[Gia92] P.D. Gianino and C.L. Woods, “Effects of spatial light modulator opaque dead zones on optical correlation,” *Applied Optics*, vol. 31, pp. 4025-4033, July 1992.

[Gia93b] P.D. Gianino and C.L. Woods, “General treatment of spatial light modulator dead-zone effects on optical correlation. II. Mathematical analysis,” *Applied Optics*, vol. 32, pp. 6536-6541, November 1993.

[Gmi80] A.F. Gmitro, J.E. Grereivenkamp, W. Swindell, H.H. Barrett, M.Y. Chiu, and S.K. Gordon, “Optical computers for reconstructing objects from their X-ray projections,” *Optical Engineering*, vol. 19, pp. 260-272, May-June 1980.

[Goo68] J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, NY, 1968.

[Goo77] J.W. Goodman, “Operations achievable with coherent optical information processing systems," *Proc. of the IEEE*, vol. 65(1), pp. 29-38, Jan. 1977.

[Gor94] C. Gorecki, I. Moreno, J. Campos, M.J. Yzuel, and K. Chalasinska-Macukow, "Real-time quantized phase-only filters generation on Epson liquid crystal TV," *Proc. of SPIE*, vol. 2297, pp. 90-101, 1994.

[Gre88] D.A. Gregory, R.D. Juday, J. Sampsell, R. Gale, R.W. Cohn, and S.E. Monroe, "Optical characteristics of a deformable-mirror spatial light modulator," *Optics Letters*, vol. 13(1), pp. 10-12, 1988.

[Gui91] P.S. Guilfoyle, “DOC II: 32-bit digital optical computer, optoelectronic hardware and software,” in *Optical Enhancements to Computing Technology*, J. Neff, Ed., Proc. SPIE, vol.1563, pp. 267-278, 1991.

[Gui93] P.S. Guilfoyle, F.F. Zeise, and J.M. Hessenbruch, “Smart optical interconnects for high speed photonic computing,” in *Optical Computing Technical Digest*, pp. 78-85, Optical Society of America, Washington DC, 1993.

[Gui98] P.S. Guilfoyle, J.M. Hessenbruch, and R.V. Stone, “Free-space interconnects for high-performance optoelectronic switching,” *IEEE Computer*, vol. 31(2), pp. 69-75, Feb. 1998.

[H&H89] P. Horowitz and W. Hill, The Art of Electronics, Second Edition, Cambridge University Press, New York, NY, 1989.

[Han94] M.W. Haney, “Compact acousto-optic processor for synthetic aperture radar image formation,” *Proc. of the IEEE*, vol. 82(11), pp. 1735-1748, Nov. 1994.

[Hau72] G. Häusler, “Optical software survey,” *Optica Acta*, vol. 24(9), pp. 965-977, Sept. 1977.

[Hay86] S. Haykin, Adaptive Filter Theory, Prentice-Hall, Englewood Cliffs, NJ 1986.

[Hec87] E. Hecht, Optics, Second Edition, Addison-Wesley, Reading, MA, 1987.

[Hor95] L.J. Hornbeck, “Digital light processing and MEMS: Timely convergence for a bright future (Invited Plenary Paper),” *Proc. of SPIE*, vol. 2639, p. 2 (abstract only), full paper available from Texas Instruments, 1995.

[Horn82] J. L. Horner, “Light utilization in optical correlators,” *Applied Optics*, vol. 21(24), pp. 4511-4514, December 1982.

[Hou62] P.V.C Hough, “Methods and means for recognizing complex patterns,” US Patent 3,069,654, 1962.

[Hsu82] Y.-N. Hsu and H.H. Arsenault, “Optical pattern recognition using circular harmonic expansion,” *Applied Optics*, vol. 21, pp. 4016-4019, 1982.

[Hud89] T.D. Hudson, D.W. Trivett, D.A. Gregory, and J.C. Kirsch, "Real time optical correlator architecture using a deformable mirror spatial light modulator," *Applied Optics*, vol. 28(22), pp. 4853-4860, 1989.

[Inf93] C. Infante, “On the modulation transfer function of matrix displays,” *Journal of the Society of Information Display*, 1/4, pp. 449-450, 1993.

[Ira98] L.J. Irakliotis and P.A. Mitkas, “Optics: A maturing technology for better computing,” Guest Editor’s Introduction to the IEEE Computer Special Issue on Optical Computing, *IEEE Computer*, vol. 31(2), pp. 36-37, Feb. 1998.

[Jav88] B. Javidi and C. Kuo, “Joint image transform correlation using a binary spatial light modulator at the Fourier plane,” *Applied Optics*, vol. 27, pp. 663-665, 1988.

[Jen77] H. Jenson, L.C Graham, L.J. Porcello, and E.N. Leith, “Side looking air-borne radar,” *Scientific American*, vol. 237(4), pp. 84-95, Oct. 1977.

[Jor94] H.F. Jordan and M.J. Murdocca, “Scanning the issue,” Introduction to the special issue on optical computing systems, *Proc. of the IEEE*, vol. 82(11), pp. 1591-1593, Nov. 1994.

[Kad94] H. Kadono, M. Ogusu, S. toyooka, “Phase shifting common path interferometer using a liquid-crystal phase modulator,” *Optics Communications* vol. 110, pp. 391-400, 1994.

[Kea98] K.J. Kearney and Z. Ninkov, “Characterization of a digital micromirror device for use as an optical mask in imaging and spectroscopy,” *Proc. of SPIE*, vol. 3292, pp. 81-92, 1998.

[Kin67] M. King, W.R. Bennett, L.B. Lambert, and M. Arm, “Real-time electrooptical signal processors with coherent detection,” *Applied Optics*, vol. 6, pp.1367-1375, Aug. 1967.

[Kir91] J.C. Kirsch, J.A. Loudin, and D.A. Gregory, "Hybrid Modulation Properties of the Epson LCTV," *Proc. of SPIE*, vol. 1558, pp. 432-441, 1991.

[Kir92] J.C. Kirsch, D.A. Gregory, M.W. Thie, and B.K. Jones, "Modulation characteristics of the Epson liquid crystal television," *Optical Engineering*, vol. 31(5), pp. 963-970, 1992.

[Kno89] J. Knopp, and R.D. Juday, "Optical joint transform correlation on the DMD," *Proc. of SPIE*, vol. 1053, pp. 208-215, 1989.

[Kno95] J. Knopp, “Modeling discrete modulators for optical correlation,” *Proc. of SPIE*, vol. 2565, pp. 185-190, 1995.

[Kno96] J. Knopp, M.V. Morelli, S.E. Monroe, R.D. Juday, "Pixel-level complex control of an Epson LCTV SLM," *Proc. of SPIE*, vol. 2752, pp. 144-152, 1996.

[Knu91] K.S. Knudsen and L.T. Bruton, “Mixed Domain Filtering of Multidimensional Signals,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 1(3), pp. 260-268, September 1991.

[Knu92a] K.S. Knudsen and L.T. Bruton, “Moving Object Detection and Trajectory Estimation in the Transform/Spatiotemporal Mixed Domain,” *Proceedings of the 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing* (San Francisco, CA), vol. 3, pp. 505-508, March 1992.

[Knu92b] K.S. Knudsen and L.T. Bruton, “Moving Object Nonlinear Trajectory Estimation in the Transform/Spatiotemporal Mixed Domain,” *Proceedings of the 1992 IEEE Custom Integrated Circuits Conference* (San Diego, CA), vol. 5, pp. 2481-2484, May 1992.

[Knu92c] K.S. Knudsen, “Multidimensional mixed domain signal processing,” Ph.D. dissertation, University of Calgary (Calgary, AB), Sept. 1992.

[Knu93a] K.S. Knudsen and L.T. Bruton, “Transform/Spatiotemporal Mixed Domain Moving Object Tracking and Enhancement,” *Proceedings of the 1993 European Conference on Circuit Theory and Design* (Davos, Switzerland), pp. 589-594, August 1993.

[Knu93b] K.S. Knudsen and L.T. Bruton, “Recursive pruning of the 2-D DFT with 3-D Signal Processing Applications,” *IEEE Transactions on Signal Processing*, vol. 41(3), pp. 1340-1356, March 1993.

[Kum86] B.V.K.V. Kumar, “Minimum variance synthetic discriminant function,” *J. Opt. Soc. Am. A*, vol. 3, pp. 1579-1584, 1986.

[Lan96a] P.M. Lane, “An implementation of an optical Fourier-transform coprocessor using a twisted-nematic liquid-crystal spatial light modulator,” M.A.Sc. thesis, Technical University of Nova Scotia (Halifax, NS), April 1996.

[Lan96b] P.M. Lane, “Development and implementation of optically based computer sub-systems,” invited paper, The Optical Processing and Computing Consortium of Canada (OPCOM) Second Annual Technical Seminar (Ottawa, ON), Nov. 6, 1996.

[Lan97a] P.M. Lane, “An optical Fourier transform coprocessor,” The Third Graduate Student Opto-Electronic Semiconductor Conference on Materials, Devices, and Applications (Vancouver, BC), June 4-6, 1997.

[Lan98] P.M. Lane, K.S. Knudsen, and M. Cada, "Moving object trajectory estimation using an optical Fourier processor," *Proc. of SPIE*, vol. 3491, pp. 939-943, July 1998.

[Lan99a] P.M. Lane and M. Cada, "An optical Fourier processor and point-diffraction interferometer for moving object trajectory estimation," *Applied Optics*, in press, 1999.

[Lan99d] P.M. Lane and M. Cada, "A novel interferometric optical Fourier transform processor," in preparation, 1999.

[Lei77] E.N. Leith and A.L. Ingalls, “Synthetic antenna data processing by wavefront reconstruction,” *Applied Optics*, vol. 7(3), pp. 539-544, March 1968.

[Lin33] W.P. Linnik, “A simple interferometer for the investigation of optical systems,” published in Russian, *C. R. Acad. Sci. URSS*, vol. 5, pp. 210, 1933.

[Lou92] J.A. Loudin and J.C. Kirsch, "Pixel level control of the Epson LCTV," *Proc. of SPIE*, vol. 1701, pp. 92-98, 1992.

[MacA98] C.E. MacAulay and A. Dlugan, “Use of digital micromirror devices in quantitative microscopy”, *Proc. of SPIE*, vol. 3260, pp. 201-206, April 1998.

[MacD99] MacDonald Dettwiler, www site, http://www.mda.ca, March 1999.

[Mah88] S.A. Mahmoud, M.S. Afifi, and R.J. Green, “Recognition and velocity computation of large moving objects in images,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36(11), pp. 1790-1791, Nov. 1988.

[Mah91] S.A. Mahmoud, “A new technique for velocity estimation of large moving objects,” *IEEE Transactions on Signal Processing*, vol. 39(3), pp. 741-743, March 1991.

[Mahal87] A. Mahalanobis, B.V.K.V. Kumar, and D.P. Casasent, “Minimum average correlation energy filters,” *Applied Optics*, vol. 26, pp. 3633-3640, 1987.

[Mal91] D. Malacara, Optical Shop Testing, Second Edition, John Wiley, New York, NY, 1991.

[Mar71] J.D. Markel, “FFT pruning,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. AU-19(4), pp. 305-311, December 1971.

[Mea90] C. Mead, “Neuromorphic electronic systems,” *Proc. of the IEEE*, vol. 78(10), pp. 1629-1636, Oct. 1990.

[Mer94] C.R. Mercer and K. Creath, “Liquid-crystal point-diffraction interferometer,” *Optics Letters*, vol.19(12), pp. 916-918, June 1994.

[Mer96] C.R. Mercer and K. Creath, “Liquid-crystal point-diffraction interferometer for wave-front measurements,” *Applied Optics*, vol. 35(10), pp. 1633-1642, April 1996.

[Moh81] N.C. Mohanty, “Computer tracking of moving point targets in space,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-3(5), pp. 606-611, Sept. 1981.

[Mor86] S. Morozumi, T. Sonehara, H. Kamakura, T. Ono, and S. Aruga, "LCD full-color projector," Society for Information Display (SID) 86 Digest, pp. 375-378, 1986.

[Nut76] A.H. Nuttall, Spectral analysis of a univariate process with bad data points via maximum entropy and linear predictive techniques, Navel Underwater Systems Center (NUSC) Scientific and Engineering Studies, Spectral Estimation, NUSC, New London, Conn., 1976.

[O&S75] A.V. Oppenheim and R.W. Schafer, Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1975.

[Pay89] J.M. Payne, “Millimeter and submillimeter wavelength radio astronomy,” *Proceedings of the IEEE*, vol. 77(7), pp. 993-1017, July 1989.

[Por90] B. Porat, and B. Friedlander, “A frequency domain algorithm for multiframe detection and estimation of dim targets,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12(4), pp. 398-401, April 1990.

[Pre93] G. Premont and Y. Sheng, “Fast design of circular-harmonic filters using simulated annealing,” *Applied Optics*, vol. 32, pp. 3116-3121, 1993.

[Ree83] I.S. Reed, R.M. Gagliardi, and H.M. Shao, “Application of three-dimensional filtering to moving target detection,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19(6), pp. 898-905, Nov. 1983.

[Rey89] G.O. Reynolds, J.B. BeVelis, G.B. Parrent, Jr., and B.J. Thompson, The New Physical Optics Notebook: Tutorials in Fourier Optics, SPIE, Bellingham, Washington, 1989.

[Ric95] N. Ricquier and B. Dierickx, “Active pixel CMOS image sensor with on-chip non-uniformity correction,” *1995 IEEE Workshop on CCDs and Advanced Image Sensors* (Dana Point, CA), April 20-22, 1995.

[S&T91] B.E.A. Saleh and M.C. Teich, Fundamentals of Photonics, John Wiley, New York, NY, 1991.

[Sch97] M. Schanz, W. Brockherde, R. Hauschild, B.J. Hosticka, and M. Schwarz, “Smart CMOS image sensors arrays*,” IEEE Transactions on Electron Devices*, vol. 44(10), pp. 1699-1705, Oct. 1997.

[Schm86] R.O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Transactions on Antennas and Propagation*, vol. AP-34(3), pp. 276-280, March 1986.

[Sco69] R. M. Scott, “Scatter plate interferometry,” *Applied Optics*, vol. 8(3), pp. 531-537, March 1969.

[Sma72] R.N. Smartt and J.Strong, “Point-Diffraction Interferometer,” *J. Opt. Soc. Am*., vol. 62, p. 737 (abstract only), 1972.

[Sma74] R.N. Smartt and W.H. Steel, “Theory and Application of Point-Diffraction Interferometers,” *Japan. J. appl. Phys.*, vol. 14 (Suppl. 14-1), pp. 351-356, 1974.

[Smit90] T. Smit, M.R. Smith, and S.T. Nichols, “Efficient sinc function interpolation technique for center padded data*,” IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp.1512-1517, Sept. 1990.

[Sou94a] C. Soutar, S.E. Monroe, and J. Knopp, "Measurement of the complex transmittance of the Epson Liquid crystal television," *Optical Engineering*, vol. 33(4), pp. 1061-1068, 1994.

[Spe79] R.J. Speer, M. Chrisp, D. Turner, S. Mrowka, and K. Tregidgo, “Grazing incidence interferometry: the use of the Linnik interferometer for testing image-forming reflection systems,” *Applied Optics*, vol. 18(12), pp. 2003-2012, June 1979.

[Sta98] M. Starzomski and M. Cada, “Application of a liquid crystal display as a multi-level quantized phase-mostly filter in a 4-*f* optical correlator,” *Proc. of SPIE*, vol. 3491, pp. 1135-1140, July 1998.

[Str96] G. Strang and T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA, 1996.

[Tam90] E.C. Tam, S.W. Tannone, F.T.S Yu, D.A. Gregory, “Closed-Loop Binary Phase Correction of an LCTV Using a Point Diffraction Interferometer*,” IEEE Photonics Technology Letters*, vol. 2(2), pp. 143-146, February 1990.

[Tan93] Q. Tang and B. Javidi, “Technique for reducing the redundant and self-correlation terms in joint transform correlators,” *Applied Optics*, vol. 32, pp. 1911-1918, 1993.

[TI96] Texas Instruments, “DLP Technology Acceptance Program Reference Manual,” P/N 4135725-0001, September 1996.

[TI99] Texas Instruments, www site, digital light processing image library, http://www.ti.com/dlp/docs/developer/resources/library/pixels_micro.shtml, January 1999.

[Tuf82] D.W. Tufts and R. Kumaresan, “Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood,” *Proceedings of the IEEE*, vol. 70(9), pp. 975-989, September 1982.

[Tur60] G.L. Turin, “An Introduction to Matched Filters,” *IRE Transactions on Information Theory*, pp. 311-329, June 1960.

[Turp81] T.M. Turpin, “Spectrum analysis using optical processing,” *Proc. of the IEEE*, vol. 69, pp. 79-92, January 1981.

[Ulr76] T.J. Ulrych and R.W. Clayton, “Time series modeling and maximum entropy,” *Phys. Earth Planet. Interiors*, vol. 12, pp. 188-200, 1976.

[Van64] A. VanderLugt, “Signal Detection by Complex Spatial Filtering,” *IEEE Transactions on Information Theory*, vol. IT-10, pp. 139-145, April 1964.

[Van74] A.B. VanderLugt, “Coherent optical processing,” *Proc. of the IEEE*, vol. 62(10), pp. 1300-1319, Oct. 1974.

[Van81] A. VanderLugt, “Interferometric spectrum analyzer,” *Applied Optics*, vol. 33(20), pp. 2770-2779, Aug. 1981.

[Van92] A. VanderLugt, Optical Signal Processing, John Wiley, New York, NY, 1992.

[Wat85] A.B. Watson and A.J. Ahumada, “Model of human visual-motion sensing,” *J. Opt. Soc. Am. A*, vol. 2(2), pp. 322-341, 1985.

[Wax85] M. Wax and T. Kailath, “Detection of Signals by Information Theoretic Criteria,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33(2), pp. 387-392, April 1985.

[Wea66] C.S. Weaver and J.W. Goodman, “A technique for optically convolving two functions,” *Applied Optics*, vol. 5(7), pp. 1248-1249, 1966.

[Yu92] F.T.S. Yu and S. Jutamulia, Optical Signal Processing, Computing, and Neural Networks, John Wiley, New York, NY, 1992.

[Zha93a] Y. Zhang, E. Kanterakis, A. Katz, and J.-M. Wang, “Smartt interferometry based optical wavelet transform and reconstruction,” *Proc. of SPIE*, vol. 1961, pp. 344-353, 1993.

[Zha93b] Y. Zhang, E. Kanterakis, A. Katz, and J.-M. Wang, “Optical wavelet processor for wavelets defined in the time domain,” *Proc. of SPIE*, vol. 2026, pp. 371-378, 1993.

[Zha94] Y. Zhang, E. Kanterakis, A. Katz, and J.-M. Wang, “Optoelectronic wavelet processors based on Smartt interferometry,” *Applied Optics*, vol. 33(23), pp. 5279-5286, August 1994.

APPENDIX

# A

# Temporal and Spatial Fourier Transforms

There is often some confusion as to the definition of the temporal and spatial Fourier transforms employed by a particular author. Definitions consistent with the engineering convention for a travelling wave have been employed in this dissertation. The 1-D temporal and 2-D spatial Fourier transforms, for continuous functions and discrete sequences, are defined in this appendix. The forward 2-D spatial transform describes the mapping provided by a Fourier transform lens. There is no inverse optical Fourier transform – repeated application of the forward transform returns a transformed image to its original spatial domain in a rotated coordinate system, $(x, y) \rightarrow (-x, -y)$.

## A.1 TEMPORAL 1-D FOURIER TRANSFORM

The forward temporal 1-D Fourier Transform of $f(t)$ is a *minus-i* transform and the inverse transform of $F(\omega)$ is a *plus-i* transform. The transform pair is defined in the conventional manner,

$$\begin{aligned} F(\omega) &= \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(x)\exp(-i\omega t)dt \\ f(t) &= \mathcal{F}^{-1}\{F(\omega)\} = (2\pi)^{-1}\int_{-\infty}^{\infty} F(\omega)\exp(i\omega t)d\omega \end{aligned} \quad \text{(A-1)}$$

where $t$ is time and $\omega$ is the temporal angular frequency.

## A.2 SPATIAL 2-D FOURIER TRANSFORM

The forward spatial 2-D Fourier Transform of $f(x,y)$ is defined with a *plus-i* exponential term and the inverse transform is defined with a *minus-i* exponential term,

$$\begin{aligned} F(k_x,k_y) &= \mathcal{F}\{f(x,y)\} = \int_{-\infty-\infty}^{\infty}\int^{\infty} f(x,y)\exp[i(k_x x+k_y y)]dx\,dy \\ f(x,y) &= \mathcal{F}^{-1}\{F(k_x,k_y)\} = (2\pi)^{-2}\int_{-\infty-\infty}^{\infty}\int^{\infty} F(k_x,k_y)\exp[-i(k_x x+k_y y)]dk_x\,dk_y \end{aligned} \quad \text{(A-2)}$$

where $x$ and $y$ are the spatial variables, and $k_x$ and $k_y$ are the spatial angular frequencies. The signs of the exponential arguments in the spatial transform pair are opposite to those of the temporal transform pair to be consistent with the engineering convention for a travelling wave. Under the engineering convention, a plane wave travelling in the positive $z$ direction has a *minus-i* spatial exponential and a *plus-i* temporal exponential [Ban91, S&T91]. The inverse transforms of Equations A-1 and A-2 are defined so that the synthesis of an arbitrary signal will have consistent signs with the engineering convention.

## A.3 DISCRETE-TIME FOURIER TRANSFORM

The discrete-time Fourier transform (DTFT) pair is defined as,

$$F(\omega) = \mathcal{F}_{DTFT}\{f(n)\} = \sum_{n=-\infty}^{\infty} f(n)\exp(-i\omega n) \quad \text{(A-3)}$$

$$f(n) \;=\; \mathscr{F}^{-1}_{DTFT}\{F(\omega)\} \;=\; (2\pi)^{-1}\int_{-\pi}^{\pi} F(\omega)\exp(i\omega n)\,d\omega$$

where $f(n)$ is a complex-valued sequence and $F(\omega)$ is a continuous $2\pi$-periodic function of spatial frequency. The symbol $n$ is employed to represent both the temporal dependence of a time-sequence $f(n)$ and the $y$ direction spatial dependence of a space-function $f(m,n)$. This was done to be consistent with conventional notation and the difference is usually obvious from the context.

## A.4 DISCRETE-SPACE FOURIER TRANSFORM

The discrete-space Fourier transform (DSFT) is analogous to the DTFT of discrete-time signal processing, however, the signs of the exponential functions have been switched to be consistent with the engineering convention for a travelling wave. The DSFT pair is defined as,

$$F(k_x,k_y) \;=\; \mathscr{F}_{DSFT}\{f(m,n)\} \;=\; \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} f(m,n)\exp\left[i(k_x m + k_y n)\right] \qquad \text{(A-4)}$$

$$f(m,n) \;=\; \mathscr{F}^{-1}_{DSFT}\{F(k_x,k_y)\} \;=\; (2\pi)^{-2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F(k_x,k_y)\exp\left[-i(k_x m + k_x n)\right]dk_x\,dk_y$$

where $f(m,n)$ is a complex 2-D sequence and $F(k_x,k_y)$ is a continuous $2\pi$-periodic function of spatial frequency. Its properties follow directly from the continuous Fourier transform and are well documented [O&S75, pp.33-34].

APPENDIX

# B

# Modified FBLP Algorithm

It is well know that the discrete Fourier transform (DFT) provides the basis for an efficient method for estimating the angular frequencies of multiple complex sinusoids embedded in noise. A high-resolution time-domain method, such as the modified forward-backward linear prediction (FBLP) algorithm [Tuf82], is preferred when the angular frequencies are more closely spaced than the reciprocal of the observation interval. The original FBLP algorithm [Ulr76, Nut76] for parametric spectral estimation was improved by Tufts and Kumaresan and is referred to as the modified FBLP algorithm. The performance of the modified algorithm for closely spaced sinusoids at low signal to noise ratios is closer to the Cramér-Rao bound than other linear prediction methods [Tuf82].

## B.1 MODIFIED FBLP METHOD

The modified forward-backward linear prediction (FBLP) method is a time-domain parametric spectral-estimation technique used to determine the angular frequency of complex sinusoids embedded in a noisy time sequence. The time sequence, which represents the superposition of $P$ component sinusoids and a white noise sequence, is modeled as,

$$u(n) \;=\; \sum_{p=1}^{P} a_p \exp(i\omega_p n) \;+\; e(n) \qquad \text{(B-1)}$$

where $\{a_p\}$ is a set of unknown complex amplitudes, $\{\omega_p\}$ is a set of unknown angular frequencies, and $e(n)$ is a white-noise sequence. The objective is to estimate the unknown angular frequencies $\omega_p$, $p = 1, 2, \ldots, P$.

## B.2 MODIFIED FBLP ALGORITHM

The algorithm presented in this appendix is based on the implementation described by Haykin [Hay86, page 349]. The input to the algorithm is an $N \times 1$ signal vector $\mathbf{u}$ and the output is a $P \times 1$ vector $\hat{\boldsymbol{\omega}}$ of angular frequency estimates. The number of component sinusoids $P$ is assumed to be known and the prediction order $K = \frac{3N}{4}$ is set to the optimum order derived empirically by Tufts and Kumaresan [Tuf82]. The algorithm is summarized in the following 5 steps.

1. Construct the $2(N-K) \times K$ data matrix $\mathbf{A}$ defined by,

$$\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} u(K) & u(K+1) & \cdots & u(N-1) & u^*(2) & u^*(3) & \cdots & u^*(N-K+1) \\ u(K-1) & u(K) & \cdots & u(N-2) & u^*(3) & u^*(4) & \cdots & u^*(N-K+2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u(1) & u(2) & \cdots & u(N-K) & u^*(M+1) & u^*(M+2) & \cdots & u^*(N) \end{bmatrix} \qquad \text{(B-2)}$$

from the $N \times 1$ signal vector $\mathbf{u}$. The superscript T denotes complex conjugate transpose and $K = \frac{3N}{4}$ represents the empirically derived optimum prediction order.

2. Determine the singular value decomposition (SVD) of the $2(N-K) \times K$ data matrix,

$$\mathbf{A} = \mathbf{Y}\mathbf{\Sigma}\mathbf{X}^{\mathsf{T}} = [\mathbf{y}_1\ \mathbf{y}_2\ \cdots\ \mathbf{y}_K] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{bmatrix} [\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_K]^{\mathsf{T}} \qquad \text{(B-3)}$$

where $\mathbf{Y}$ is a $2(N-K) \times K$ unitary matrix, $\mathbf{\Sigma}$ is a $K \times K$ diagonal matrix, and $\mathbf{X}$ is a $K \times K$ unitary matrix. The columns $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ of matrix $\mathbf{Y}$ are called the left singular vectors of $\mathbf{A}$, the elements $\lambda_1, \lambda_2, \ldots, \lambda_K$ on the main diagonal of $\mathbf{\Sigma}$ are called the singular values of $\mathbf{A}$, and the columns $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K$ of $\mathbf{X}$ are called the right singular vectors of $\mathbf{A}$. The right singular vectors and singular values correspond, respectively, to the eigenvalues and eigenvectors of the $K \times K$ correlation matrix $\mathbf{R} = \mathbf{A}^{\mathsf{T}}\mathbf{A}$.

3. Identify the $P$ largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_P$ in $\mathbf{\Sigma}$ which correspond to the $P$ complex sinusoids components of the signal vector $\mathbf{u}$. Use the eigenvalues and eigenvectors to compute an estimate of the $K \times 1$ tap-weight vector for the linear predictor,

$$\hat{\mathbf{w}} = \sum_{p=1}^{P} \frac{\mathbf{x}_p}{\lambda_p} \mathbf{x}_p \mathbf{A}^{\mathsf{T}} \mathbf{b} \qquad \text{(B-4)}$$

where $\mathbf{b}$ is the $2(N-K) \times 1$ desired response vector defined by,

$$\mathbf{b}^{\mathrm{T}} = \left[u(K+1),\, u(K+2),\, \ldots,\, u(N),\, u^*(1),\, u^*(2),\, \ldots,\, u^*(N-K)\right] \qquad \text{(B-5)}$$

4. Compute the $(K+1)\times 1$ tap-weight vector for the prediction-error filter,

$$\mathbf{a} = \begin{bmatrix} 1 \\ -\hat{\mathbf{w}} \end{bmatrix} \qquad \text{(B-6)}$$

5. Retrieve the angular frequency estimates of the complex sinusoids from the sample spectrum of the prediction-error filter. First determine the $K$ roots $z_1, z_2, \ldots, z_K$ of the polynomial,

$$a_1 z^K + a_2 z^{K-1} + \cdots + a_{K-1} z^2 + a_K z + a_{K+1} \qquad \text{(B-7)}$$

from the denominator of the prediction-error filter system function. The $P$ roots $z_1, z_2, \ldots, z_P$ which are on (or closest to) the unit circle on the $z$-plane correspond to the estimates of the angular frequencies,

$$\hat{\omega}_p = \operatorname{atan}\left(\operatorname{Im} z_p, \operatorname{Re} z_p\right), \quad p = 1, 2, \ldots, P \qquad \text{(B-8)}$$

APPENDIX

# C

# JTI Equations and Algorithms

Selected JTI equations and the algorithms necessary to extract the real and imaginary parts of the source spectrum were developed and presented in Section 4.3. The equations and algorithms of all possible source-reference image configurations are summarized in this appendix. The sine and cosine spatial modulation functions are summarized in Section C.1. The JTI spectra are divided into five classes depending on the spatial configuration of the source and reference images and the number of interfering images. Expressions for the JTI spectra, and the corresponding extraction algorithms, for the five different classes, are presented in Sections C.2 through C.6.

## C.1 MODULATION FUNCTIONS

The real and imaginary parts of the source-image spectrum are modulated by cosine and sine modulation functions, respectively. The modulation functions are simple $\sin$ and $\cos$ functions for horizontal and vertical JTI, however they are more complicated for the other JTI classes.

### C.1.1 Sine Modulation Functions

$$
\begin{aligned}
s_1 &= \sin(k_x - k_y) \\
s_2 &= \sin(k_x + k_y) \\
s_3 &= \sin(k_x) - \sin(k_y) \\
s_4 &= \sin(k_x) + \sin(k_y) \\
s_5 &= \sin(k_x) - \sin(k_y) + \sin(k_x - k_y) = s_1 + s_3 \\
s_6 &= \sin(k_x) + \sin(k_y) + \sin(k_x + k_y) = s_2 + s_4
\end{aligned}
\qquad \text{(C-1)}
$$

### C.1.2 Cosine Modulation Functions

$$
\begin{aligned}
c_1 &= \cos(k_x - k_y) \\
c_2 &= \cos(k_x + k_y) \\
c_3 &= \cos(k_x) - \cos(k_y) \\
c_4 &= \cos(k_x) + \cos(k_y) \\
c_5 &= \cos(k_x) + \cos(k_y) + \cos(k_x - k_y) = c_1 + c_4 \\
c_6 &= \cos(k_x) + \cos(k_y) + \cos(k_x + k_y) = c_2 + c_4
\end{aligned}
\qquad \text{(C-2)}
$$

## C.2 HORIZONTAL AND VERTICAL JTI

A single source image is interfered with a single reference image. The source and reference images are interfered with respect to a line which can be either horizontal or vertical. If the source is displaced to the left or right of the reference, the interference is called vertical JTI, however, if the source is displaced above or below the reference, it is called horizontal JTI.

### C.2.1 JTI Equations

$$\begin{aligned}
I_{\substack{s\,r\\0\,0}} &= I_{\substack{0\,0\\s\,r}} = I_s + I_r + 2\left[\cos(k_x)\,\Re(SR^*) + \sin(k_x)\,\Im(SR^*)\right] \\
I_{\substack{r\,s\\0\,0}} &= I_{\substack{0\,0\\r\,s}} = I_s + I_r + 2\left[\cos(k_x)\,\Re(SR^*) - \sin(k_x)\,\Im(SR^*)\right] \\
I_{\substack{s\,0\\r\,0}} &= I_{\substack{0\,s\\0\,r}} = I_s + I_r + 2\left[\cos(k_y)\,\Re(SR^*) + \sin(k_y)\,\Im(SR^*)\right] \\
I_{\substack{r\,0\\s\,0}} &= I_{\substack{0\,r\\0\,s}} = I_s + I_r + 2\left[\cos(k_y)\,\Re(SR^*) - \sin(k_y)\,\Im(SR^*)\right]
\end{aligned} \tag{C-3}$$

where the source and reference intensities are given by,

$$\begin{aligned}
I_s &= I_{\substack{0\,0\\s\,0}} = I_{\substack{0\,0\\0\,s}} = I_{\substack{0\,s\\0\,0}} = I_{\substack{s\,0\\0\,0}} \\
I_r &= I_{\substack{0\,0\\r\,0}} = I_{\substack{0\,0\\0\,r}} = I_{\substack{0\,r\\0\,0}} = I_{\substack{r\,0\\0\,0}}
\end{aligned} \tag{C-4}$$

### C.2.2 Algorithms for Imaginary Part of the Spectrum

$$\begin{aligned}
4\sin(k_x)\,\Im(SR^*) &= I_{\substack{s\,r\\0\,0}} - I_{\substack{r\,s\\0\,0}} \\
4\sin(k_y)\,\Im(SR^*) &= I_{\substack{s\,0\\r\,0}} - I_{\substack{r\,0\\s\,0}}
\end{aligned} \tag{C-5}$$

### C.2.3 Algorithms for Real Part of the Spectrum

$$\begin{aligned}
4\cos(k_x)\,\Re(SR^*) &= I_{\substack{s\,r\\0\,0}} + I_{\substack{r\,s\\0\,0}} - 2I_s - 2I_r \\
4\cos(k_y)\,\Re(SR^*) &= I_{\substack{s\,0\\r\,0}} + I_{\substack{r\,0\\s\,0}} - 2I_s - 2I_r
\end{aligned} \tag{C-6}$$

## C.3 DIAGONAL JTI

A single source image is diagonally interfered with a single reference image. The source and reference images may be interfered with respect to a line which is diagonal with either positive or negative slope. If the source and reference are displaced along a diagonal line with 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).

### C.3.1 JTI Equations

$$
\begin{aligned}
I_{\substack{s\,0\\0\,r}} &= I_s + I_r + 2\left[c_2\,\Re\left(SR^*\right) + s_2\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\,0\\0\,s}} &= I_s + I_r + 2\left[c_2\,\Re\left(SR^*\right) - s_2\,\Im\left(SR^*\right)\right] \\
I_{\substack{0\,r\\s\,0}} &= I_s + I_r + 2\left[c_1\,\Re\left(SR^*\right) + s_1\,\Im\left(SR^*\right)\right] \\
I_{\substack{0\,s\\r\,0}} &= I_s + I_r + 2\left[c_1\,\Re\left(SR^*\right) - s_1\,\Im\left(SR^*\right)\right]
\end{aligned}
\qquad \text{(C-7)}
$$

### C.3.2 Algorithms for the Imaginary Part of the Spectrum

$$
\begin{aligned}
4s_2\,\Im\left(SR^*\right) &= I_{\substack{s\,0\\0\,r}} - I_{\substack{r\,0\\0\,s}} \\
4s_1\,\Im\left(SR^*\right) &= I_{\substack{0\,r\\s\,0}} - I_{\substack{0\,s\\r\,0}}
\end{aligned}
\qquad \text{(C-8)}
$$

### C.3.3 Algorithms for the Real Part of the Spectrum

$$
\begin{aligned}
4c_2\,\Re\left(SR^*\right) &= I_{\substack{s\,0\\0\,r}} + I_{\substack{r\,0\\0\,s}} - 2I_s - 2I_r \\
4c_1\,\Re\left(SR^*\right) &= I_{\substack{0\,r\\s\,0}} + I_{\substack{0\,s\\r\,0}} - 2I_s - 2I_r
\end{aligned}
\qquad \text{(C-9)}
$$

## C.4 THREE-FUNCTION JTI

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.

### C.4.1 JTI Equations

$$
\begin{aligned}
I_{\substack{0\,s\\ s\,r}} &= 2(1+c_1)I_s + I_r + 2\left[c_4\,\Re\left(SR^*\right) + s_4\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\,s\\ s\,0}} &= 2(1+c_1)I_s + I_r + 2\left[c_4\,\Re\left(SR^*\right) - s_4\,\Im\left(SR^*\right)\right] \\
I_{\substack{s\,r\\ 0\,s}} &= 2(1+c_2)I_s + I_r + 2\left[c_4\,\Re\left(SR^*\right) + s_3\,\Im\left(SR^*\right)\right] \\
I_{\substack{s\,0\\ r\,s}} &= 2(1+c_2)I_s + I_r + 2\left[c_4\,\Re\left(SR^*\right) - s_3\,\Im\left(SR^*\right)\right] \\
I_{\substack{s\,r\\ r\,0}} &= I_s + 2(1+c_1)I_r + 2\left[c_4\,\Re\left(SR^*\right) + s_4\,\Im\left(SR^*\right)\right] \\
I_{\substack{0\,r\\ r\,s}} &= I_s + 2(1+c_1)I_r + 2\left[c_4\,\Re\left(SR^*\right) - s_4\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\,0\\ s\,r}} &= I_s + 2(1+c_2)I_r + 2\left[c_4\,\Re\left(SR^*\right) + s_3\,\Im\left(SR^*\right)\right] \\
I_{\substack{r\,s\\ 0\,r}} &= I_s + 2(1+c_2)I_r + 2\left[c_4\,\Re\left(SR^*\right) - s_3\,\Im\left(SR^*\right)\right]
\end{aligned}
\tag{C-10}
$$

where the double-source and double-reference intensities are given by,

$$
\begin{aligned}
I_{\substack{0\,s\\ s\,0}} &= 2(1+c_1)I_s \\
I_{\substack{0\,r\\ r\,0}} &= 2(1+c_1)I_r \\
I_{\substack{r\,0\\ 0\,r}} &= 2(1+c_2)I_r \\
I_{\substack{s\,0\\ 0\,s}} &= 2(1+c_2)I_s
\end{aligned}
\tag{C-11}
$$

### C.4.2 Algorithms for the Imaginary Part of the Spectrum

$$
\begin{aligned}
4s_4\,\Im\left(SR^*\right) &= I_{\substack{0\,s\\ s\,r}} - I_{\substack{r\,s\\ s\,0}} \\
4s_3\,\Im\left(SR^*\right) &= I_{\substack{s\,r\\ 0\,s}} - I_{\substack{s\,0\\ r\,s}} \\
4s_4\,\Im\left(SR^*\right) &= I_{\substack{s\,r\\ r\,0}} - I_{\substack{0\,r\\ r\,s}} \\
4s_3\,\Im\left(SR^*\right) &= I_{\substack{r\,0\\ s\,r}} - I_{\substack{r\,s\\ 0\,r}}
\end{aligned}
\tag{C-12}
$$

### C.4.3 Algorithms for the Real Part of the Spectrum

$$
\begin{aligned}
4c_4\,\Re\left(SR^*\right) &= I_{\substack{0\ s\\ s\ r}} + I_{\substack{r\ s\\ s\ 0}} - 2I_{\substack{0\ s\\ s\ 0}} - 2I_r \\
4c_4\,\Re\left(SR^*\right) &= I_{\substack{s\ r\\ 0\ s}} + I_{\substack{s\ 0\\ r\ s}} - 2I_{\substack{0\ s\\ 0\ s}} - 2I_r \\
4c_4\,\Re\left(SR^*\right) &= I_{\substack{s\ r\\ r\ 0}} + I_{\substack{0\ r\\ r\ s}} - 2I_s - 2I_{\substack{0\ r\\ r\ 0}} \\
4c_4\,\Re\left(SR^*\right) &= I_{\substack{r\ 0\\ s\ r}} + I_{\substack{r\ s\\ 0\ r}} - 2I_s - 2I_{\substack{r\ 0\\ 0\ r}}
\end{aligned}
\qquad \text{(C-13)}
$$

## C.5 SYMMETRIC JTI

Symmetric JTI is a special case of 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 due to the special symmetry of the joint image.

### C.5.1 JTI Equations

$$
\begin{aligned}
I_{\substack{s\ r\\ r\ s}} &= 2(1+c_2)I_s + 2(1+c_1)I_r + 4c_4\,\Re\left(SR^*\right) \\
I_{\substack{r\ s\\ s\ r}} &= 2(1+c_1)I_s + 2(1+c_2)I_r + 4c_4\,\Re\left(SR^*\right)
\end{aligned}
\qquad \text{(C-14)}
$$

### C.5.2 Algorithms for the Real Part of the Spectrum

$$
\begin{aligned}
4c_4\,\Re\left(SR^*\right) &= I_{\substack{s\ r\\ r\ s}} - I_{\substack{s\ 0\\ 0\ s}} - I_{\substack{0\ r\\ r\ 0}} \\
4c_4\,\Re\left(SR^*\right) &= I_{\substack{r\ s\\ s\ r}} - I_{\substack{0\ s\\ s\ 0}} - I_{\substack{r\ 0\\ 0\ r}}
\end{aligned}
\qquad \text{(C-15)}
$$

## C.6 FOUR-FUNCTION JTI

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.

### C.6.1 JTI Equations

$$
\begin{aligned}
I_{\substack{s\,s\\s\,r}} &= (3+2c_5)I_s + I_r + 2\left[c_6\,\Re\left(SR^*\right) + s_6\,\Im\left(SR^*\right)\right]\\
I_{\substack{r\,s\\s\,s}} &= (3+2c_5)I_s + I_r + 2\left[c_6\,\Re\left(SR^*\right) - s_6\,\Im\left(SR^*\right)\right]\\
I_{\substack{s\,r\\s\,s}} &= (3+2c_6)I_s + I_r + 2\left[c_5\,\Re\left(SR^*\right) + s_5\,\Im\left(SR^*\right)\right]\\
I_{\substack{s\,s\\r\,s}} &= (3+2c_6)I_s + I_r + 2\left[c_5\,\Re\left(SR^*\right) - s_5\,\Im\left(SR^*\right)\right]\\
I_{\substack{s\,r\\r\,r}} &= I_s + (3+2c_5)I_r + 2\left[c_6\,\Re\left(SR^*\right) + s_6\,\Im\left(SR^*\right)\right]\\
I_{\substack{r\,r\\r\,s}} &= I_s + (3+2c_5)I_r + 2\left[c_6\,\Re\left(SR^*\right) - s_6\,\Im\left(SR^*\right)\right]\\
I_{\substack{r\,r\\s\,r}} &= I_s + (3+2c_6)I_r + 2\left[c_5\,\Re\left(SR^*\right) + s_5\,\Im\left(SR^*\right)\right]\\
I_{\substack{r\,s\\r\,r}} &= I_s + (3+2c_6)I_r + 2\left[c_5\,\Re\left(SR^*\right) - s_5\,\Im\left(SR^*\right)\right]
\end{aligned}
\qquad \text{(C-16)}
$$

where the triple-source and triple-reference intensities are given by,

$$
\begin{aligned}
I_{\substack{0\,s\\s\,s}} &= (3+2c_5)I_s\\
I_{\substack{s\,s\\0\,s}} &= (3+2c_6)I_s\\
I_{\substack{0\,r\\r\,r}} &= (3+2c_5)I_r\\
I_{\substack{r\,r\\0\,r}} &= (3+2c_6)I_r
\end{aligned}
\qquad \text{(C-17)}
$$

### C.6.2 Algorithms for the Real Part of the Spectrum

$$
\begin{aligned}
4s_6\,\Im\left(SR^*\right) &= I_{\substack{s\ s\\ s\ r}} - I_{\substack{r\ s\\ s\ s}} \\
4s_5\,\Im\left(SR^*\right) &= I_{\substack{s\ r\\ s\ s}} - I_{\substack{s\ s\\ r\ s}} \\
4s_6\,\Im\left(SR^*\right) &= I_{\substack{s\ r\\ r\ r}} - I_{\substack{r\ r\\ r\ s}} \\
4s_5\,\Im\left(SR^*\right) &= I_{\substack{r\ r\\ s\ r}} - I_{\substack{r\ s\\ r\ r}}
\end{aligned}
\qquad \text{(C-18)}
$$

### C.6.3 Algorithms for the Imaginary Part of the Spectrum

$$
\begin{aligned}
4c_6\,\Re\left(SR^*\right) &= I_{\substack{s\ s\\ s\ r}} + I_{\substack{r\ s\\ s\ s}} - 2I_{\substack{0\ s\\ s\ s}} - 2I_r \\
4c_5\,\Re\left(SR^*\right) &= I_{\substack{s\ r\\ s\ s}} + I_{\substack{s\ s\\ r\ s}} - 2I_{\substack{s\ s\\ 0\ s}} - 2I_r \\
4c_6\,\Re\left(SR^*\right) &= I_{\substack{s\ r\\ r\ r}} + I_{\substack{r\ r\\ r\ s}} - 2I_s - 2I_{\substack{0\ r\\ r\ r}} \\
4c_5\,\Re\left(SR^*\right) &= I_{\substack{r\ r\\ s\ r}} + I_{\substack{r\ s\\ r\ r}} - 2I_s - 2I_{\substack{r\ r\\ 0\ r}}
\end{aligned}
\qquad \text{(C-19)}
$$