Low-Complexity Multi-Dimensional Filters for Plenoptic Signal Processing

by

Chamira Udaya Shantha Edussooriya
B.Sc.Eng., University of Moratuwa, Sri Lanka, 2008
M.A.Sc., University of Victoria, Canada, 2012

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

Low-Complexity Multi-Dimensional Filters for Plenoptic Signal Processing

by

Chamira Udaya Shantha Edussooriya
B.Sc.Eng., University of Moratuwa, Sri Lanka, 2008
M.A.Sc., University of Victoria, Canada, 2012

Supervisory Committee

Dr. Leonard T. Bruton, Co-Supervisor
(Department of Electrical and Computer Engineering)

Dr. Panajotis Agathoklis, Co-Supervisor
(Department of Electrical and Computer Engineering)

Dr. Yang Shi, Outside Member
(Department of Mechanical Engineering)

Supervisory Committee

Dr. Leonard T. Bruton, Co-Supervisor
(Department of Electrical and Computer Engineering)

Dr. Panajotis Agathoklis, Co-Supervisor
(Department of Electrical and Computer Engineering)

Dr. Yang Shi, Outside Member
(Department of Mechanical Engineering)

ABSTRACT

Five-dimensional (5-D) light field video (LFV) (also known as plenoptic video) is a more powerful form of representing information of dynamic scenes compared to conventional three-dimensional (3-D) video. In this dissertation, the spectra of moving objects in LFVs are analyzed, and it is shown that such moving objects can be enhanced based on their depth and velocity by employing 5-D digital filters, what is defined as *depth-velocity* filters. In particular, the spectral region of support (ROS) of a Lambertian object moving with constant velocity and at constant depth is shown to be a skewed 3-D hyperfan in the 5-D frequency domain. Furthermore, it is shown that the spectral ROS of a Lambertian object moving at non-constant depth can be approximated as a sequence of ROSs, each of which is a skewed 3-D hyperfan, in the 5-D continuous frequency domain.

Based on the spectral analysis, a novel 5-D finite-extent impulse response (FIR) depth-velocity filter and a novel ultra-low complexity 5-D infinite-extent impulse response (IIR) depth-velocity filter are proposed for enhancing objects moving with constant velocity and at constant depth in LFVs. Furthermore, a novel ultra-low complexity 5-D IIR adaptive depth-velocity filter is proposed for enhancing objects moving at non-constant depth in LFVs. Also, an ultra-low complexity 3-D linear-phase IIR velocity filter that can be incorporated to design 5-D IIR depth-velocity

filters is proposed. To the best of the author's knowledge, the proposed 5-D FIR and IIR depth-velocity filters and the proposed 5-D IIR adaptive depth-velocity filter are the first such 5-D filters applied for enhancing moving objects in LFVs based on their depth and velocity.

Numerically generated LFVs and LFVs of real scenes, generated by means of a commercially available Lytro light field (LF) camera, are used to test the effectiveness of the proposed 5-D depth-velocity filters. Numerical simulation results indicate that the proposed 5-D depth-velocity filters outperform the 3-D velocity filters and the four-dimensional (4-D) depth filters in enhancing moving objects in LFVs. More importantly, the proposed 5-D depth-velocity filters are capable of exposing heavily occluded parts of a scene and of attenuating noise significantly. Considering the ultra-low complexity, the proposed 5-D IIR depth-velocity filter and the proposed 5-D IIR adaptive depth-velocity filter have significant potentials to be employed in real-time applications.

Contents

Supervisory Committee ii

Abstract iii

Table of Contents v

List of Tables ix

List of Figures x

List of Abbreviations xvii

Acknowledgements xviii

Dedication xix

1 Introduction 1

1.1 Related Work . . . 3

1.1.1 3-D Velocity Filtering . . . 4

1.1.2 4-D Depth Filtering . . . 4

1.1.3 5-D Light Field Videos . . . 5

1.1.4 Spectral Analysis of the 7-D Plenoptic Function . . . 5

1.2 Contributions of the Dissertation . . . 6

1.3 Outline of the Dissertation . . . 9

2 Analysis of the Spectra of Moving Objects in Light Field Videos 11

2.1 Introduction . . . 11

2.1.1 Notation . . . 12

2.2 Review of LF Representation of a Lambertian Scene and Its Spectrum 13

2.2.1 Two-Plane Parameterization of a LF . . . 13

2.2.2 LF Representation of a Lambertian Point Source and Its Spectrum . . . 15
2.2.3 ROS of the Spectrum . . . 16
2.3 Analysis of the Spectrum of a Lambertian Object Moving with Constant Velocity . . . 18
2.3.1 LFV Representation of a Lambertian Point Source . . . 18
2.3.2 Spectrum of a Lambertian Point Source Moving at a Constant Depth . . . 21
2.3.3 ROS of the Spectrum of a Lambertian Object Moving at a Constant Depth . . . 23
2.3.4 ROS of the Spectrum Corresponding to a Discrete-Domain LFV 24
2.3.5 Numerical Simulation of the ROS of the Spectrum . . . 26
2.4 Summary . . . 27

3 Novel 5-D FIR Depth-Velocity Filter 30
3.1 Introduction . . . 30
3.2 Design of the 5-D FIR Depth-Velocity Filter . . . 31
3.2.1 Design of the 5-D FIR Filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ . . . 32
3.2.2 Design of the 5-D FIR Filter $H_{uvt}(\mathbf{z})$ . . . 35
3.3 Experimental Results . . . 37
3.3.1 Numerically Generated LFV . . . 37
3.3.2 Lytro-LF-Camera-Based LFV . . . 40
3.3.3 Performance Limitations of Proposed 5-D FIR Depth-Velocity Filters . . . 46
3.4 Effects of Non-Lambertian Reflections and Occlusion . . . 47
3.5 Summary . . . 48

4 Novel Ultra-Low Complexity 5-D IIR Depth-Velocity Filter 50
4.1 Introduction . . . 50
4.2 Review of Design of the 3-D IIR Planar Filters . . . 51
4.3 Design of the 5-D IIR Depth-Velocity Filter . . . 54
4.4 Experimental Results . . . 57
4.4.1 Numerically Generated LFV . . . 57
4.4.2 Lytro-LF-Camera-Based LFV . . . 59

4.4.3 Performance Limitations of Proposed 5-D IIR Depth-Velocity Filters . . . 63
4.5 Summary . . . 64

5 Novel Ultra-Low Complexity 5-D IIR Adaptive Depth-Velocity Filter 66
5.1 Introduction . . . 66
5.2 Analysis of the Spectrum of a Lambertian Object Moving at Non-Constant Depth . . . 67
5.2.1 Approximate Representation of a Lambertian Point Source Moving at Non-Constant Depth Using Time-Variant 3-D Hyperplanes 67
5.2.2 Approximate Spectrum of a Lambertian Object Moving at Non-Constant Depth and Its ROS . . . 69
5.3 Design of the 5-D IIR Adaptive Depth-Velocity Filter . . . 70
5.4 Experimental Results . . . 73
5.4.1 Numerically Generated LFV . . . 73
5.4.2 Lytro-LF-Camera-Based LFV . . . 77
5.5 Summary . . . 81

6 Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter 82
6.1 Introduction . . . 82
6.2 Review of LT Signals and Their Spectra . . . 85
6.3 Proposed Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter 86
6.4 Design and Efficient Implementation of the Ultra-Low Complexity 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank . . . 88
6.4.1 Design of the 1-D Temporal Modified DFT Filter Bank . . . 89
6.4.2 Design of the 2-D Spatial Filters . . . 90
6.4.3 Near-Perfect Reconstruction of the 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank . . . 91
6.4.4 Efficient Implementation of the 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank . . . 91
6.5 Numerical Simulation Results . . . 94
6.5.1 A Design Example of the Proposed Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter . . . 94

6.5.2 Performance in Enhancing a Heavily Corrupted Approximately LT Signal . . . 94
6.5.3 Attenuation of Sunlight Flicker Patterns in Shallow Underwater Videos . . . 97
6.5.4 Comparison with Previously Reported 3-D Velocity Filters . . 100
6.6 Summary . . . 102

7 Conclusions and Future Work **104**
7.1 Conclusions . . . 104
7.2 Future Work . . . 107

A Derivation of the Spectrum of a Lambertian Point Source Moving at a Constant Depth **109**

B Derivation of the Ideal Infinite-Extent Impulse Response of $H_{xu}(\mathbf{z})$ **111**

C Derivation of the Ideal Infinite-Extent Impulse Response of $H_{uvt}(\mathbf{z})$ **113**

D Proof of the Near-Perfect Reconstruction of the 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank **115**

Bibliography **118**

List of Tables

Table 3.1 Specifications of the Lambertian planar objects. 38

Table 5.1 Specifications of the Lambertian planar objects. 73

Table 6.1 Nontrivial real multiplications and additions required to process a sample by the different blocks of the 3-D wide-angle linear-phase IIR cone filter bank. 93

Table 6.2 Numbers of nontrivial real multiplications and additions required to process a sample by different 3-D velocity filters. 101

List of Figures

Figure 1.1 The 7-D plenoptic function describes the intensity of light rays passing through the center of an ideal camera at every possible location in the 3-D space (x, y, z), at every possible angle (θ, ϕ), for every wavelength λ and at every time t. 2

Figure 1.2 (a) The Stanford camera array (Source - http://graphics.stanford.edu/); (b) a Lytro Illum LF camera (Source - https://www.lytro.com/illum/); (c) a Raytrix LFV camera (Source - http://www.raytrix.de/). 3

Figure 2.1 The two-plane parameterization of a Lambertian scene comprised of an object having Lambertian surfaces; (a) with globally defined image coordinates (u, v); (b) with locally defined image coordinates (u, v). 14

Figure 2.2 (a) Two-plane parametrization of a Lambertian point source of intensity l_0; (b) the representation of the Lambertian point source in the xu subspace. 16

Figure 2.3 The ROS of the spectrum of a Lambertian point source, $\mathcal{P}_{4C}$, (a) $\mathcal{H}_{4C,xu}$ in the $\Omega_x\Omega_u$ subspace; (b) $\mathcal{H}_{4C,yv}$ in the $\Omega_y\Omega_v$ subspace; as z_0 varies in the range $(0, \infty)$, α varies in the range $(0°, 90°)$. 17

Figure 2.4 The ROS of the spectrum of a Lambertian object, $\mathcal{O}_{4C}$, (a) in the $\Omega_x\Omega_u$ subspace; (b) in the $\Omega_y\Omega_v$ subspace. The angle of the 3-D hyperfan depends on the depth range $z_0 \in [d_{min}, d_{max}]$ occupied by the Lambertian object. 18

Figure 2.5 (a) A Lambertian point source of intensity l_0 moves with the constant velocity $\mathbf{V} = [V_x, V_y, V_z]^{\mathrm{T}}$; (b) the representation of the Lambertian point source in the xu subspace. 19

Figure 2.6 The ROS of the spectrum, $\mathcal{P}_{5C}$, (a) $\mathcal{H}_{5C,xu}$ in the $\Omega_x\Omega_u$ subspace; (b) $\mathcal{H}_{5C,yv}$ in the $\Omega_y\Omega_v$ subspace; as z_0 varies in the range $(0,\infty)$, α varies in the range $(0°, 90°)$. (c) $\mathcal{H}_{5C,uvt}$ in the $\Omega_u\Omega_v\Omega_t$ subspace. 22
Figure 2.7 The ROS of the spectrum, $\mathcal{O}_{5C}$, (a) in the $\Omega_x\Omega_u$ subspace; (b) in the $\Omega_y\Omega_v$ subspace; (c) in the $\Omega_u\Omega_v\Omega_t$ subspace. 24
Figure 2.8 The epipolar-plane images of the generated LFV (a) 10th frame; (b) 40th frame. 26
Figure 2.9 The ROS of the spectrum of the numerically generated LFV. The magnitude of the spectrum is normalized, and the iso-surface is drawn at 0.05. 28

Figure 3.1 Structure of the 5-D FIR depth-velocity filter. 32
Figure 3.2 The ROS of the spectrum of the object of interest (solid) and the passband (cross-hatched) of (a) $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane; (b) $H_{yv}(\mathbf{z})$ on the $\omega_y\omega_v$ plane; (c) $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space. 33
Figure 3.3 Frequency-domain specifications of the ROS of $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane. The parameter a_u determines the orientation of the 4-D hyperplaner passband. b_x and b_u are the bandwidths of the filter along the ω_x and ω_u dimensions, respectively, and $b_u = b_x/a_u$. 34
Figure 3.4 The ROS of $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane for $a_u \in (0,1)$. 35
Figure 3.5 Frequency-domain specifications of the ROS of $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space. The parameters a_u and a_v determine the orientation of the 4-D hyperplaner passband. b_u, b_v and b_t are the bandwidths of the filter along the ω_u, ω_v and ω_t dimensions, respectively. Note that $b_u = b_t/|a_u|$ and $b_v = b_t/|a_v|$. 36
Figure 3.6 The 15th frame of the numerically generated LFV corresponding to the central 25 sub-apertures ($n_x, n_y = 3, 4, \ldots, 7$). The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with a different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at a different depth) of the object of interest. . . 39

Figure 3.7 Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane. 40
Figure 3.8 Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$. 40
Figure 3.9 The 15th frame of (a) the original, (b) the 4-D FIR depth filtered, (c) the 3-D FIR velocity filtered and (d) the 5-D FIR depth-velocity filtered (proposed) LFVs corresponding to the central sub-aperture ($n_x, n_y = 5$). The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with a different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at a different depth) of the object of interest. 41
Figure 3.10 The experimental setup employed to generate the LFV of a real scene. A Lytro LF camera was employed to capture the individual frames of the scene. The white truck (object of interest) and the red truck (moving interfering object) move at approximately the same depth. The fence is a static interfering object. 42
Figure 3.11 The 20th frame of the LFV corresponding to the central 15 sub-apertures ($n_x = 4, 5, 6$ and $n_y = 3, 4, \ldots, 7$). The white truck is the object of interest; the red truck is the moving interfering object; the fence is the static interfering object. 43
Figure 3.12 Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane. 44
Figure 3.13 Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$. 44
Figure 3.14 Three frames of the original LFV (top row) and the filtered LFV (bottom row), $n_x, n_y = 5$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame. 45
Figure 3.15 Three frames of the corrupted LFV (top row) and the filtered LFV (bottom row), $n_x, n_y = 5$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame. 45

Figure 4.1 First-order 3-D pseudo-passive RL network. 52
Figure 4.2 Structure of the 5-D IIR depth-velocity filter. 54
Figure 4.3 The ROS of the spectrum of the object of interest (solid) and the passbands of the 5-D IIR filters (cross-hatched); (a) $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane; (b) $H_{yv}(\mathbf{z})$ on the $\omega_y\omega_v$ plane; (c) $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space. 55
Figure 4.4 Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane. 58
Figure 4.5 Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$. 58
Figure 4.6 The 15th frame of (a) the original, (b) the 4-D IIR depth filtered [44], (c) the 3-D IIR velocity filtered [24] and (d) the 5-D IIR depth-velocity filtered (proposed) LFVs corresponding to the central sub-aperture $(n_x, n_y) = (5, 5)$. The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with a different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at a different depth) of the object of interest. 60
Figure 4.7 Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane. 61
Figure 4.8 Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$. 61
Figure 4.9 Three frames of the original LFV (top row) and the 5-D IIR depth-velocity filtered LFV (bottom row), $(n_x, n_y) = (7, 7)$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame. 62
Figure 4.10 Three frames of the corrupted LFV (top row) and the 5-D IIR depth-velocity filtered LFV (bottom row), $(n_x, n_y) = (7, 7)$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame. 62

Figure 5.1 Structure of the proposed 5D IIR adaptive depth-velocity filter. 70

Figure 5.2 The 10th frame of the numerically generated LFV corresponding to the central 25 sub-apertures ($n_x, n_y = 3, 4, \ldots, 7$). The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at different depth) of the object of interest. 74

Figure 5.3 Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 50$ in the $\omega_x\omega_u$ plane. 75

Figure 5.4 Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 50$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$. 75

Figure 5.5 The 10th frame of (a) the original, (b) the 4-D IIR adaptive depth filtered (c) the 3-D IIR adaptive velocity filtered [25] and (d) the 5-D IIR adaptive depth-velocity filtered (proposed) LFVs corresponding to the central sub-aperture ($n_x, n_y = 5$). The middle object is the object of interest; the top object is the interfering object moving at the same apparent velocity (but at different depth) of the object of interest; the bottom object is the interfering object moving at the same depth (but with different apparent velocity) of the object of interest. 76

Figure 5.6 The experimental setup employed to generate the LFV of a real scene. A Lytro LF camera was employed to capture the individual frames of the scene. The red truck is the object of interest whereas the white truck and the fence are a moving interfering object and a static interfering object, respectively. . 77

Figure 5.7 The 40th frame of the LFV corresponding to the central 15 sub-apertures ($n_x = 4, 5, 6$ and $n_y = 3, 4, \ldots, 7$). The red truck is the object of interest; the white truck is the moving interfering object; the fence is the static interfering object. 78

Figure 5.8 Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 30$ in the $\omega_x\omega_u$ plane. 79

Figure 5.9 Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 30$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$. 79

Figure 5.10 Three frames of the original LFV (top row) and the 5-D IIR adaptive depth-velocity filtered LFV (bottom row), $(n_x, n_y) = (7, 7)$; (a) and (d) 30th frame; (b) and (e) 40th frame; (c) and (f) 50th frame. 80

Figure 6.1 The ROS of the spectrum of the object of interest (solid) and the 5-D hyperfan-shaped passband of the 5-D filter $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space. 83

Figure 6.2 (a) An LT signal corresponding to an object moving with a constant 2-D spatial velocity $[v_x, v_y]^{\mathrm{T}}$; (b) the region of support of its spectrum inside the principal Nyquist cube. 86

Figure 6.3 Proposed ultra-low complexity 3-D linear-phase IIR velocity filter. 87

Figure 6.4 Ideal passband of the 3-D IIR filter $C(z_x, z_y, z_t)$, which is the exterior of a wide-angle cone, where ϵ is the angle between the $\omega_x\omega_y$ plane and the surface of the wide-angle cone. 87

Figure 6.5 (a) The 3-D wide-angle linear-phase IIR cone filter bank. (b) Approximation of the passband using a 2-D spatial allpass filter and $2M_h$ 2-D spatial highpass filters; a planar view on the $\omega_x\omega_t$ plane. 89

Figure 6.6 Efficient realization of the 3-D wide-angle linear-phase IIR cone filter bank. The 1-D temporal modified DFT filter bank is realized by two 1-D temporal DFT-polyphase filter banks. . . 93

Figure 6.7 Magnitude response of the 3-D wide-angle linear-phase IIR cone filter bank (a) -3 dB iso-surface; (b) a slice at $\omega_t = 0.0313\pi$; (c) a slice at $\omega_y = 0$. 95

Figure 6.8 The 275th frame of (a) input video (b) output video. 96

Figure 6.9 The 100th frame of the (a) "pool" underwater video; (b) "caesarea" underwater video. 98

Figure 6.10 The 100th frame of the processed "pool" underwater video (a) proposed method; (b) method in [109]. 98

Figure 6.11 The 100th frame of the processed "caesarea" underwater video (a) proposed method; (b) method in [109]. 99

Figure 6.12 The average PSNR of the "pool" and "caesarea" underwater videos obtained with different levels of random 2-D spatial-shift errors, which are uniformly distributed in the range $[-b, b]$ pixels. 100

List of Abbreviations

1-D	One-Dimensional
2-D	Two-Dimensional
3-D	Three-Dimensional
4-D	Four-Dimensional
5-D	Five-Dimensional
6-D	Six-Dimensional
7-D	Seven-Dimensional
AWGN	Additive White Gaussian Noise
BIBO	Bounded-Input Bounded-Output
BRDF	Bidirectional Reflectance Distribution Function
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
FIR	Finite-Extent Impulse Response
IDFT	Inverse Discrete Fourier Transform
IIR	Infinite-Extent Impulse Response
LF	Light Field
LFV	Light Field Video
LT	Linear Trajectory
PSNR	Peak-Signal-to-Noise Ratio
RGB	Red, Green and Blue
RL	Inductor-Resistor
ROS	Region of Support
SINR	Signal-to-Interference-and-Noise Ratio

ACKNOWLEDGEMENTS

First, I would like to express my heartfelt gratitude to the co-supervisors Dr. Leonard T. Bruton and Dr. Panajotis Agathoklis for their mentorship, advice, inspiring discussions and patience. Furthermore, I really appreciate their kind support and the encouragement provided to me during hard times. Besides those, I admire their endless commitment to the advancement of the field of multidimensional signal processing.

My strength is my family. I take this moment to express my heartfelt thanks to my beloved wife Adeehsa, parents, parents-in-law, sister and brothers-in-law for their unconditional love, constant support and encouragement.

Next, I wish to thank the course instructors Dr. Jens Bornemann, Dr. Michael Adams, Dr. Nikitas Dimopoulos, and especially Dr. Andreas Antoniou and Dr. Wu-Sheng Lu, for their outstanding teaching and inspiration. I also greatly acknowledge the assistance received from the staff of the Department of Electrical and Computer Engineering including Ms. Amy Rowe, Ms. Janice Closson, Mr. Dan Mai, Ms. Tanya Threlfall, Ms. Moneca Bracken, Ms. Vicky Smith and Ms. Lynne Barrett. Furthermore, special thanks go to Mr. Kevin Jones and Mr. Erik Laxdal for their kind assistance provided to me whenever I had technical issues.

Also, I take this opportunity to thank Dr. Chulantha Kulasekere and Dr. Rohan Munasinghe at the University of Moratuwa, Sri Lanka, and Dr. Arjuna Madanayake at the University of Calgary (currently at the University of Akron) for all the assistance provided me to open the door to graduate studies.

Victoria and UVic itself are gorgeous places. However, life would have been dull and boring if I had not had companionship with a nice group of humans: my colleagues and friends. I take this opportunity to express sincere gratitude to the colleague Ioana Sevcenco for her support, encouragement and wonderful friendship, and to thank the colleagues Dr. Soltan Alharbi, Hussam Shubayli, Dr. Iman Moazzen, Rajat Gupta, Himika Rahman, Le Liang, Leyuan Pan, Yongyu Dai and Hongrui Wang. Furthermore, I wish to thank the families of Dr. Deepal Samarajeewa, Mr. Gamini Fonseka and Mr. Sisira Kosgoda for all the support given to me during my stay in Victoria.

Finally, I greatly acknowledge the financial support received from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Victoria to pursue this endeavour.

DEDICATION

To schools
Royal College, Horana, Sri Lanka
and
Taxila Central College, Horana, Sri Lanka,
where I received the primary and the secondary education, respectively,

and

to the most pleasant person I have ever met,
my maternal grandpa.

Chapter 1

Introduction

Light is a fundamental form of conveying information. The light rays emanating from a scene is completely described by the seven-dimensional (7-D) *plenoptic function*, proposed by Adelson and Bergen [1]. The name *plenoptic* has been derived by combining the Latin term *plenus* (meaning complete or full) with the term *optic* [1]. More specifically, the 7-D plenoptic function describes the intensity of light rays passing through the center of an ideal camera at every possible location in the three-dimensional (3-D) space (x, y, z), at every possible angle (θ, ϕ), for every wavelength λ and at every time t [1], as shown in Figure 1.1.

The two-dimensional (2-D) images and panoramas, the 3-D videos and concentric mosaics, the four-dimensional (4-D) light fields (LFs) and the five-dimensional (5-D) light field videos (LFVs) (also known as plenoptic videos) are simplified forms of the 7-D plenoptic function [2] [3]. In particular, the 5-D LFV is derived from the 7-D plenoptic function by assuming the intensity of a light ray does not change along its direction of propagation and the red, green and blue (RGB) colour components are used instead of the wavelength [2]. An LFV may be considered as a 2-D array of 3-D conventional videos and can be generated by employing a specifically designed LFV camera or a collection of conventional 3-D video cameras arranged as a 2-D planar array. A typical planar array of 3-D conventional video cameras, a commercially-available LF camera and a LFV camera are shown in Figure 1.2. Ideally, an LFV of a dynamic scene in free space contains all information of the scene because it captures all the light rays emanating from the scene. This richness of information may be exploited to accomplish novel tasks that are not possible with conventional 3-D videos such as digital refocusing and *depth-velocity filtering*, a technique proposed in this dissertation to enhance moving objects in LFVs based on their velocity and

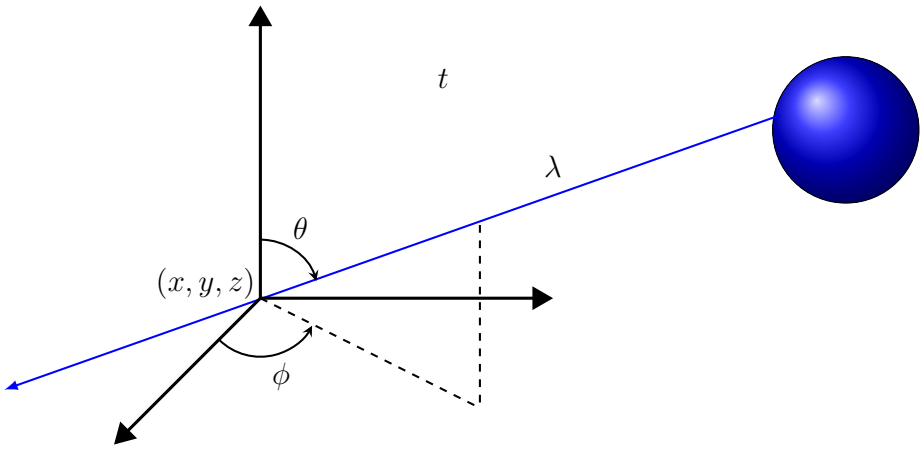
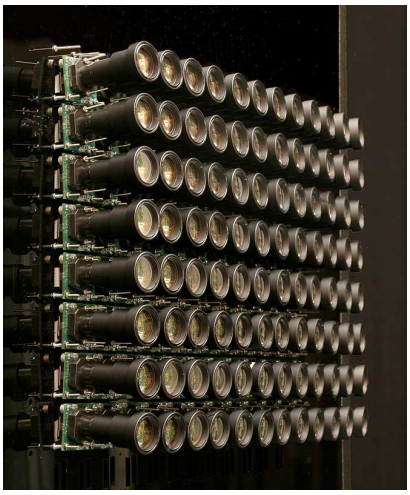


Figure 1.1: The 7-D plenoptic function describes the intensity of light rays passing through the center of an ideal camera at every possible location in the 3-D space (x, y, z), at every possible angle (θ, ϕ), for every wavelength λ and at every time t.

depth by employing 5-D *depth-velocity* filters.

The 5-D depth-velocity filters have significant potential to be employed in a multitude of applications. An exiting feature offered by the 5-D depth-velocity filters is the capability of exposing heavily occluded moving objects in a dynamic scene. Thanks to this exiting feature, the 5-D depth-velocity filters naturally fit for intelligent surveillance and security systems [4] [5] [6] [7]. For example, in such systems, the 5-D depth-velocity filters can be used as a preprocessing technique in tracking of moving objects (e.g., humans and vehicles) regardless of occlusion. Furthermore, the 5-D depth-velocity filters can be employed as a preprocessing technique in various computer vision applications. For example, they may be employed in the vision system of a robot [8] [9]. Another potential application is the refocusing of LFVs based on both depth and velocity; in other words, controlling the defocus (or out-of-focus) blur and the motion-blur of a dynamic scene [10] [11]. For this case, the frequency response of the 5-D depth-velocity filters needs to be slightly modified so that objects lying outside the focal region are blurred rather than completely attenuated. Furthermore, the focal region may be selected, for example, as an elliptical region around a human, a circular region around a vehicle or an arbitrary shape around an object of interest. Such refocusing has been exploited to improve the artistic quality of 3-D movies and 3-D television shows [12] [13]. Furthermore, considering the fact that an LFV camera gathers more light compared to a conventional video camera [14], LFVs have been emerged as an attractive replacement for conventional videos especially un-



(a)

(b)

(c)

Figure 1.2: (a) The Stanford camera array (Source - http://graphics.stanford.edu/); (b) a Lytro Illum LF camera (Source - https://www.lytro.com/illum/); (c) a Raytrix LFV camera (Source - http://www.raytrix.de/).

der low light conditions, for example in underwater scenes [15] [16] [17]. This opens an another potential area for the 5-D depth-velocity filtering.

1.1 Related Work

A brief review of previous work related to 3-D velocity filters, 4-D depth filters, 5-D LFVs and the spectral analysis of the 7-D plenoptic function is presented in this section.

1.1.1 3-D Velocity Filtering

In many applications including traffic analysis, motion detection, radar tracking and computer vision, enhancement of moving objects is generally accomplished by employing 3-D velocity filters (also known as 3-D linear-trajectory filters) [18] [19] [20] [21] [22] [23]. Such filters selectively filter moving objects based on their 2-D spatial velocities. A number of design methods for 3-D velocity filters are found in the literature. In [24], a 3-D infinite-extent impulse response (IIR) filter having a planar passband has been proposed based on the concept of network resonance to enhance moving objects having linear trajectories, i.e. objects moving with constant velocities. In [25], the work presented in [24] has been extended to 3-D IIR adaptive filters for tracking and enhancing moving objects with general trajectories. Another IIR filter design technique for tracking moving objects with linear or nonlinear trajectories has been presented in [26]. In [27], an optimal 3-D finite-extent impulse response (FIR) filter for detecting moving objects with linear trajectories has been proposed. Two 3-D IIR filter banks having wedge-shaped (exterior of a wide-angle cone) passbands have been reported in [28] and [29]. An IIR velocity filter and an IIR velocity filter bank designed using multidimensional wave digital filters have been presented in [22] and [30], respectively.

1.1.2 4-D Depth Filtering

4-D LFs have been used for image-based rendering systems [31] [32], refocusing in digital photography [33] [34] [35] [36] [37] and exposing occluded regions in a scene [38] [39] [40] [41] [42]. Another important application is enhancing objects of a scene based on their depth. This has been first demonstrated in [43] using a 4-D depth filter having a planar passband. Similar results have been reported in [38] [44] and [45]. In particular, an efficient 4-D IIR filter was proposed in [44]. In [46], a 4-D IIR dual-fan filter is proposed to selective filtering of objects occupied in a range of depths. A 4-D hyperfan all-in-focus filter is employed in [47] in order to denoise LFs. Moreover, in [37], 4-D depth filters are utilized to obtain volumetric focus for LFs instead of traditional planar focus. Two hardware implementations of first-order 4-D IIR hyperplanar filters that can be employed for depth filtering of LFs have been proposed in [48] and [49]. Furthermore, differential-form and integral-form 4-D IIR depth filters and their hardware implementations have been presented in [41] and [50], respectively. Recently, a 4-D FIR filter, which can be employed for real-time refocusing

of LFs, and its hardware architecture are proposed in [51].

1.1.3 5-D Light Field Videos

Most of the early work on LFVs was about capturing and rendering systems. In [52] and [53], a system consisting of 100 conventional video cameras, arranged rectangularly on a plane, is presented. In [54], another system has been described. This system was comprised of eight conventional video cameras arranged in a line, and real-time LFV compression and decompression has also been discussed. Recently, an LFV capturing system, where conventional video cameras were arranged in a hemisphere, has been presented in [55]. This system is capable of recoding LFVs at 30 fps with a resolution of 9000×2400 pixels. A true LFV camera, especially suitable for smart phones, is reported in [56]. A motion-aware LFV camera that can capture high resolution LFVs by exploiting scene-specific redundancy in space, time and angle is presented in [57]. Another LFV recoding system consisting of a camera array is reported in [58]. In that system, a spatio-temporal exposure pattern is employed to capture high dynamic range LFVs. In the context of motion analysis in LFVs, 3-D motion estimation has been studied in [59] and [60]. In addition, recently, closed-form solutions for visual odometry for LFV cameras have been reported in [61].

1.1.4 Spectral Analysis of the 7-D Plenoptic Function

Some of the earliest work in spectral analysis of the plenoptic function has been presented in [62] and [63] in the context of LFs and image-based rendering. In particular, for a Lambertian scene, it is shown that the region of support (ROS) of the spectrum of an LF is bounded by the minimum and maximum depths. In [34] and [64], the ROS of the spectrum of an LF is shown to be a 3-D manifold in the 4-D frequency space, which is denoted as a 3-D hyperfan in [47]. In [65], the spectral analysis of the plenoptic function was extended to non-Lambertian and occluded scenes exploiting the concept of the so-called six-dimensional (6-D) surface plenoptic function[1]. It is shown that non-Lambertian reflections and occlusion broaden the ROS of the spectrum of an LF. Furthermore, it was asserted that, in most cases, the spectrum is not band limited even though the surface plenoptic function may be band limited. In [66], the bandwidth of the plenoptic function has been more precisely analyzed,

[1]The 6-D surface plenoptic function is a reparameterization of the 7-D plenoptic function under the assumption of the intensity of a light ray does not change along its direction of propagation [65].

and shown that the plenoptic function is not band limited unless the scene is a flat surface. Moreover, it is shown that the bandwidth of the plenoptic function depends on the maximum surface slope in addition to the minimum and maximum depths of the scene and the maximum frequency of the texture of the scene. In [67] and [68], the bandwidth of the plenoptic function has been examined under the finite field-of-view constraint of cameras and finite scene width. It is shown that these finite constraints lead to band-unlimited plenoptic spectra even for Lambertian scenes having only a flat surface. Spectrum broadening due to non-Lambertian reflections and occlusion is studied in [69] and [70], and a new sampling rate and new reconstruction filters have been proposed. Furthermore, in [71], the effect of the resolution of the sampling cameras in a camera array is studied.

1.2 Contributions of the Dissertation

In this dissertation, the spectra of moving objects in LFVs are analyzed, and it is shown that such moving objects can be enhanced based on their depth and velocity by employing 5-D depth-velocity filters. In particular, it is shown that the spectral ROS of a Lambertian object moving with constant velocity and at constant depth is a skewed 3-D hyperfan in the 5-D frequency domain. Furthermore, a novel 5-D FIR depth-velocity filter, a novel ultra-low complexity 5-D IIR depth-velocity filter and a novel ultra-low complexity 5-D IIR adaptive depth-velocity filter are proposed. *To the best of the author's knowledge, these three 5-D depth-velocity filters are the first such 5-D filters applied to enhance moving objects in LFVs based on depth and velocity.* Moreover, an ultra-low complexity 3-D IIR velocity filter that can be incorporated to design 5-D IIR depth-velocity filters is proposed and used in enhancing shallow underwater videos. In what follows is detailed descriptions of the contributions of the dissertation.

Spectral Analysis of Moving Objects in LFVs (part of this work has been published in [72])

The spectrum of a Lambertian object moving with constant velocity in an LFV is analyzed. First, it is shown that a Lambertian point source moving with a constant velocity is represented as a 3-D hypersurface of constant value in the continuous-domain LFVs. Next, it is shown that when the motion of the Lambertian point

source is parallel to the camera plane (i.e., the Lambertian point source moves at a constant depth), the 3-D hypersurface is reduced to a 3-D hyperplane. For this case, the spectrum of the LFV and its ROS are derived in *closed form*. The ROS of the spectrum is a plane through the origin in the 5-D continuous frequency domain. Based on the analysis for a Lambertian point source, the ROS of the spectrum of a continuous-domain LFV, corresponding to a Lambertian object moving with a constant velocity and at a constant depth, is derived. For this case, it is shown that the ROS of the spectrum is a *skewed 3-D hyperfan* in the 5-D continuous frequency domain. The degree of skewness of the 3-D hyperfan depends on both velocity and depth of the moving object whereas the angle of the 3-D hyperfan depends on the depth range occupied by the moving object. The 3-D hyperfans corresponding to the ROSs of the spectra of Lambertian objects moving with different constant velocities or at different constant depths do not overlap except at the origin in the 5-D frequency domain. This allows enhancement of such objects based on depth and velocity by employing 5-D digital filters, what is defined as *depth-velocity* filters. Furthermore, it is shown that the essential bandwidth of the spectrum is finite along the temporal frequency dimension and, therefore, the corresponding discrete-domain LFV can be generated with negligible aliasing by employing a sufficiently high temporal sampling rate. The analysis is concluded by deriving the ROS of the spectrum for a discrete-domain LFV sampled with negligible aliasing and illustrating it through numerical simulations.

Novel 5-D FIR Depth-Velocity Filter (part of this work has been published in [72])

A novel 5-D linear-phase FIR depth-velocity filter having a planar passband in the 5-D frequency domain is proposed. The planar passband of the proposed 5-D FIR depth-velocity filter is realized by cascading three 5-D linear-phase FIR filters having 4-D hyperplanar passbands of appropriate orientations in the 5-D discrete frequency domain. The 5-D FIR filters having 4-D hyperplanar passbands are designed by employing the so-called windowing method [73](ch. 3.3) [74](ch. 5.1). The performance of the proposed 5-D FIR depth-velocity filter in enhancing moving objects in LFVs is tested by employing a numerically generated LFV and an LFV of a real scene, generated by means of a commercially available Lytro LF camera. Experimental results confirm the effectiveness of the proposed 5-D FIR depth-velocity filters over 4-D FIR

depth filters and 3-D FIR velocity filters.

Novel Ultra-Low Complexity 5-D IIR Depth-Velocity Filter (part of this work has been published in [75])

A novel ultra-low complexity 5-D IIR depth-velocity filter is proposed for enhancing objects moving with constant velocity and at constant depth in LFVs. The proposed 5-D IIR depth-velocity filter is realized by cascading three first-order 5-D IIR filters having 4-D hyperplanar passbands of appropriate orientations. The first-order 5-D IIR filters are designed by appropriately extending the first-order 3-D IIR planar filter design method proposed in [24]. The proposed 5-D IIR depth-velocity filter is practical bounded-input bounded-output (BIBO) stable. Numerical simulation results indicate that the proposed 5-D IIR depth-velocity filter outperforms the 3-D IIR velocity filters [24] and the 4-D IIR depth filters [44] in enhancing moving objects in LFVs. Furthermore, by employing the LFV generated using the commercially available Lytro LF camera, it is shown that the performance of the proposed 5-D IIR depth-velocity filter is comparable to that of the proposed 5-D FIR depth-velocity filter. *Most importantly, the proposed 5-D IIR depth-velocity filter requires less than* 1% *of the arithmetic operations required by the 5-D FIR depth-velocity filter to process a sample.* Considering the ultra-low complexity, the proposed 5-D IIR depth-velocity filter has a significant potential to be employed in *real-time* applications.

Novel Ultra-Low Complexity 5-D IIR Adaptive Depth-Velocity Filter (part of this work has been published in [76])

The 5-D depth-velocity filtering technique is extended to a more general case, where objects moving with constant velocity but at *non-constant depth* are enhanced. First, the spectrum of a Lambertian object moving in an LFV at non-constant depth is analyzed, and it is shown that the ROS of the spectrum can be approximated as a sequence of ROSs, each of which is a skewed 3-D hyperfan, in the 5-D continuous frequency domain. Based on this analysis, a novel ultra-low complexity 5-D IIR adaptive depth-velocity filter is proposed for enhancing such moving objects. The proposed 5-D IIR adaptive depth-velocity filter is realized by cascading three first-order 5-D IIR adaptive filters having time-variant 4-D hyperplanar passbands of appropriate orientations. The first-order 5-D IIR adaptive filters are designed by appropriately extending the first-order 3-D IIR adaptive planar filter design method proposed in [25]. The

time-variant coefficients of the three first-order 5-D IIR adaptive filters are derived in *closed form*. The performance of the proposed 5-D IIR adaptive depth-velocity filter is confirmed by employing a numerically generated LFV and an LFV of a real scene, generated by means of a commercially available Lytro LF camera. Experimental results indicate that the proposed 5-D IIR adaptive depth-velocity filter outperforms the 3-D IIR adaptive velocity filters [25] and the 4-D IIR adaptive depth filters in enhancing moving objects in LFVs. Considering the ultra-low complexity and the availability of the closed-form expressions for the time-variant coefficients, the proposed 5-D IIR adaptive depth-velocity filter has a significant potential to be employed in *real-time* applications.

Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter (part of this work has been published in [77] and [78])

An ultra-low complexity 3-D linear-phase IIR velocity filter that can be incorporated to design 5-D IIR depth-velocity filters is proposed. The proposed 3-D linear-phase IIR velocity filter consists of an ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank between two 2-D spatial variable-shift filters. The ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank is designed by employing a one-dimensional (1-D) temporal modified discrete Fourier transform (DFT) filter bank and 2-D spatial allpass, IIR highpass and allstop filters. The linear phase response is achieved by employing zero-phase filtering for the 2-D spatial IIR highpass filters. A typical 3-D linear-phase IIR velocity filter of order $4 \times 4 \times 510$ requires only 26 real multiplications and 60 real additions to process a sample, which is significantly lower compared to the previously reported 3-D FIR and IIR velocity filters [27] [28] [29]. In order to illustrate the performance of the proposed 3-D linear-phase IIR velocity filter in enhancing videos, it is employed to attenuate sunlight flicker patterns in shallow underwater videos. Experimental results confirm the effectiveness of the 3-D linear-phase IIR velocity filter in attenuating the sunlight flicker patterns and its robustness to motion estimation errors.

1.3 Outline of the Dissertation

The organization of the rest of the dissertation is briefly presented in this section, and a detailed outline of a chapter is presented in the beginning of each chapter.

In Chapter 2, the analysis of the spectra of object moving with constant velocity in LFVs is presented. The novel 5-D FIR depth-velocity filter is described in detail in Chapter 3. Furthermore, the experimental results obtained using a numerically generated LFV and a Lytro-LF-camera-based LFV are presented. In Chapter 4, the novel ultra-low complexity 5-D IIR depth-velocity filter and the experimental results obtained using the numerically generated LFV and the Lytro-LF-camera-based LFV are presented. Next, in Chapter 5, the approximate form of the spectrum of a Lambertian object moving in an LFV at non-constant depth is presented. More over, the design of the novel ultra-low complexity 5-D IIR adaptive depth-velocity filter is described in detail. Also, the experimental results obtained using a numerically generated LFV and an Lytro-LF-camera-based LFV are presented. In Chapter 6, the ultra-low complexity 3-D linear-phase IIR velocity filter and the experimental results corresponding to the enhancement of shallow underwater videos are presented. Finally, conclusions and future work are presented in Chapter 7.

Chapter 2

Analysis of the Spectra of Moving Objects in Light Field Videos

2.1 Introduction

The spectrum of an LFV that corresponds to a Lambertian object moving with constant velocity is analyzed in this chapter. Furthermore, it is shown that such moving objects can be enhanced based on their depth and velocity by employing 5-D depth-velocity filters.

A Lambertian object may be assumed to be comprised of Lambertian surfaces[1]. A Lambertian surface may be considered as a collection of Lambertian point sources and, therefore, the LF of a Lambertian object is given by the superposition of the LFs of corresponding Lambertian point sources. Consequently, in the analysis presented in this chapter, we mainly pay our attention to the LF and LFV representations of *Lambertian point sources* and their spectra.

We begin the analysis by showing that a Lambertian point source moving with a constant velocity (i.e., having a linear trajectory) is represented as a 3-D hypersurface of constant value in the continuous-domain LFVs. Next, it is shown that when the motion of the Lambertian point source is parallel to the camera plane (i.e., the Lambertian point source moves at a constant depth), the 3-D hypersurface is reduced to a 3-D hyperplane. For this case, we derive the spectrum of the LFV and its ROS in closed form. The ROS of the spectrum is a plane through the origin in the 5-D

[1] A Lambertian surface scatters incoming light uniformly in all directions. In other words, the bidirectional reflectance distribution function (BRDF) of a Lambertian surface is constant [79](ch. 2.2).

continuous frequency domain. Based on the analysis for a Lambertian point source, the ROS of the spectrum of a continuous-domain LFV, corresponding to a Lambertian object moving with a constant velocity and at a constant depth, is derived. For this case, it is shown that the ROS of the spectrum is a skewed 3-D hyperfan in the 5-D continuous frequency domain. The degree of skewness of the 3-D hyperfan depends on both velocity and depth of the moving object whereas the angle of the 3-D hyperfan depends on the depth range occupied by the moving object. Furthermore, it is shown that the essential bandwidth of the spectrum is finite along the temporal frequency dimension and, therefore, the corresponding discrete-domain LFV can be generated with negligible aliasing by employing a sufficiently high temporal sampling rate. The analysis is concluded by deriving the ROS of the spectrum for a discrete-domain LFV sampled with negligible aliasing and illustrating it through numerical simulations.

The rest of the chapter is organized as follows. In Section 2.2, LF parameterization, the LF representation of a Lambertian scene and its spectrum are reviewed. In Section 2.3, a detailed analysis of the LFV representation of a Lambertian point source moving with constant velocity and its spectrum is presented. Furthermore, the ROS of the spectrum of an LFV corresponding to a Lambertian object moving with constant velocity and at constant depth is derived and illustrated through numerical simulations. Finally, a summary of the chapter is presented in Section 2.4.

2.1.1 Notation

The following notation scheme is employed in this dissertation. Lowercase letters are used to denote 4-D LF and 5-D LFV signals whereas uppercase counterparts are used to denote their spectra. In addition, an alphanumeric subscript scheme is utilized in order to avoid the ambiguity of the dimension and domain of the LF and LFV signals (and their spectra) denoted by the same letter. The alphanumeric subscripts are comprised of two elements: a number, denoting the dimension of the signal, followed by the uppercase letter "C" or "D", denoting the continuous and discrete domains, respectively. For example, $l_{4\mathrm{C}}(\cdot)$ is a 4-D continuous-domain LF signal whereas $l_{5\mathrm{D}}(\cdot)$ is a 5-D discrete-domain LFV signal. Their spectra are denoted by $L_{4\mathrm{C}}(\cdot)$ and $L_{5\mathrm{D}}(\cdot)$, respectively. Vectors and matrices are denoted by lowercase and uppercase bold letters, respectively. The superscript "T" is employed to denote the transpose of a vector or a matrix. Furthermore, uppercase calligraphic letters in conjunction with the above mentioned alphanumeric subscript scheme are employed

to denote sets.

2.2 Review of LF Representation of a Lambertian Scene and Its Spectrum

We review the LF representation of a Lambertian scene and its spectrum, i.e. the Fourier transform, in this section. To this end, we first consider the standard *two-plane parameterization* [31] [32] of a LF.

2.2.1 Two-Plane Parameterization of a LF

The light rays of a LF can be parameterized in number of ways [2] [3] [80](ch. 2.1). The most widely employed LF parameterization is the standard two-plane parameterization, where each light ray is parameterized by its intersections with two parallel planes: the camera plane and the image plane. Note that, in a more general two-plane parameterization, the camera and image planes are not necessary to be parallel [3] [80](ch. 2.1). Two other possible LF parameterizations are the two-sphere parametrization [81] [82] and the sphere-plane parameterization [81].

Two variants of the two-plane parameterization are shown in Figures 2.1 (a) and 2.1 (b). The main difference between the two variants is the image coordinates (u, v) are defined *globally* in one case and *locally* in the other case, with respect to the camera position (x, y). Note that, in the case of a camera array, the camera plane xy corresponds to the plane containing the principal planes[2] of the camera lenses, and the image plane uv corresponds to the coplanar focal planes of the cameras. Under the two-plane parameterization with the locally defined image coordinates, all rays with identical (u, v) values are parallel whereas all rays with $u = v = 0$ are perpendicular to the two planes, which results more compact expressions for the LF representation [61]. Consequently, the two-plane parameterization with the locally defined image coordinates, henceforth referred to as simply the two-plane parameterization for brevity, is employed throughout in this dissertation unless otherwise specified.

[2]The principal planes of an optical system are the two hypothetical planes at which the lateral magnification is equal to unity [83](ch. 4.3). These planes can be obtained by the locus of intersection of incident light rays parallel to the axis and by the light rays directed to the focus [84](ch. 7.1). For a thin lens system in air, the principal planes are coincident and pass through the optical center of the thin lens [84](ch. 7.1).

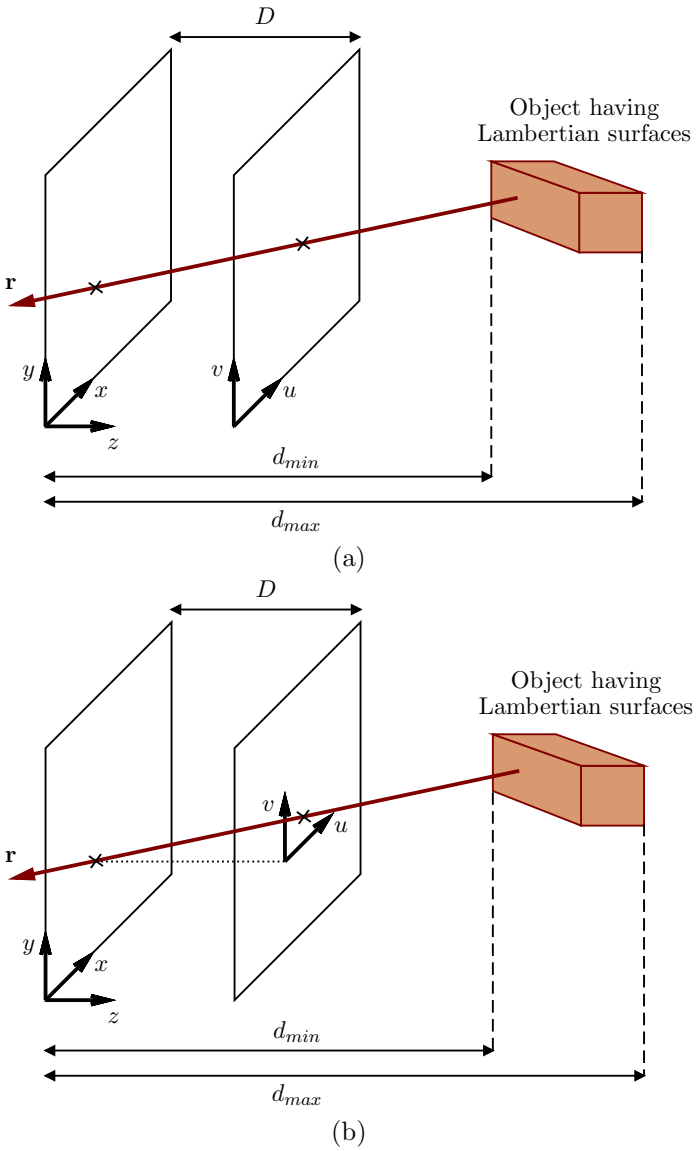


Figure 2.1: The two-plane parameterization of a Lambertian scene comprised of an object having Lambertian surfaces; (a) with globally defined image coordinates (u, v); (b) with locally defined image coordinates (u, v).

The pinhole camera model [84](ch. 6.2) is employed in the two-plane parameterization, and each ray $\mathbf{r}$ is represented by the 4-tuple $(x, y, u, v) \in \mathbb{R}^4$, where the coordinates (x, y) and (u, v) represent the *position* and *direction* of the ray, respectively. In the following review and in the analysis presented in Section 2.3, for simplicity, the scene is assumed to have no occlusion. Furthermore, windowing effects due to the finite number of cameras and due to the limited field-of-view of each camera are ignored, i.e. both the camera plane xy and the image plane uv are assumed to be of

infinite extent.

2.2.2 LF Representation of a Lambertian Point Source and Its Spectrum

Consider the two-plane parametrization of a Lambertian point source of intensity (or radiance) l_0 shown in Figure 2.2. In this case, the Lambertian point source is represented as a *plane* of constant value l_0 in the corresponding 4-D continuous-domain LF $l_{4\mathrm{C}}(x, y, u, v)$ [44]. The plane is given by the intersection of the two 3-D hyperplanes[3]

$$mx + u + c_x = 0 \tag{2.1a}$$

$$my + v + c_y = 0, \tag{2.1b}$$

having normal vectors $\mathbf{c}_{4\mathrm{C},xu} = [m, 0, 1, 0]^{\mathrm{T}}$ and $\mathbf{c}_{4\mathrm{C},yv} = [0, m, 0, 1]^{\mathrm{T}}$, respectively, where

$$m = \frac{D}{z_0} \tag{2.2a}$$

$$c_x = \frac{-Dx_0}{z_0} \tag{2.2b}$$

$$c_y = \frac{-Dy_0}{z_0}, \tag{2.2c}$$

where $(x_0, y_0, z_0) \in \mathbb{R}^2 \times \mathbb{R}^+$ is the position of the Lambertian point source and D is the distance between the camera plane xy and the image plane uv [44]. The LF $l_{4\mathrm{C}}(x, y, u, v)$ may be expressed as

$$l_{4\mathrm{C}}(x, y, u, v) = l_0\, \delta(mx + u + c_x)\, \delta(my + v + c_y), \tag{2.3}$$

where $\delta(\cdot)$ is the 1-D continuous-domain impulse function [86](ch. 6.2) [87](ch. 2.1).

The spectrum of $l_{4\mathrm{C}}(x, y, u, v)$, $L_{4\mathrm{C}}(\Omega_x, \Omega_y, \Omega_u, \Omega_v)$ can be obtained as [44] [63]

$$\begin{aligned} L_{4\mathrm{C}}(\Omega_x, \Omega_y, \Omega_u, \Omega_v) = 4\pi^2 l_0\, \delta(\Omega_x - m\Omega_u)\, \delta(\Omega_y - m\Omega_v) \\ \times \mathrm{e}^{j(\Omega_u c_x + \Omega_v c_y)}, \end{aligned} \tag{2.4}$$

[3] An n-D hyperplane, where $n = 3, 4$, means an n-D manifold in the $(n+1)$-D space that is uniquely determined by the $(n+1)$-D normal vector and an $(n+1)$-D point on the n-D hyperplane [85](ch. 66).

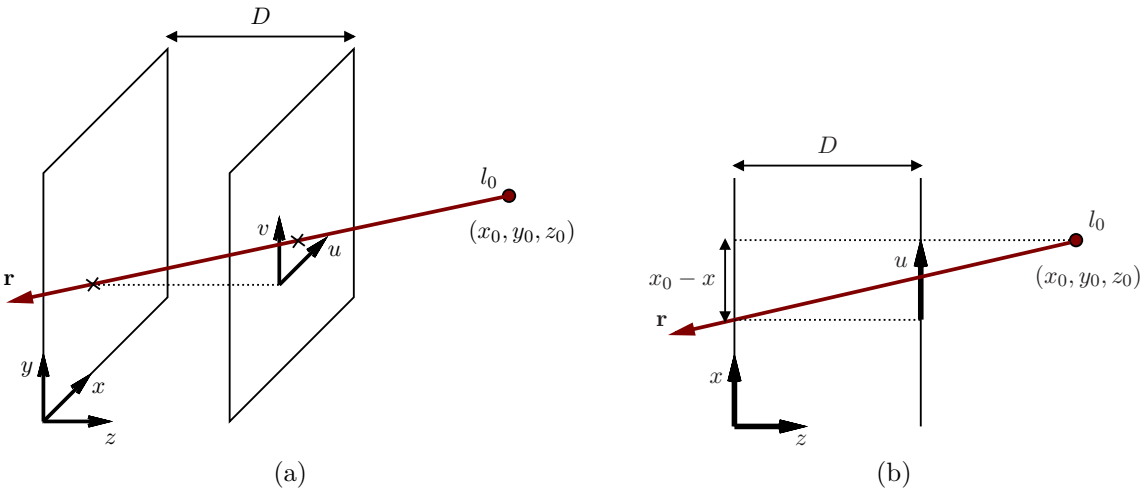


Figure 2.2: (a) Two-plane parametrization of a Lambertian point source of intensity l_0; (b) the representation of the Lambertian point source in the xu subspace.

where $(\Omega_x, \Omega_y, \Omega_u, \Omega_v) \in \mathbb{R}^4$.

2.2.3 ROS of the Spectrum

The ROS $\mathcal{P}_{4C}$ of the spectrum $L_{4C}(\Omega_x, \Omega_y, \Omega_u, \Omega_v)$ can be obtained from (2.4) as [44] [63]

$$\mathcal{P}_{4C} = \mathcal{H}_{4C,xu} \cap \mathcal{H}_{4C,yv}, \tag{2.5}$$

where

$$\mathcal{H}_{4C,xu} = \{(\Omega_x, \Omega_y, \Omega_u, \Omega_v) \in \mathbb{R}^4 \mid \Omega_x - m\Omega_u = 0\} \tag{2.6a}$$
$$\mathcal{H}_{4C,yv} = \{(\Omega_x, \Omega_y, \Omega_u, \Omega_v) \in \mathbb{R}^4 \mid \Omega_y - m\Omega_v = 0\}. \tag{2.6b}$$

The ROS $\mathcal{P}_{4C}$, illustrated in Figure 2.3, is a plane through the origin in the 4-D continuous frequency domain, which is given by the intersection of the two 3-D hyperplanes

$$\Omega_x - m\Omega_u = 0 \tag{2.7a}$$
$$\Omega_y - m\Omega_v = 0 \tag{2.7b}$$

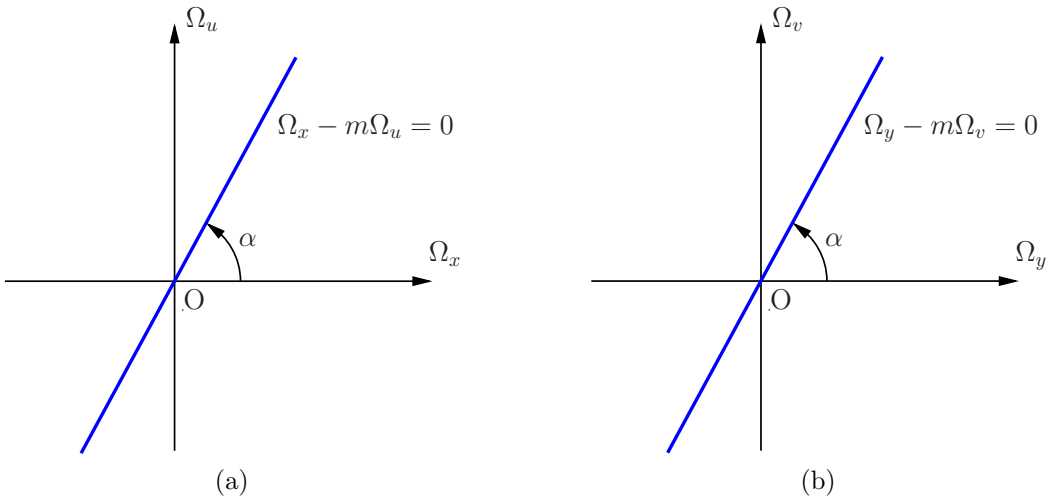


Figure 2.3: The ROS of the spectrum of a Lambertian point source, $\mathcal{P}_{4C}$, (a) $\mathcal{H}_{4C,xu}$ in the $\Omega_x\Omega_u$ subspace; (b) $\mathcal{H}_{4C,yv}$ in the $\Omega_y\Omega_v$ subspace; as z_0 varies in the range $(0, \infty)$, α varies in the range $(0°, 90°)$.

having normal vectors $\mathbf{d}_{4C,xu} = [1, 0, -m, 0]^T$ and $\mathbf{d}_{4C,yv} = [0, 1, 0, -m]^T$, respectively. Note that, the ROS $\mathcal{P}_{4C}$ depends only on the *depth* z_0 of the Lambertian point source [44] [63].

In the case of a Lambertian object, where the depth varies in a range, i.e. $z_0 \in [d_{min}, d_{max}]$ (see Figure 2.1), the ROS of the spectrum, $\mathcal{O}_{4C}$, is obtained as

$$\begin{aligned} \mathcal{O}_{4C} &= \bigcup_{z_0} \mathcal{P}_{4C} \\ &= \bigcup_{z_0} \left(\mathcal{H}_{4C,xu} \cap \mathcal{H}_{4C,yv}\right), \end{aligned} \tag{2.8}$$

which is a *3-D hyperfan* in the 4-D continuous frequency domain [9](ch. 4) [47]. The angle of the 3-D hyperfan depends on the depth range occupied by the Lambertian object. Figure 2.4 illustrates the ROS $\mathcal{O}_{4C}$.

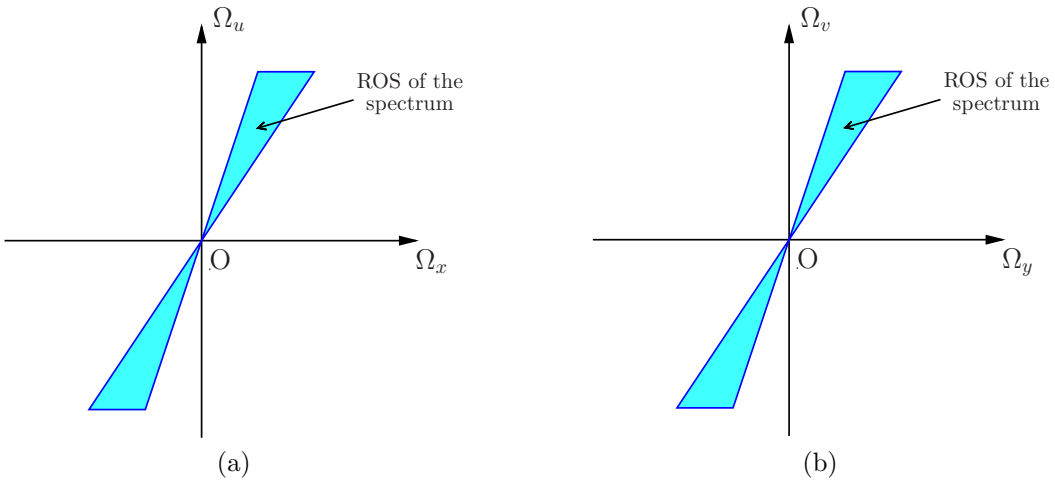


Figure 2.4: The ROS of the spectrum of a Lambertian object, $\mathcal{O}_{4\mathrm{C}}$, (a) in the $\Omega_x\Omega_u$ subspace; (b) in the $\Omega_y\Omega_v$ subspace. The angle of the 3-D hyperfan depends on the depth range $z_0 \in [d_{min}, d_{max}]$ occupied by the Lambertian object.

2.3 Analysis of the Spectrum of a Lambertian Object Moving with Constant Velocity

2.3.1 LFV Representation of a Lambertian Point Source

Consider the case shown in Figure 2.5, where a Lambertian point source of intensity l_0 moves with the *constant* velocity $\mathbf{V} = [V_x, V_y, V_z]^{\mathrm{T}}$. Similar to the 4-D continuous-domain LF representation of a Lambertian point source, as discussed in the previous section, the 5-D continuous-domain LFV $l_{5\mathrm{C}}(\mathbf{x})$, $\mathbf{x} = [x, y, u, v, t]^{\mathrm{T}} \in \mathbb{R}^5$, may be expressed as

$$l_{5\mathrm{C}}(\mathbf{x}) = l_0\, \delta(m(t)x + u + c_x(t))\, \delta(m(t)y + v + c_y(t)), \tag{2.9}$$

where

$$m(t) = \frac{D}{z_p(t)} \tag{2.10a}$$

$$c_x(t) = \frac{-Dx_p(t)}{z_p(t)} \tag{2.10b}$$

$$c_y(t) = \frac{-Dy_p(t)}{z_p(t)}. \tag{2.10c}$$

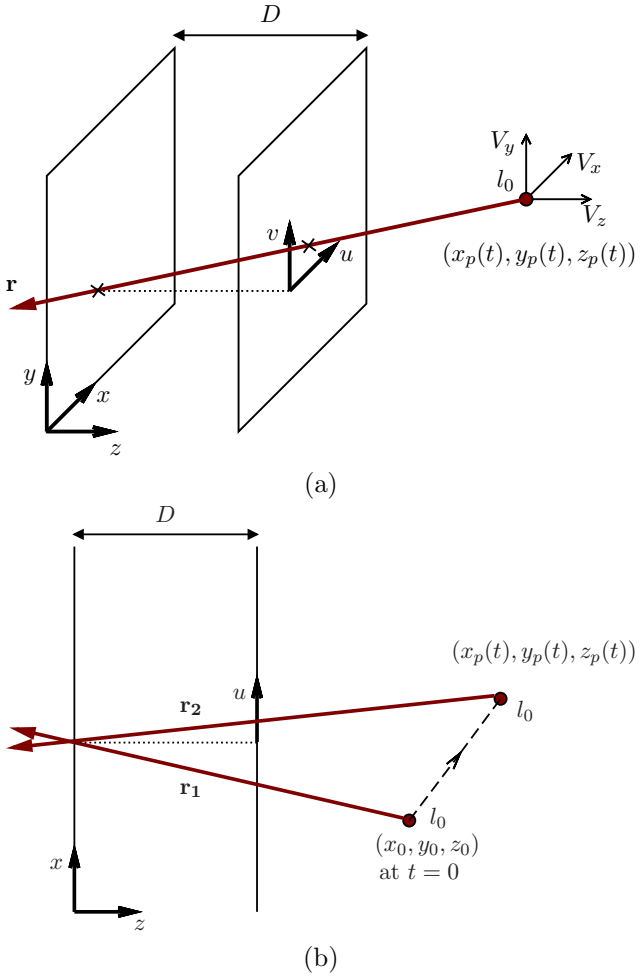


Figure 2.5: (a) A Lambertian point source of intensity l_0 moves with the constant velocity $\mathbf{V} = [V_x, V_y, V_z]^{\mathrm{T}}$; (b) the representation of the Lambertian point source in the xu subspace.

In this case, similar to the *constant intensity assumption* [88](ch. 2.3) employed in analysis of moving objects in conventional 3-D videos, we assume that the intensity l_0 of the Lambertian point source does not change with time, i.e. the scene is under *homogeneous ambient illumination*. Because

$$x_p(t) = V_x t + x_0 \tag{2.11a}$$

$$y_p(t) = V_y t + y_0 \tag{2.11b}$$

$$z_p(t) = V_z t + z_0, \tag{2.11c}$$

where $(x_0, y_0, z_0) \in \mathbb{R}^2 \times \mathbb{R}^+$ is the position of the Lambertian point source at $t = 0$, after some manipulation, (2.9) can be rewritten as

$$\begin{aligned} l_{5\mathrm{C}}(\mathbf{x}) = l_0 \, &\delta(mx + u + k_x t - k_z u t + c_x) \\ &\times \delta(my + v + k_y t - k_z v t + c_y), \end{aligned} \tag{2.12}$$

where m, c_x and c_y are given by (2.2a), (2.2b) and (2.2c), respectively, and

$$k_x = \frac{-DV_x}{z_0} \tag{2.13a}$$

$$k_y = \frac{-DV_y}{z_0} \tag{2.13b}$$

$$k_z = \frac{-V_z}{z_0}. \tag{2.13c}$$

According to (2.12), the Lambertian point source is represented in the LFV $l_{5\mathrm{C}}(\mathbf{x})$ as a *3-D hypersurface of constant value* l_0, which is given by the intersection of the two 4-D hypersurfaces

$$mx + u + k_x t - k_z u t + c_x = 0 \tag{2.14a}$$

$$my + v + k_y t - k_z v t + c_y = 0. \tag{2.14b}$$

In the case, where the Lambertian point source moves at a *constant depth* z_0, i.e. $V_z = 0$, (2.12) reduces to

$$l_{5\mathrm{C}}(\mathbf{x}) = l_0 \, \delta(mx + u + k_x t + c_x) \, \delta(my + v + k_y t + c_y). \tag{2.15}$$

In this case, the Lambertian point source is represented in the LFV $l_{5\mathrm{C}}(\mathbf{x})$ as a *3-D hyperplane of constant value* l_0, which is given by the intersection of the two 4-D hyperplanes

$$mx + u + k_x t + c_x = 0 \tag{2.16a}$$

$$my + v + k_y t + c_y = 0 \tag{2.16b}$$

having normal vectors $\mathbf{c}_{5\mathrm{C},xut} = [m, 0, 1, 0, k_x]^{\mathrm{T}}$ and $\mathbf{c}_{5\mathrm{C},yvt} = [0, m, 0, 1, k_y]^{\mathrm{T}}$, respectively.

2.3.2 Spectrum of a Lambertian Point Source Moving at a Constant Depth

In this subsection, the spectrum of $l_{5\mathrm{C}}(\mathbf{x})$ is derived in closed form for the case $V_z = 0$. The spectrum $L_{5\mathrm{C}}(\mathbf{\Omega})$, $\mathbf{\Omega} = [\Omega_x, \Omega_y, \Omega_u, \Omega_v, \Omega_t]^{\mathrm{T}} \in \mathbb{R}^5$, can be obtained as

$$\begin{aligned} L_{5\mathrm{C}}(\mathbf{\Omega}) = 8\pi^3 l_0 \, \delta(\Omega_x - m\Omega_u) \, \delta(\Omega_y - m\Omega_v) \\ \times \, \delta(\Omega_t - k_x\Omega_u - k_y\Omega_v) \, \mathrm{e}^{j(\Omega_u c_x + \Omega_v c_y)} \end{aligned} \tag{2.17}$$

as derived in Appendix A. Note that, similar to the spectrum of an object moving with constant 2-D spatial velocity in a conventional 3-D video [88] (ch. 2.3), (2.17) can be expressed as

$$L_{5\mathrm{C}}(\mathbf{\Omega}) = 2\pi \, L_{4\mathrm{C}}(\Omega_x, \Omega_y, \Omega_u, \Omega_v) \, \delta(\Omega_t - k_x\Omega_u - k_y\Omega_v), \tag{2.18}$$

where $L_{4\mathrm{C}}(\Omega_x, \Omega_y, \Omega_u, \Omega_v)$ is the spectrum of the *time-invariant* LF of the Lambertian point source as given in (2.4).

From (2.17), the ROS of the spectrum, $\mathcal{P}_{5\mathrm{C}}$, can be obtained as

$$\mathcal{P}_{5\mathrm{C}} = \mathcal{H}_{5\mathrm{C},xu} \cap \mathcal{H}_{5\mathrm{C},yv} \cap \mathcal{H}_{5\mathrm{C},uvt}, \tag{2.19}$$

where

$$\mathcal{H}_{5\mathrm{C},xu} = \{\mathbf{\Omega} \in \mathbb{R}^5 \mid \Omega_x - m\Omega_u = 0\} \tag{2.20a}$$

$$\mathcal{H}_{5\mathrm{C},yv} = \{\mathbf{\Omega} \in \mathbb{R}^5 \mid \Omega_y - m\Omega_v = 0\} \tag{2.20b}$$

$$\mathcal{H}_{5\mathrm{C},uvt} = \{\mathbf{\Omega} \in \mathbb{R}^5 \mid \Omega_t - k_x\Omega_u - k_y\Omega_v = 0\}, \tag{2.20c}$$

which are illustrated in Figures 2.6(a)–2.6(c), respectively. The ROS $\mathcal{P}_{5\mathrm{C}}$ is a *plane through the origin* in the 5-D continuous frequency domain, which is given by the intersection of the three 4-D hyperplanes

$$\Omega_x - m\Omega_u = 0 \tag{2.21a}$$

$$\Omega_y - m\Omega_v = 0 \tag{2.21b}$$

$$\Omega_t - k_x\Omega_u - k_y\Omega_v = 0 \tag{2.21c}$$

having normal vectors $\mathbf{d}_{5\mathrm{C},xu} = [1, 0, -m, 0, 0]^{\mathrm{T}}$, $\mathbf{d}_{5\mathrm{C},yv} = [0, 1, 0, -m, 0]^{\mathrm{T}}$ and $\mathbf{d}_{5\mathrm{C},uvt} =$

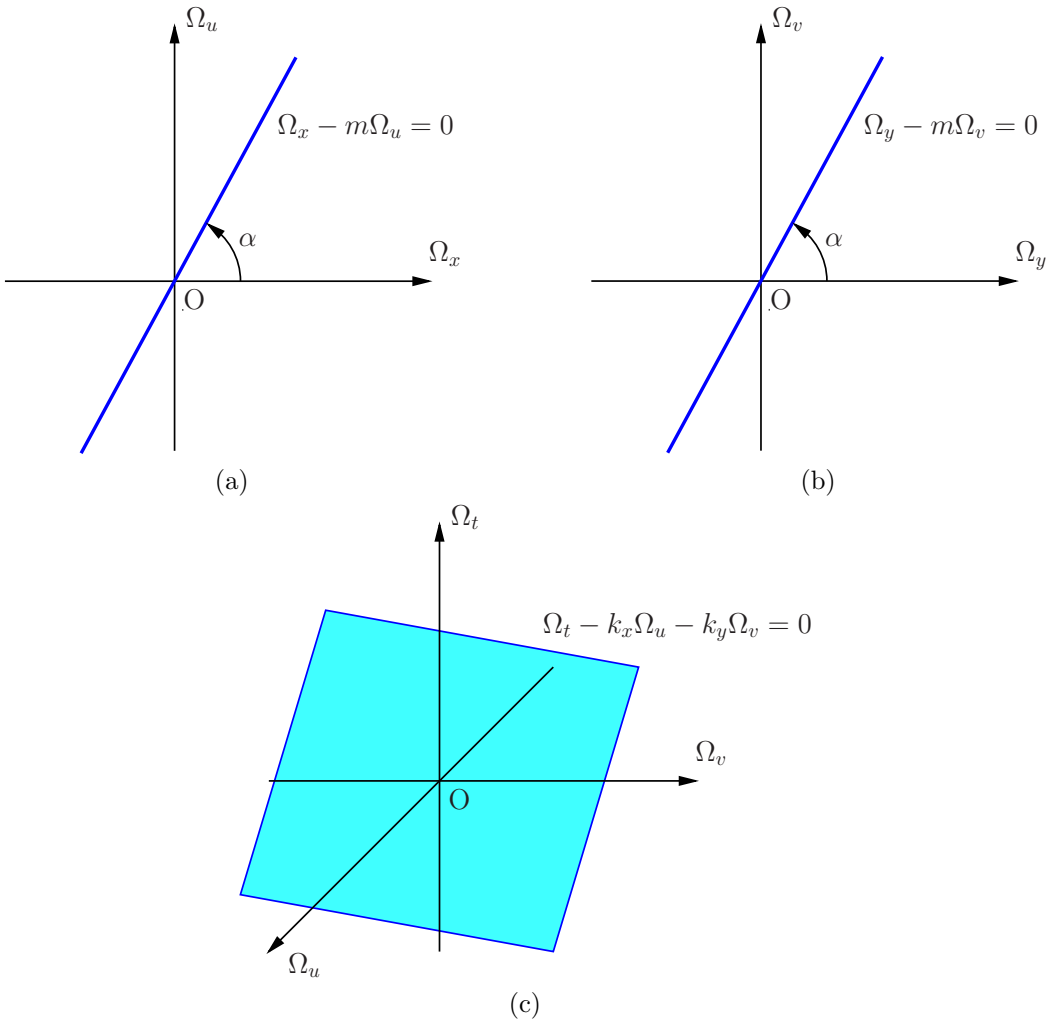


Figure 2.6: The ROS of the spectrum, $\mathcal{P}_{5\mathrm{C}}$, (a) $\mathcal{H}_{5\mathrm{C},xu}$ in the $\Omega_x\Omega_u$ subspace; (b) $\mathcal{H}_{5\mathrm{C},yv}$ in the $\Omega_y\Omega_v$ subspace; as z_0 varies in the range $(0, \infty)$, α varies in the range $(0°, 90°)$. (c) $\mathcal{H}_{5\mathrm{C},uvt}$ in the $\Omega_u\Omega_v\Omega_t$ subspace.

$[0, 0, -k_x, -k_y, 1]^{\mathrm{T}}$, respectively. Note that the ROS $\mathcal{P}_{5\mathrm{C}}$ depends on both the *constant velocity vector* $\mathbf{V} = [V_x, V_y, 0]^{\mathrm{T}}$ and the *depth* z_0 of the Lambertian point source, implying that Lambertian point sources moving with different constant velocities or at different depths ideally have *non-overlapping* ROSs except at the origin. *This fact is the basis for enhancing moving objects in LFVs by employing 5-D depth-velocity filters described in this dissertation.*

2.3.3 ROS of the Spectrum of a Lambertian Object Moving at a Constant Depth

In the context of filter design, the ROS of the spectrum is more important than the spectrum itself since the most of the design specifications, such as cutoff frequencies, of a filter are determined based on the ROS of the spectra of signals to be filtered. In this subsection, we derive the ROS of the spectrum of a Lambertian object moving with $\mathbf{V} = [V_x, V_y, 0]^{\mathrm{T}}$.

The Lambertian object may be considered as a collection of Lambertian point sources having depth $z_0 \in [d_{min}, d_{max}]$ (see Figure 2.1). Therefore, the ROS of the spectrum, $\mathcal{O}_{5\mathrm{C}}$, can be obtained as the *union* of the ROSs of the spectra of the corresponding Lambertian point sources because of the linearity of the Fourier transform [73](ch. 1.3) [74](ch. 1.2), i.e.

$$\begin{aligned}\mathcal{O}_{5\mathrm{C}} &= \bigcup_{z_0} \mathcal{P}_{5\mathrm{C}} \\ &= \bigcup_{z_0} \left(\mathcal{H}_{5\mathrm{C},xu} \cap \mathcal{H}_{5\mathrm{C},yv} \cap \mathcal{H}_{5\mathrm{C},uvt}\right). \end{aligned} \tag{2.22}$$

The ROS $\mathcal{O}_{5\mathrm{C}}$, illustrated in Figure 2.7, is a *skewed 3-D hyperfan* in the 5-D continuous frequency domain. Here, 3-D hyperfan means a 3-D manifold that can be formed by sweeping a plane through the 4-D or 5-D space [9](ch. 4) [47]. *The degree of skewness of the 3-D hyperfan depends on both velocity and depth of the moving object whereas the angle of the 3-D hyperfan depends on the depth range occupied by the moving object.* Note that, similar to 4D LFs, we observe a *dimensionality gap* in the ROS of the spectrum; that is the 5-D LFV is reduced to a 3-D manifold in the 5-D continuous frequency domain. In this case, one dimension is reduced due to the fact that the Lambertian scene is in fact in the 3-D space even though it is represented as 4-D in the corresponding LF [64]. In other words, the parameter m in (2.16a) and (2.16b) is the same for both 4-D hyperplanes for a point in the 3-D space. The other dimension is reduced as a consequence of the constant intensity assumption.

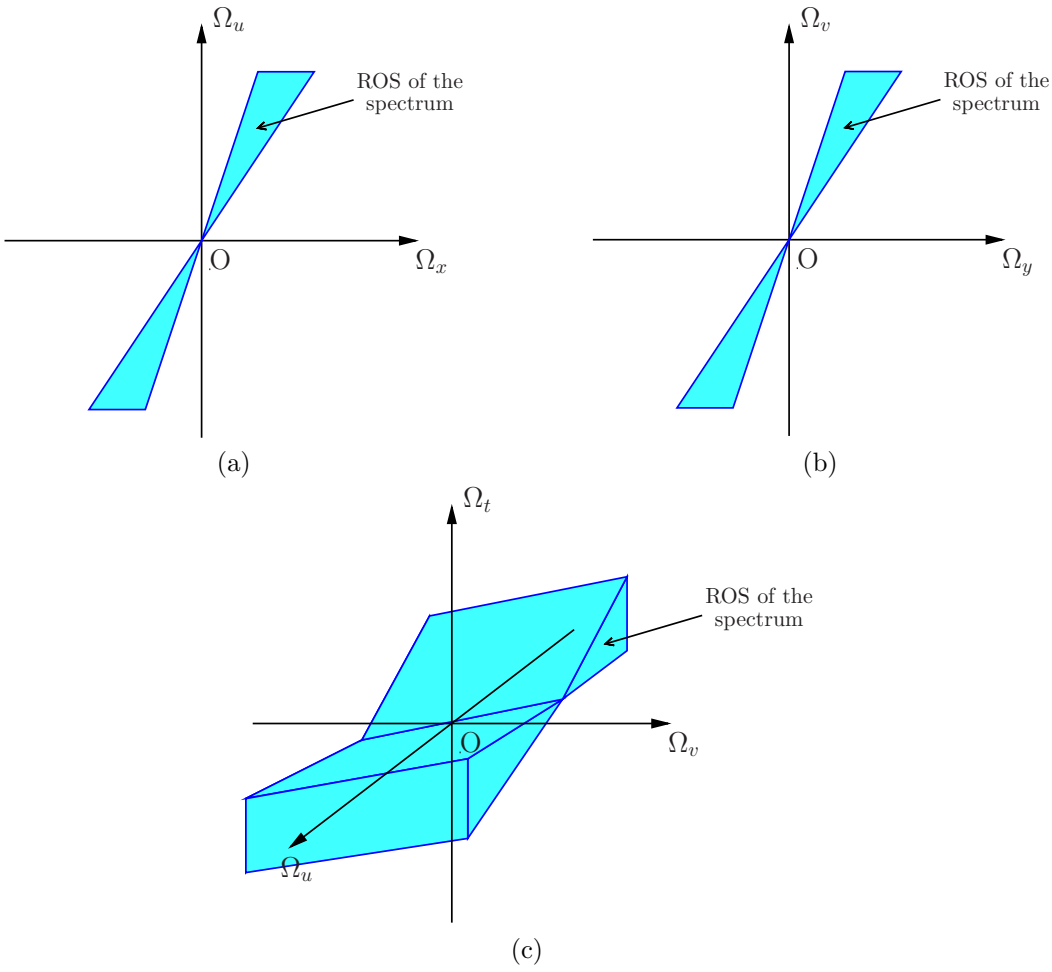


Figure 2.7: The ROS of the spectrum, $\mathcal{O}_{5\mathrm{C}}$, (a) in the $\Omega_x\Omega_u$ subspace; (b) in the $\Omega_y\Omega_v$ subspace; (c) in the $\Omega_u\Omega_v\Omega_t$ subspace.

2.3.4 ROS of the Spectrum Corresponding to a Discrete-Domain LFV

An LFV obtained from an LFV camera is in fact a discrete-domain (sampled) signal. Consequently, processing of LFVs is carried out in the discrete-domain. In this subsection, we derive the ROS of the spectrum corresponding to a *discrete-domain* LFV. To this end, we assume that the discrete-domain LFV $l_{5\mathrm{D}}(\mathbf{n})$, $\mathbf{n} = [n_x, n_y, n_u, n_v, n_t]^{\mathrm{T}} \in \mathbb{Z}^5$, is generated by rectangularly sampling the corresponding continuous-domain LFV $l_{5\mathrm{C}}(\mathbf{x})$, with the sampling matrix

$$\mathbf{\Delta} = \mathrm{diag}[\Delta x, \Delta y, \Delta u, \Delta v, \Delta t], \tag{2.23}$$

where the term "diag" denotes a diagonal matrix and $\Delta i,\ i = x, y, u, v, t,$ is the sampling interval along the corresponding dimension. Furthermore, the *5-D principal Nyquist hypercube* in the 5-D discrete frequency domain $\boldsymbol{\omega}$ $(= [\omega_x, \omega_y, \omega_u, \omega_v, \omega_t]^{\mathrm{T}} \in \mathbb{R}^5$, where $\omega_i = \Omega_i \Delta i,\ i = x, y, u, v, t)$ is defined as

$$\mathcal{N}_{5\mathrm{D}} \triangleq \left\{ \boldsymbol{\omega} \in \mathbb{R}^5 \mid -\pi \leq \omega_i \leq \pi,\ i = x, y, u, v, t \right\}. \tag{2.24}$$

The discrete-domain spectrum $L_{5\mathrm{D}}(\boldsymbol{\omega})$ is a periodic extension of the corresponding continuous-domain spectrum $L_{5\mathrm{C}}(\boldsymbol{\Omega})$ and is given by [73](ch. 1.4) [74](ch. 2.2)

$$L_{5\mathrm{D}}(\boldsymbol{\omega}) = \frac{1}{|\mathrm{det}\boldsymbol{\Delta}|} \sum_{\mathbf{k}} L_{5\mathrm{C}} \left(\boldsymbol{\Omega} - 2\pi \boldsymbol{\Delta}^{-1} \mathbf{k} \right), \tag{2.25}$$

where $\mathbf{k} \in \mathbb{Z}^5$ and the term "det" denotes the determinant of a matrix. Even though the spectrum of an LF is, in general, band unlimited, most of the energy of the spectrum resides within a finite 4-D hypervolume denoted as the *essential bandwidth* [66] [67]. Therefore, most LFs can be sampled with negligible aliasing. It is clear from (2.17) and especially from (2.18) that the maximum temporal frequency of $L_{5\mathrm{C}}(\boldsymbol{\Omega})$ depends only on the maximum values of k_x and k_y, and the essential bandwidth of $L_{4\mathrm{C}}(\Omega_x, \Omega_y, \Omega_u, \Omega_v)$ on the Ω_u and Ω_v dimensions. Consequently, with finite k_x and k_y, we can conclude that the essential bandwidth of $L_{5\mathrm{C}}(\boldsymbol{\Omega})$ on the Ω_t dimension is *finite*. Therefore, the continuous-domain LFV $l_{5\mathrm{C}}(\mathbf{x})$ can be sampled with negligible aliasing by employing a sufficiently high temporal sampling rate and, for this case, the ROS of the spectrum, $\mathcal{O}_{5\mathrm{D}}$, inside $\mathcal{N}_{5\mathrm{D}}$ is obtained as

$$\mathcal{O}_{5\mathrm{D}} = \bigcup_{z_0} \left(\mathcal{H}_{5\mathrm{D},xu} \cap \mathcal{H}_{5\mathrm{D},yv} \cap \mathcal{H}_{5\mathrm{D},uvt} \right), \tag{2.26}$$

where

$$\mathcal{H}_{5\mathrm{D},xu} = \left\{ \boldsymbol{\omega} \in \mathcal{N}_{5\mathrm{D}} \,\middle|\, \omega_x - \left(\frac{m\Delta x}{\Delta u} \right) \omega_u = 0 \right\} \tag{2.27a}$$

$$\mathcal{H}_{5\mathrm{D},yv} = \left\{ \boldsymbol{\omega} \in \mathcal{N}_{5\mathrm{D}} \,\middle|\, \omega_y - \left(\frac{m\Delta y}{\Delta v} \right) \omega_v = 0 \right\} \tag{2.27b}$$

$$\mathcal{H}_{5\mathrm{D},uvt} = \left\{ \boldsymbol{\omega} \in \mathcal{N}_{5\mathrm{D}} \,\middle|\, \omega_t - \left(\frac{k_x \Delta t}{\Delta u} \right) \omega_u - \left(\frac{k_y \Delta t}{\Delta v} \right) \omega_v = 0 \right\}. \tag{2.27c}$$

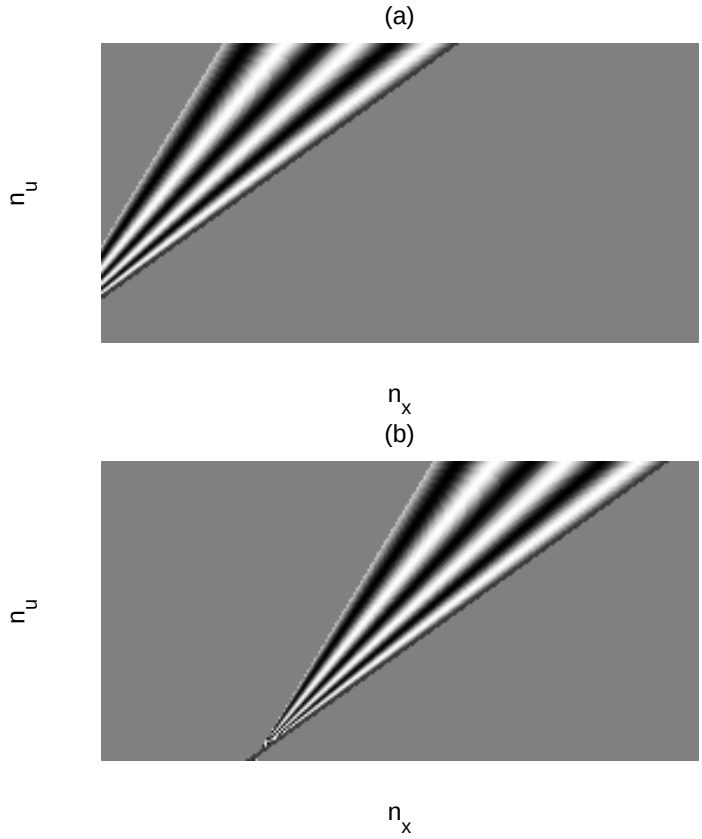


Figure 2.8: The epipolar-plane images of the generated LFV (a) 10th frame; (b) 40th frame.

2.3.5 Numerical Simulation of the ROS of the Spectrum

In this subsection, we present the numerically simulated ROS of the spectrum of a discrete-domain LFV that corresponds to a Lambertian object moving with a constant velocity and at a constant depth. Note that, for illustration purposes, the LFV is restricted to the x, u and t dimensions. The time-invariant intensity pattern of the Lambertian object is selected as sinusoidal in space and the velocity V_x is selected as 3 pixels/frame. Furthermore, the distance D between the camera plane xy and the image plane uv is selected as 50 cm, and the Lambertian object occupies the depth range $[30, 70]$ cm. The LFV of size $256 \times 128 \times 64$ is numerically generated. The epipolar-plane images [89] corresponding to the 10th and 40th frames of the LFV are shown in Figures 2.8 (a) and 2.8 (b), respectively.

In the present case, where only the x, u and t dimensions are incorporated, the ROS of the spectrum, $\mathcal{O}_{3\mathrm{D}}$, inside $\mathcal{N}_{3\mathrm{D}}$ can be obtained, by following a procedure

similar to that employed in the previous subsection, as

$$\mathcal{O}_{\text{3D}} = \bigcup_{z_0} \left(\mathcal{H}_{\text{3D},xu} \cap \mathcal{H}_{\text{3D},ut} \right), \tag{2.28}$$

where

$$\mathcal{H}_{\text{3D},xu} = \left\{ \boldsymbol{\omega} \in \mathcal{N}_{\text{3D}} \;\middle|\; \omega_x - \left(\frac{m\Delta x}{\Delta u} \right) \omega_u = 0 \right\} \tag{2.29a}$$

$$\mathcal{H}_{\text{3D},ut} = \left\{ \boldsymbol{\omega} \in \mathcal{N}_{\text{3D}} \;\middle|\; \omega_t - \left(\frac{k_x \Delta t}{\Delta u} \right) \omega_u = 0 \right\}. \tag{2.29b}$$

Note that, in this case, $\boldsymbol{\omega} = [\omega_x, \omega_u, \omega_t]^{\text{T}} \in \mathbb{R}^3$, and $\mathcal{N}_{\text{3D}}$ is the principal Nyquist cube in the 3-D discrete frequency domain. The ROS of the spectrum corresponding to a single depth z_0 is a straight line through the origin of which the orientation is determined by the depth z_0 and the velocity V_x. For a Lambertian object where depth $z_0 \in [d_{min}, d_{max}]$, the ROS is a collection of straight lines that resembles a skewed fan-shaped surface inside $\mathcal{N}_{\text{3D}}$, with the degree of skewness depending on the velocity and depth. For example, for $V_x = 0$, the fan lies on the $\omega_x\omega_u$ plane, and as V_x increases, the angle between the fan and the $\omega_x\omega_u$ plane increases.

The ROS of the spectrum of the numerically generated LFV is shown in Figure 2.9. It is observed that the ROS of the spectrum is approximately a fan-shaped surface inside $\mathcal{N}_{\text{3D}}$. Consequently, the numerically obtained ROS of the spectrum is consistent with the theoretically predicted ROS of the spectrum. The slight deviations of the numerically obtained ROS of the spectrum from an ideal fan-shaped surface are mainly due to windowing effects caused by the finite number of samples available for x and u dimensions.

2.4 Summary

A LF can be parameterized in number of ways. The most widely employed LF parameterization is the standard two-plane parameterization, where a light ray is parameterized by its intersections with two parallel planes: the camera plane xy and the image plane uv. In the two-plane parameterization, each light ray is represented by the 4-tuple (x, y, u, v), where the coordinate (x, y) represents the position of the light ray while the coordinate (u, v) represents the direction of the light ray.

A Lambertian point source is represented as a plane of constant value in the

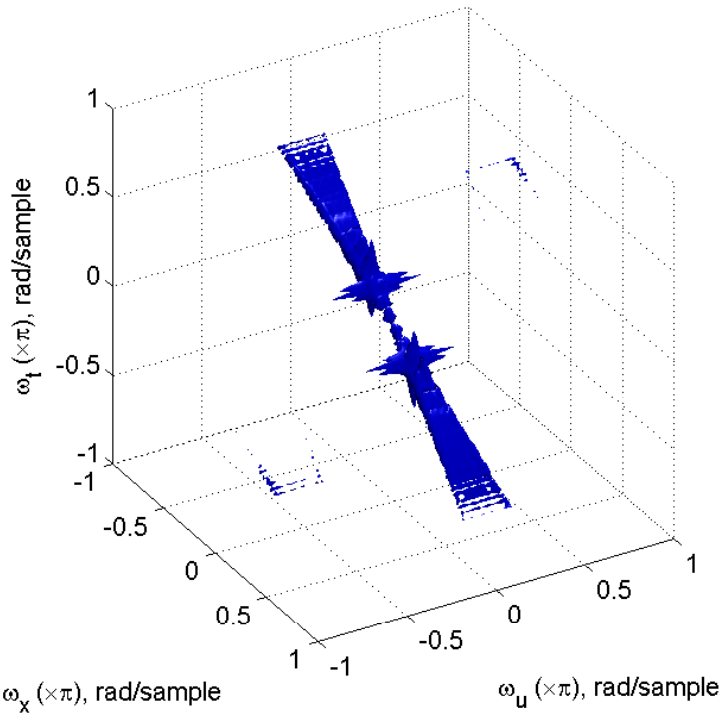


Figure 2.9: The ROS of the spectrum of the numerically generated LFV. The magnitude of the spectrum is normalized, and the iso-surface is drawn at 0.05.

corresponding 4-D continuous-domain LF $l_{4C}(x, y, u, v)$. The ROS of the spectrum $L_{4C}(\Omega_x, \Omega_y, \Omega_u, \Omega_v)$ is a plane through the origin in the 4-D continuous frequency domain. Furthermore, the ROS depends only on the depth of the Lambertian point source. In the case of a Lambertian object, the ROS of the spectrum is a 3-D hyperfan in the 4-D continuous frequency domain. The angle of the 3-D hyperfan depends on the depth range occupied by the Lambertian object.

A Lambertian point source moving with a constant velocity is represented as a 3-D hypersurface of constant value in the corresponding 5-D continuous-domain LFV $l_{5C}(\mathbf{x})$. When the motion of the Lambertian point source is parallel to the camera plane, i.e. the Lambertian point source moves at a constant depth, the 3-D hypersurface is reduced to a 3-D hyperplane. For this case, the spectrum $L_{5C}(\mathbf{\Omega})$ and its ROS are derived in closed form. The ROS of the spectrum is a plane through the origin in the 5-D continuous frequency domain. The ROS of the spectrum of a continuous-domain LFV corresponding to a Lambertian object moving with a constant velocity and at a constant depth is derived in closed form. It is shown that

the ROS is a skewed 3-D hyperfan in the 5-D continuous frequency domain. The degree of skewness of the 3-D hyperfan depends on both velocity and depth of the moving object whereas the angle of the 3-D hyperfan depends on the depth range occupied by the moving object. The 3-D hyperfans corresponding to the ROSs of the spectra of Lambertian objects moving with different constant velocities or at different constant depths do not overlap except at the origin in the 5-D continuous frequency domain. This allows enhancement of such objects based on their depth and velocity by employing 5-D depth-velocity filters.

The essential bandwidth of the spectrum is finite along the temporal frequency dimension Ω_t and, consequently, the corresponding discrete-domain LFV can be generated with negligible aliasing by employing a sufficiently high temporal sampling rate. The ROS of the spectrum of a discrete-domain LFV corresponding to a Lambertian object moving with a constant velocity and at a constant depth is derived in closed form. Furthermore, the ROS of the spectrum is numerical simulated. The numerically simulated ROS of the spectrum well agrees with the theoretically predicted ROS of the spectrum.

Chapter 3

Novel 5-D FIR Depth-Velocity Filter

3.1 Introduction

Multi-dimensional FIR filters are widely employed in image, 3-D conventional video and LF processing applications [51] [73] [74] [88] [90]. Major reasons for their popularity in those applications are: the design of multi-dimensional FIR filters is simple compared to IIR counterparts, and they can be designed to have zero- or linear-phase responses. The other prominent feature of the multi-dimensional FIR filters over IIR counterparts is that they are always stable.

It is clear from the spectral analysis presented in Chapter 2 that the ROSs of the spectra of objects moving with different constant velocities or at different depths are distinct and do not overlap except at the origin. Consequently, by employing a 5-D depth-velocity filter of which the passband encompasses the ROS of the spectrum of the object of interest and the stopband encompasses the ROSs of the spectra of interfering objects, the object of interest can be enhanced while attenuating the interfering objects.

A novel 5-D linear-phase FIR depth-velocity filter having a planar passband in the 5-D frequency domain is proposed in this chapter. The planar passband of the proposed 5-D FIR depth-velocity filter is realized by cascading three 5-D FIR filters having 4-D hyperplanar passbands of appropriate orientations in the 5-D discrete frequency domain. The 5-D FIR filters having 4-D hyperplanar passbands are designed by employing the so-called windowing method [73](ch. 3.3) [74](ch. 5.1). The per-

formance of the proposed 5-D FIR depth-velocity filter in enhancing moving objects in LFVs is tested by employing a numerically generated LFV and an LFV of a real scene, generated by means of a commercially available Lytro LF camera. Experimental results confirm the effectiveness of the proposed 5-D FIR depth-velocity filters over 4-D FIR depth filters and 3-D FIR velocity filters.

The rest of the chapter is organized as follows. The design of the proposed 5-D FIR depth-velocity filter is described in detail in Section 3.2. In Section 3.3, experimental results of enhancing moving objects in LFVs by using the proposed 5-D FIR depth-velocity filter are presented. Furthermore, we discuss the performance limitations of the proposed 5-D FIR depth-velocity filter. In Section 3.4, the effects of non-Lambertian reflections and occlusion on the performance of the 5-D depth-velocity filters are briefly discussed. Finally, a summary of the chapter is presented in Section 3.5.

3.2 Design of the 5-D FIR Depth-Velocity Filter

The structure of the 5-D FIR depth-velocity filter $H(\mathbf{z})$, $\mathbf{z} = [z_x, z_y, z_u, z_v, z_t]^{\mathrm{T}} \in \mathbb{C}^5$, is shown in Figure 3.1. The 5-D FIR depth-velocity filter is designed as a cascade of three 5-D FIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$, and the transfer function $H(\mathbf{z})$ is given by

$$H(\mathbf{z}) = H_{xu}(\mathbf{z})\, H_{yv}(\mathbf{z})\, H_{uvt}(\mathbf{z}). \tag{3.1}$$

The passband of each of the three 5-D FIR filters is selected as 4-D hyperplanes inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5\mathrm{D}}$. In particular, the passbands of $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ encompass the 5-D hypervolumes given by

$$\mathcal{B}_{5\mathrm{D},xu} = \bigcup_{z_0} \mathcal{H}_{5\mathrm{D},xu} \tag{3.2a}$$

$$\mathcal{B}_{5\mathrm{D},yv} = \bigcup_{z_0} \mathcal{H}_{5\mathrm{D},yv} \tag{3.2b}$$

$$\mathcal{B}_{5\mathrm{D},uvt} = \bigcup_{z_0} \mathcal{H}_{5\mathrm{D},uvt} \tag{3.2c}$$

respectively, as illustrated in Figure 3.2. Recall that $\mathcal{H}_{5\mathrm{D},xu}$, $\mathcal{H}_{5\mathrm{D},yv}$ and $\mathcal{H}_{5\mathrm{D},uvt}$ are the 4-D hyperplanes corresponding to the ROS of the spectrum of the object of interest

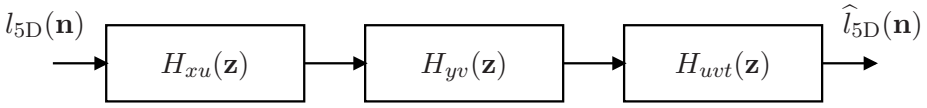


Figure 3.1: Structure of the 5-D FIR depth-velocity filter.

and are given by (2.27a), (2.27b) and (2.27c), respectively. The overall passband $\mathcal{B}_{5\mathrm{D}}$ of $H(\mathbf{z})$, which is a 5-D hypervolume inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5\mathrm{D}}$, is given by the intersection of $\mathcal{B}_{5\mathrm{D},xu}$, $\mathcal{B}_{5\mathrm{D},yv}$ and $\mathcal{B}_{5\mathrm{D},uvt}$, i.e.

$$\mathcal{B}_{5\mathrm{D}} = \left(\bigcup_{z_0} \mathcal{H}_{5\mathrm{D},xu}\right) \bigcap \left(\bigcup_{z_0} \mathcal{H}_{5\mathrm{D},yv}\right) \bigcap \left(\bigcup_{z_0} \mathcal{H}_{5\mathrm{D},uvt}\right). \tag{3.3}$$

Note that the 3-D hyperfan corresponding to the ROS $\mathcal{O}_{5\mathrm{D}}$ of the spectrum of the object of interest, given by (2.26), is a *proper subset* of $\mathcal{B}_{5\mathrm{D}}$. Consequently, the passband of $H(\mathbf{z})$ completely encompasses the ROS of the spectrum of the object of interest. Furthermore, most of the ROSs of the spectra of interfering objects lie outside the passband of $H(\mathbf{z})$.

3.2.1 Design of the 5-D FIR Filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$

The passbands of the 5-D FIR filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ have similar characteristics. In particular, they occupy the similar areas of parallelograms on the $\omega_x\omega_u$ and $\omega_y\omega_v$ planes, respectively. Moreover, the passbands of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ are independent of three dimensions inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5\mathrm{D}}$. Consequently, the design methods of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ are similar, and we present the design method only for $H_{xu}(\mathbf{z})$ in detail.

The frequency-domain specifications of the ROS of $H_{xu}(\mathbf{z})$ are shown in Figure 3.3. The parameter $a_u \in (0, \infty)$ determines the orientation of the 4-D hyperplaner passband whereas b_x and b_u, where $b_x, b_u \in \mathbb{R}^+$, are the bandwidths of the filter along the ω_x and ω_u dimensions, respectively. Note that $b_u = b_x/a_u$. The ideal frequency response of $H_{xu}(\mathbf{z})$, inside $\mathcal{N}_{5\mathrm{D}}$, may be expressed as

$$H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}) = \begin{cases} 1, & a_u\omega_u - b_x \leq \omega_x \leq a_u\omega_u + b_x \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

The ideal infinite-extent impulse response $h^I_{xu}(\mathbf{n})$ of $H_{xu}(\mathbf{z})$ is derived in Appendix B.

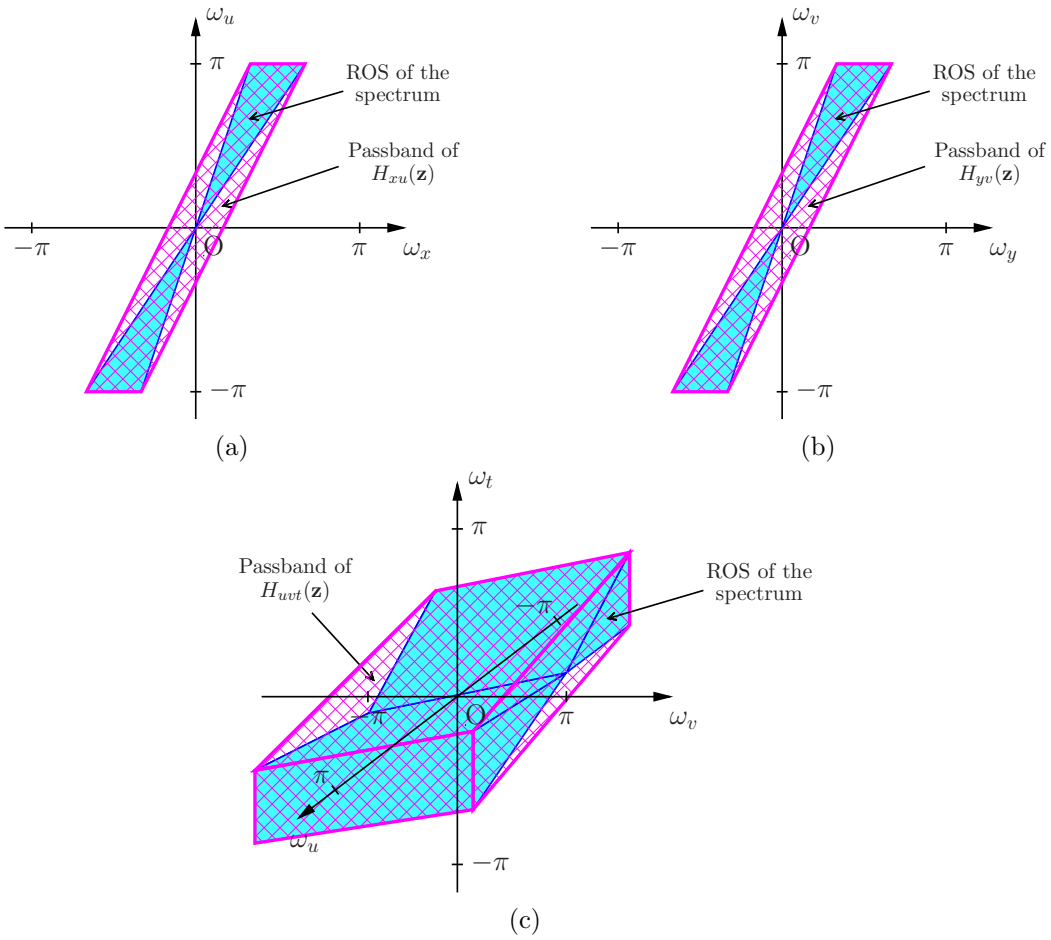


Figure 3.2: The ROS of the spectrum of the object of interest (solid) and the passband (cross-hatched) of (a) $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane; (b) $H_{yv}(\mathbf{z})$ on the $\omega_y\omega_v$ plane; (c) $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space.

Then, the finite-extent impulse response $h_{xu}(\mathbf{n})$ of $H_{xu}(\mathbf{z})$ is obtained as

$$\begin{aligned} h_{xu}(\mathbf{n}) &= h^I_{xu}(\mathbf{n})\, w_{xu}(n_x, n_u) \\ &= \left[g^I_{xu}(n_x, n_u)\, w_{xu}(n_x, n_u)\right] \delta(n_y)\delta(n_v)\delta(n_t), \end{aligned} \tag{3.5}$$

where $w_{xu}(n_x, n_u)$ is a 2-D separable or circular window of size $(M_x + 1) \times (M_u + 1)$ [73](ch. 3.3) [74](ch. 5.1), where $M_x, M_u \in \mathbb{N}$. Note that the order of $H_{xu}(\mathbf{z})$ is $M_x \times 0 \times M_u \times 0 \times 0$.

With the passband considered in deriving $h_{xu}(\mathbf{n})$ (see Figure 3.3), for the range $(0, 1)$ of the parameter a_u, the passband of $H_{xu}(\mathbf{z})$ has undesired ROSs as shown in Figure 3.4. Nevertheless, this drawback can be avoided by employing the symme-

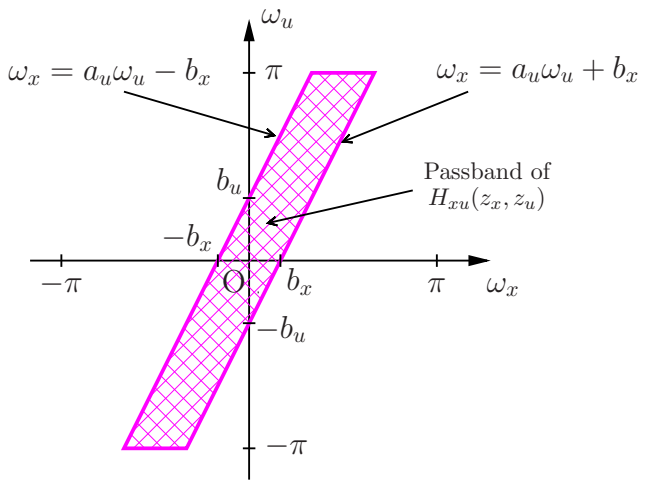


Figure 3.3: Frequency-domain specifications of the ROS of $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane. The parameter a_u determines the orientation of the 4-D hyperplaner passband. b_x and b_u are the bandwidths of the filter along the ω_x and ω_u dimensions, respectively, and $b_u = b_x/a_u$.

try of the passbands corresponding to a_u and $1/a_u$ around the diagonal defined by the line $\omega_x = \omega_u$ inside the 2-D principal Nyquist square $\mathcal{N}_{2\mathrm{D}}$ on the $\omega_x\omega_u$ plane. More specifically, in calculating $g^I_{xu}(n_x, n_u)$ from (B.3a)–(B.3d), a_u and b_x should be replaced with $1/a_u$ and b_t, respectively, and $\left[g^I_{xu}(n_x, n_u)\, w_{xu}(n_x, n_u)\right]^{\mathrm{T}}$ should be used instead of $\left[g^I_{xu}(n_x, n_u)\, w_{xu}(n_x, n_u)\right]$ in (3.5). Furthermore, in this case, the 2-D window should be of size $(M_u+1)\times(M_x+1)$ rather than of size $(M_x+1)\times(M_u+1)$.

The design of $H_{yv}(\mathbf{z})$ can be carried out similar to the design of $H_{xu}(\mathbf{z})$. In particular, the finite-extent impulse response $h_{yv}(\mathbf{n})$ of $H_{yv}(\mathbf{z})$ of order $0\times M_y\times 0\times M_v\times 0$ can obtained as

$$\begin{aligned} h_{yv}(\mathbf{n}) &= h^I_{yv}(\mathbf{n})\, w_{yv}(n_y, n_v) \\ &= \left[g^I_{yv}(n_y, n_v)\, w_{yv}(n_y, n_v)\right] \delta(n_x)\delta(n_u)\delta(n_t), \end{aligned} \tag{3.6}$$

where $g^I_{yv}(n_y, n_v)$ can be obtained from (B.3a)–(B.3d) by replacing n_x and n_u with n_y and n_v, respectively, and a_u and b_x with the counterpart parameters a_v and b_y, respectively. Note that $w_{yv}(n_y, n_v)$ is a 2-D separable or circular window of size $(M_y+1)\times(M_v+1)$ [73](ch. 3.3) [74](ch. 5.1), where $M_y, M_v \in \mathbb{N}$, and $b_v = b_y/a_v$.

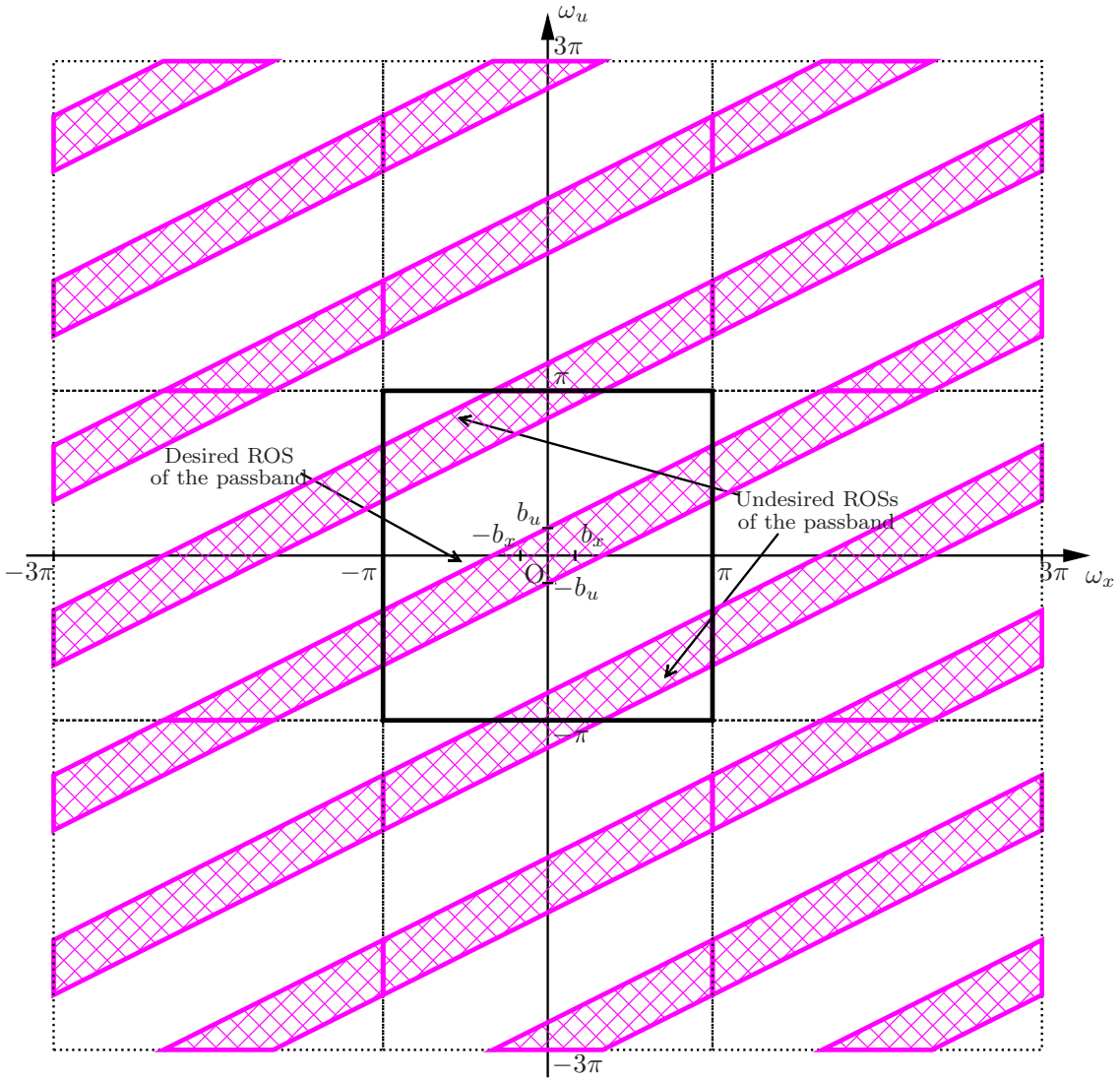


Figure 3.4: The ROS of $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane for $a_u \in (0, 1)$.

3.2.2 Design of the 5-D FIR Filter $H_{uvt}(\mathbf{z})$

The frequency-domain specifications of the ROS of $H_{uvt}(\mathbf{z})$ are shown in Figure 3.5. The parameters $a_u \in (-\infty, \infty)$ and $a_v \in (-\infty, \infty)$ determine the orientation of the 4-D hyperplaner passband whereas b_u, b_v and b_t, where $b_u, b_v, b_t \in \mathbb{R}^+$, are the bandwidths of the filter along the ω_u, ω_v and ω_t dimensions, respectively. Note that $b_u = b_t/|a_u|$ and $b_v = b_t/|a_v|$. The ideal frequency response of $H_{uvt}(\mathbf{z})$, inside $\mathcal{N}_{5\mathrm{D}}$,

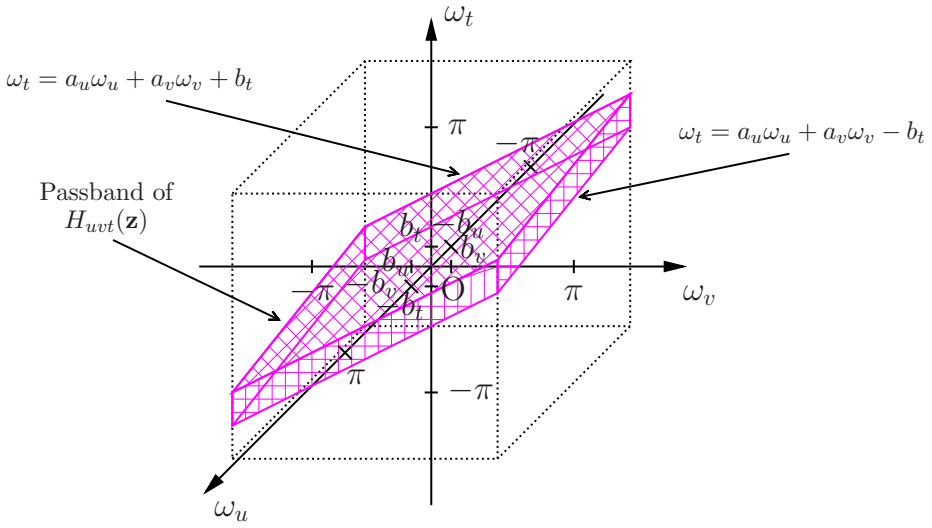


Figure 3.5: Frequency-domain specifications of the ROS of $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space. The parameters a_u and a_v determine the orientation of the 4-D hyperplaner passband. b_u, b_v and b_t are the bandwidths of the filter along the ω_u, ω_v and ω_t dimensions, respectively. Note that $b_u = b_t/|a_u|$ and $b_v = b_t/|a_v|$.

may be expressed as

$$H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}) = \begin{cases} 1, & a_u\omega_u + a_v\omega_v - b_t \le \omega_t \le a_u\omega_u + a_v\omega_v + b_t \\ 0, & \text{otherwise.} \end{cases} \tag{3.7}$$

The ideal infinite-extent impulse response $h^I_{uvt}(\mathbf{n})$ of $H_{uvt}(\mathbf{z})$ is derived in Appendix C. Then, the finite-extent impulse response $h_{uvt}(\mathbf{n})$ of $H_{uvt}(\mathbf{z})$ can be obtained as

$$\begin{aligned} h_{uvt}(\mathbf{n}) &= h^I_{uvt}(\mathbf{n})\, w_{uvt}(n_u, n_v, n_t) \\ &= \left[g^I_{uvt}(n_u, n_v, n_t)\, w_{uvt}(n_u, n_v, n_t)\right] \delta(n_x)\delta(n_y), \end{aligned} \tag{3.8}$$

where $w_{uvt}(n_u, n_v, n_t)$ is a 3-D separable or spherical window of size $(M_u+1)\times(M_v+1)\times(M_t+1)$ [73](ch. 3.3) [74](ch. 5.1), where $M_u, M_v, M_t \in \mathbb{N}$. Note that the order of $H_{uvt}(\mathbf{z})$ is $0\times 0\times M_u\times M_v\times M_t$.

Similar to the passbands of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$, the passband of $H_{uvt}(\mathbf{z})$ has undesired ROSs for some values of a_u and a_v. This undesired property of $H_{uvt}(\mathbf{z})$ is studied using extensive numerical simulations. The passband of $H_{uvt}(\mathbf{z})$ is free from the undesired ROSs, if $|a_u|, |a_v| \in [0, 0.5]$. Furthermore, if $|a_u|, |a_v| \in [2, \infty)$ (or $1/|a_u|, 1/|a_v| \in (0, 0.5]$), $H_{uvt}(\mathbf{z})$ can be designed to have no undesired ROSs

in the passband by employing the symmetry of the passbands inside the 3-D principal Nyquist cube $\mathcal{N}_{3D}$ in the $\omega_u\omega_v\omega_t$ space and appropriate permutations of the dimensions, as employed in the design of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$. Unfortunately, if $|a_u|, |a_v| \in (0.5, 2)$, the passband of $H_{uvt}(\mathbf{z})$ suffers from the undesired ROSs, which may degrade the performance of the 5-D FIR depth-velocity filters in attenuating interfering objects and noise. Nevertheless, thanks to the fact that the undesired ROSs of the passband act like a highpass filter and most of natural scenes are lowpass signals, the effects of the undesired ROSs of the passband may be negligible for most of natural scenes. Furthermore, it is worthwhile to note that the multi-dimensional FIR filters having planar passbands (without undesired ROSs) can be designed by appropriately rotating the passband of a 1-D FIR lowpass filter. This fact is further discussed in the section "Future Work" in Chapter 7.

3.3 Experimental Results

The concept of the 5-D depth-velocity filtering of LFVs to enhance moving objects is experimentally proved in this section. First, experimental results obtained for a numerically generated LFV are presented, and the effectiveness of the proposed 5-D FIR depth-velocity filter in enhancing moving objects is confirmed. Then, the proposed 5-D FIR depth-velocity filter is employed to enhance a moving object in an LFV of a real scene, which is generated using a Lytro LF camera, and it is shown that the proposed 5-D FIR depth-velocity filter is capable of exposing heavily occluded parts of a scene and of attenuating noise significantly.

3.3.1 Numerically Generated LFV

The numerically generated LFV contains three Lambertian planar objects moving with constant velocity and at constant depth. All the three planar objects have a Lambertian intensity pattern that resembles a checkerboard of 8×8 squares. The other specifications of the three planar objects are presented in Table 3.1. The distance D between the camera plane and the image plane is selected as 50 cm. Note that the object of interest and the first interfering object move at the same depth so that the first interfering object can not be attenuated using 4-D depth filters. On the other hand, the object of interest and the second interfering object move with the same apparent velocity so that the second interfering object can not be attenuated using 3-D

Object type	Size height × width (cm^2)	Depth z_0 (cm)	True velocity $[V_x, V_y, V_z]$ (cm/frame)	Apparent velocity $[-k_x, -k_y, -k_z]$ (pixels/frame)
Object of interest	50×50	50	$[1, 1, 0]$	$[1, 1, 0]$
Interfering object 1	50×50	50	$[1, -2, 0]$	$[1, -2, 0]$
Interfering object 2	10×10	10	$[0.2, 0.2, 0]$	$[1, 1, 0]$

Table 3.1: Specifications of the Lambertian planar objects.

velocity filters. The generated LFV is of the size $9 \times 9 \times 256 \times 256 \times 64$, and the 15th frame of the LFV corresponding to the central 25 sub-apertures ($n_x, n_y = 3, 4, \ldots, 7$) is shown in Figure 3.6. Note that all three objects appear with the same apparent size in the LFV because the smaller object (interfering object 2) moves closer to the camera plane.

The filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ are designed with the parameters $a_u = a_v = 1$ and $b_x = b_y = 0.02\pi$ rad/sample and by using a 2-D rectangular window of size 9×41. The orders of the two filters are $8 \times 0 \times 40 \times 0 \times 0$ and $0 \times 8 \times 0 \times 40 \times 0$, respectively. The magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 3.7. The magnitude response of $H_{yv}(\mathbf{z})$ is similar to that of $H_{xu}(\mathbf{z})$. The filter $H_{uvt}(\mathbf{z})$ of order $0 \times 0 \times 40 \times 40 \times 40$ is designed with the parameters $a_u, a_v = -1$ and $b_t = 0.02\pi$ rad/sample and by using a 3-D rectangular window of size $41 \times 41 \times 41$. The magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 3.8. Note that two undesired ROSs exist in addition to the desired passband. Also, note that $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ has wider transition bands (similar for $|H_{yv}(\mathrm{e}^{j\boldsymbol{\omega}})|$) compared to the transition bands of $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ because the order of $H_{xu}(\mathbf{z})$ for the n_x dimension (similarly, the order of $H_{yv}(\mathbf{z})$ for the n_y dimension) is substantially lower compared to the orders of $H_{uvt}(\mathbf{z})$ for the n_u, n_v and n_t dimensions. The order of the 5-D FIR depth-velocity filter $H(\mathbf{z})$ is $8 \times 8 \times 80 \times 80 \times 40$.

The numerically generated LFV is filtered with the proposed 5-D FIR depth-velocity filter, a 4-D FIR depth filter and a 3-D FIR velocity filter. The 4-D FIR depth filtered LFV is obtained by filtering the original LFV only with $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$. Similarly, the 3-D FIR velocity filtered LFV is obtained by filtering the original LFV only with $H_{uvt}(\mathbf{z})$. The 15th frame of the original, the 4-D FIR depth filtered, the 3-D FIR velocity filtered and the 5-D FIR depth-velocity filtered LFVs corresponding to the central sub-aperture ($n_x, n_y = 5$) are shown in Figure 3.9. It

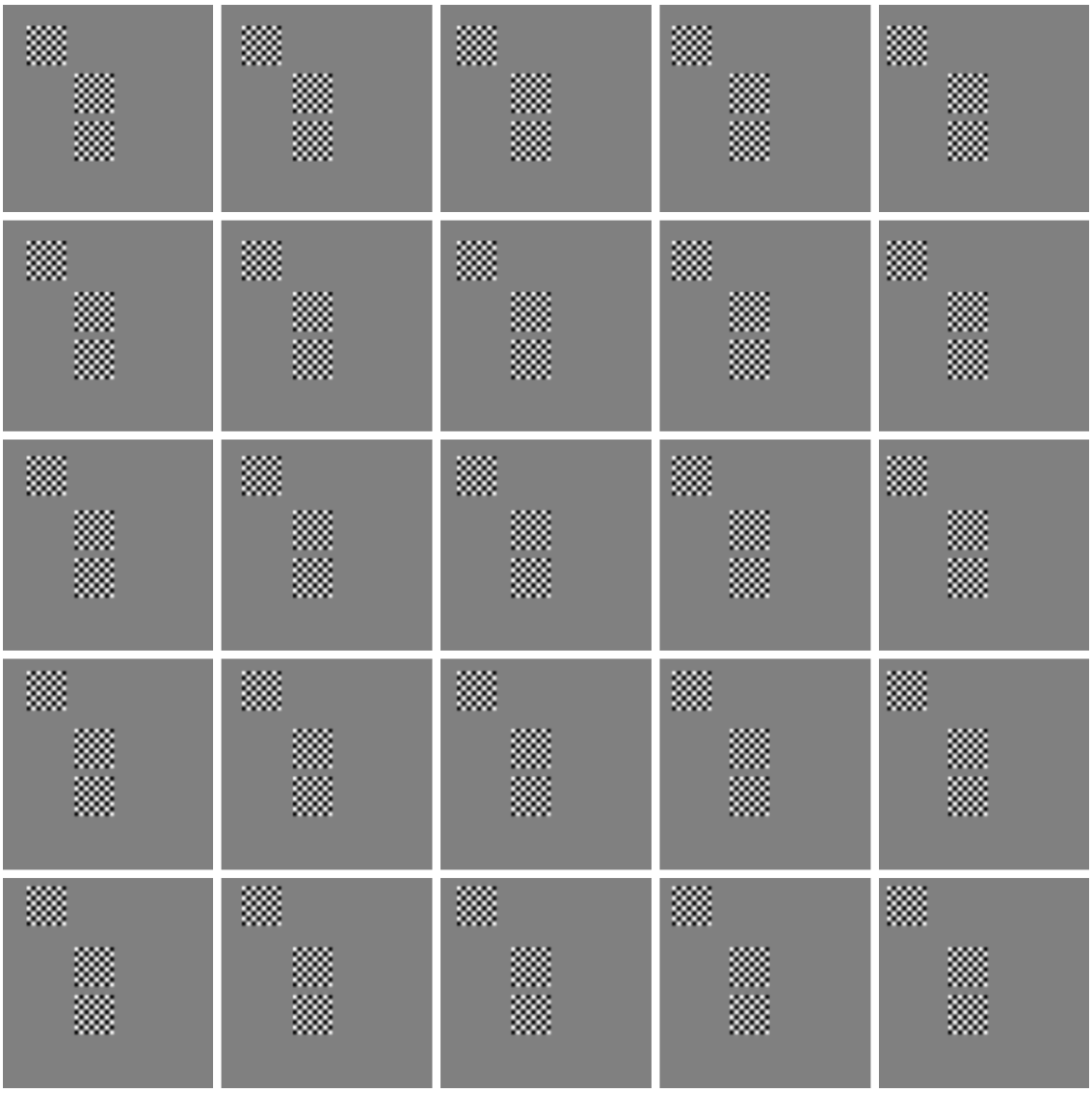


Figure 3.6: The 15th frame of the numerically generated LFV corresponding to the central 25 sub-apertures ($n_x, n_y = 3, 4, \ldots, 7$). The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with a different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at a different depth) of the object of interest.

can be clearly observed that the first interfering object (bottom object) moving at the same depth of the object of interest (middle object) is not attenuated by the 4-D FIR depth filter. Similarly, the second interfering object (top object) moving with the same apparent velocity of the object of interest is not attenuated by the 3-D FIR velocity filter. On the other hand, both interfering objects are attenuated

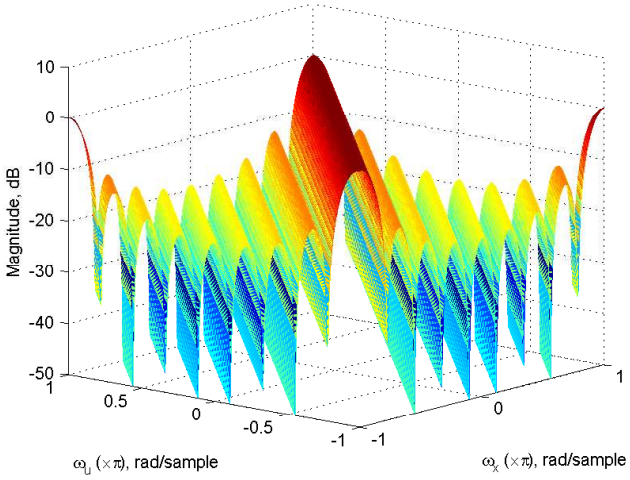


Figure 3.7: Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane.

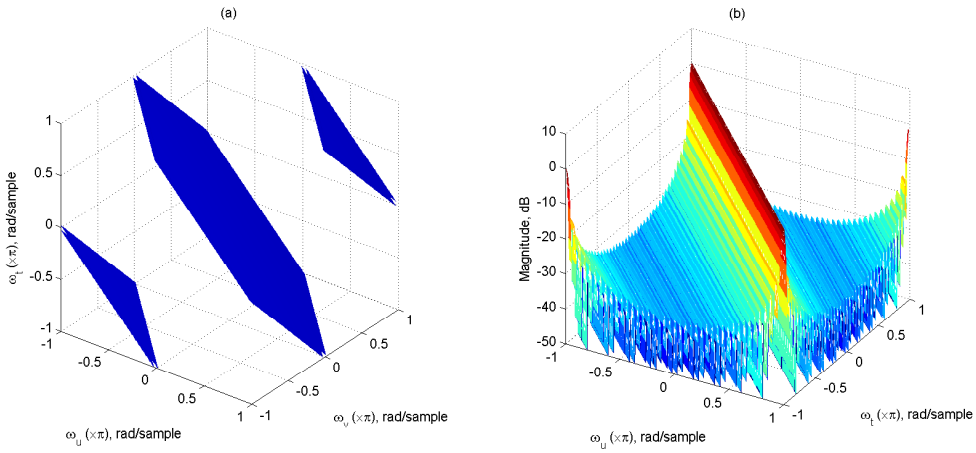


Figure 3.8: Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$.

by the proposed 5-D FIR depth-velocity filter. *This confirms the effectiveness of the proposed 5-D FIR depth-velocity filters, compared to the 4-D FIR depth filters and the 3-D FIR velocity filters, in enhancing objects moving with constant velocity and at constant depth in LFVs.*

3.3.2 Lytro-LF-Camera-Based LFV

The experimental setup employed to generate the LFV is shown in Figure 3.10. It is comprised of a white truck (object of interest), a red truck (moving interfering object)

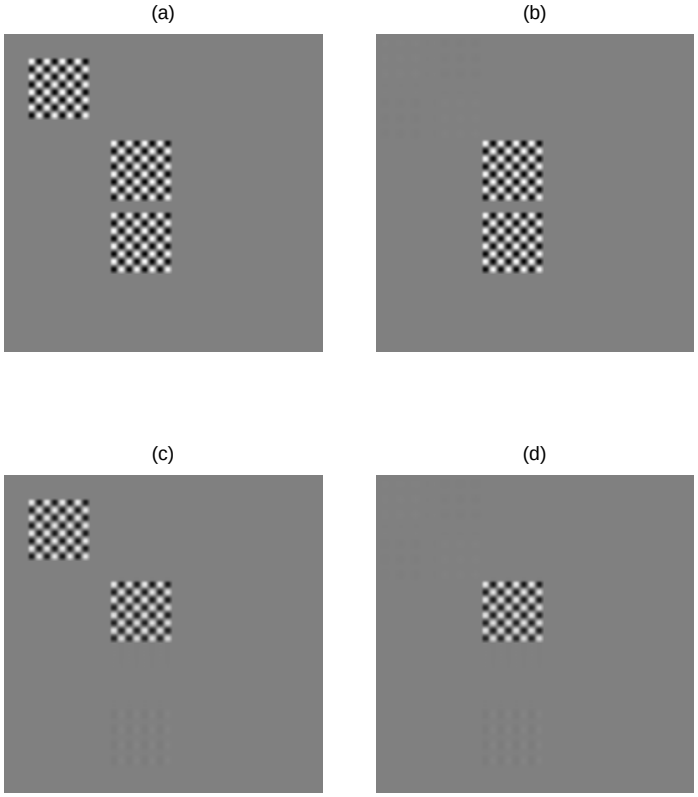


Figure 3.9: The 15th frame of (a) the original, (b) the 4-D FIR depth filtered, (c) the 3-D FIR velocity filtered and (d) the 5-D FIR depth-velocity filtered (proposed) LFVs corresponding to the central sub-aperture ($n_x, n_y = 5$). The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with a different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at a different depth) of the object of interest.

and a fence (static interfering object). The white and red trucks were manually moved in planes approximately parallel to the camera plane, and a commercially available Lytro LF camera [91] [92] was employed to capture the static LFs of the scene. The scene is under homogeneous ambient illumination so that non-Lambertian reflections are minimum. The depth ranges occupied by the white truck and the red truck are $[160, 163]$ cm and $[157, 160]$ cm, respectively, whereas the fence is located at 8 cm deep from the camera plane. The white truck was manually moved from right to left with the speed of 1 cm per frame whereas the red truck was manually moved from left to right with the speed of 1 cm per frame. Sixty LFs were captured, and each LF was decoded by using the MATLAB LFToolbox (Version 0.3) [93]. The generated LFV is of the size $9 \times 9 \times 220 \times 360 \times 60$. The 20th frame of the LFV corresponding to the

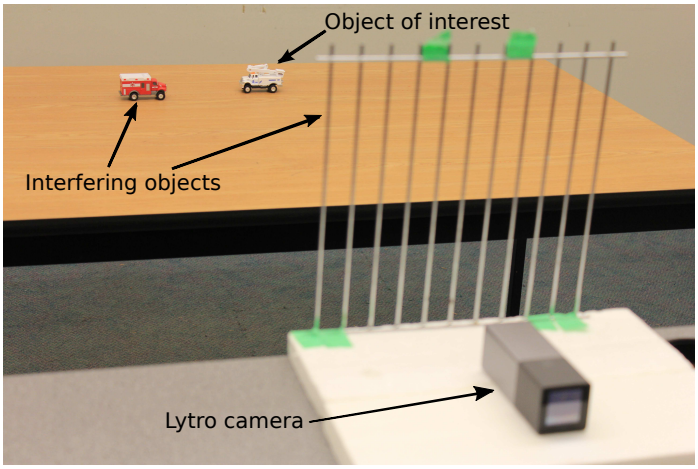


Figure 3.10: The experimental setup employed to generate the LFV of a real scene. A Lytro LF camera was employed to capture the individual frames of the scene. The white truck (object of interest) and the red truck (moving interfering object) move at approximately the same depth. The fence is a static interfering object.

central 15 sub-apertures ($n_x = 4, 5, 6$ and $n_y = 3, 4, \ldots, 7$) is shown in Figure 3.11.

The frequency specifications for the 5-D FIR depth-velocity filter are derived based on the parameterization of the decoded LFV [93] which is a two-plane parameterization different than the one used for the analysis presented in Chapter 2. This leads to the filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ being designed with the parameters $a_u = a_v = -0.33$ and $b_x = b_y = 0.01\pi$ rad/sample. A 2-D rectangular window of size 9×41 is employed in the design, and the order of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ is $8 \times 0 \times 40 \times 0 \times 0$ and $0 \times 8 \times 0 \times 40 \times 0$, respectively. The magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 3.12. The magnitude response of $H_{yv}(\mathbf{z})$ is the same as that of $H_{xu}(\mathbf{z})$. The velocity V_y of the object of interest employed in the experiment is zero (hence, $k_y = 0$). In this special case, $\mathcal{B}_{\mathrm{5D},uvt}$, given in (3.2c), is independent of the ω_v dimension. Consequently, the passband of $H_{uvt}(\mathbf{z})$ is independent of the ω_v dimension, and the finite-extent impulse response $h_{uvt}(\mathbf{n})$, given in (3.8), can be expressed as

$$h_{uvt}(\mathbf{n}) = \left[g^I_{ut}(n_u, n_t)\, w_{ut}(n_u, n_t)\right] \delta(n_x)\delta(n_y)\delta(n_v), \tag{3.9}$$

where $g^I_{ut}(n_u, n_t)$ can be obtained from $g^I_{xu}(n_x, n_u)$ presented in Appendix B, by replacing (n_x, n_u) with (n_u, n_t) and the parameters a_u and b_x with the counterpart parameters a_t and b_u, respectively. A filter $H_{uvt}(\mathbf{z})$ of order $0 \times 0 \times 40 \times 0 \times 40$, designed with the parameters $a_t = 0.25$ and $b_u = 0.005\pi$ rad/sample and by using a 2-D rectangular window of size 41×41 for $w_{ut}(n_u, n_t)$, is employed in the experiment.

Figure 3.11: The 20th frame of the LFV corresponding to the central 15 sub-apertures ($n_x = 4, 5, 6$ and $n_y = 3, 4, \ldots, 7$). The white truck is the object of interest; the red truck is the moving interfering object; the fence is the static interfering object.

The magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 3.13. The order of the 5-D FIR depth-velocity filter $H(\mathbf{z})$ is $8 \times 8 \times 80 \times 40 \times 40$.

Three frames of the original and the filtered LFVs, corresponding to the central sub-aperture ($n_x, n_y = 5$), are shown in Figure 3.14. It can be observed that the object of interest (white truck) is enhanced with slight distortion while the two interfering objects (red truck and the fence) are substantially attenuated. *Most importantly, it*

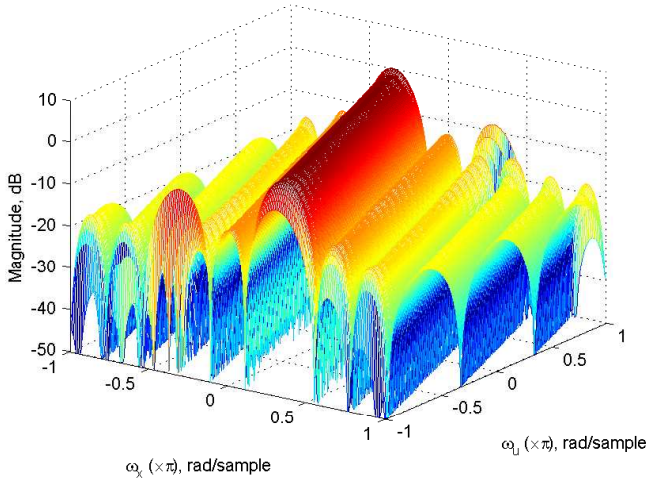


Figure 3.12: Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane.

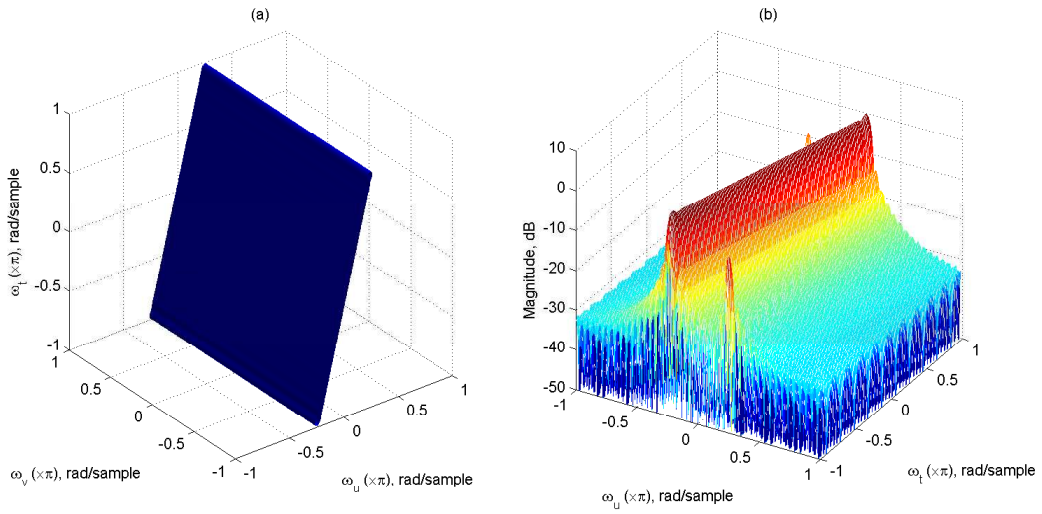


Figure 3.13: Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D FIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$.

can be seen in the 29*th and* 37*th frames that the object of interest, which is considerably occluded in the original LFV, is substantially exposed in the filtered LFV.* These results confirm the successful enhancement of the object of interest based on depth *and* velocity by the 5-D FIR depth-velocity filter.

Next, we evaluate the *de-noising* performance of the proposed 5-D FIR depth-velocity filter. To this end, an additive white Gaussian noise (AWGN) signal of size $9 \times 9 \times 220 \times 360 \times 60$ is added to the original LFV. The standard deviation

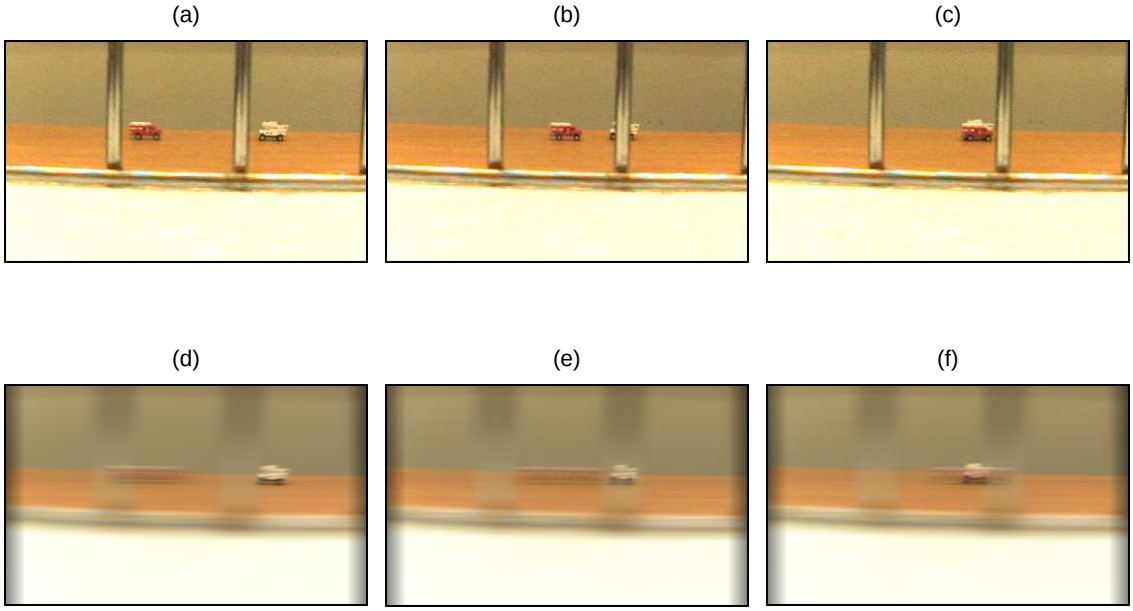


Figure 3.14: Three frames of the original LFV (top row) and the filtered LFV (bottom row), $n_x, n_y = 5$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame.

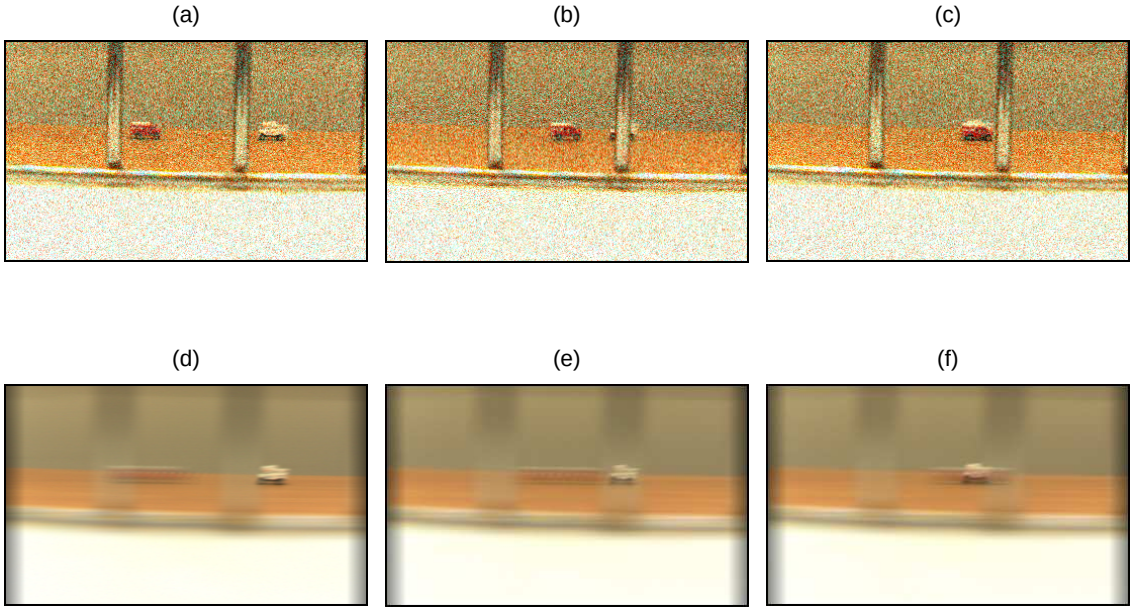


Figure 3.15: Three frames of the corrupted LFV (top row) and the filtered LFV (bottom row), $n_x, n_y = 5$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame.

of the AWGN (σ_{in}) is selected as 40 units to represent a heavily corrupted LFV. Note that the maximum intensity value for the RGB colour components is 255 units.

The average peak-signal-to-noise ratio (PSNR) [88](ch. 1.5) of the input LFV is 16.09 dB/frame. The corrupted LFV is filtered with the proposed 5-D FIR depth-velocity filter to enhance the white truck. Three frames of the corrupted and the filtered LFVs, corresponding to the central sub-aperture ($n_x, n_y = 5$), are shown in Figure 3.15. It is observed that the 5-D FIR depth-velocity filter has attenuated the AWGN significantly so that there is *no visual difference* between the outputs of the noise-free (see Figure 3.14) and the noisy LFVs. In order to numerically evaluate the de-noising performance of the 5-D FIR depth-velocity filter, the standard deviation of the AWGN (σ_{out}) and the average PSNR of the output LFV are calculated. They are found to be 0.73 units and 50.82 dB/frame, respectively. In calculating the average PSNR of the output LFV, the output of the noise-free LFV is considered as the ground truth. The comparatively very high value of the average PSNR of the output LFV confirms the negligible visual difference between the outputs of the noise-free and the noisy LFVs. Furthermore, the *AWGN attenuation* (in dB) is defined as

$$\text{AWGN Attenuation} = 20\log_{10}\left(\frac{\sigma_{\text{in}}}{\sigma_{\text{out}}}\right). \tag{3.10}$$

The AWGN attenuation is 34.73 dB, which is a significantly high value for a de-noising method. Such a high noise attenuation is obtained due to the fact that most of the noise energy reside in the 5-D hypervolume corresponding to the stopband of the 5-D FIR depth-velocity filter. *These results indicate that the proposed 5-D FIR depth-velocity filter provides an excellent de-noising performance.*

3.3.3 Performance Limitations of Proposed 5-D FIR Depth-Velocity Filters

It can be seen from Figures 3.7 and 3.8(b) (also from Figures 3.12 and 3.13(b)) that the selectivity of $H_{xu}(\mathbf{z})$ (and $H_{yv}(\mathbf{z})$) is relatively poor compared to the selectivity of $H_{uvt}(\mathbf{z})$ as a result of the wider transition bands of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$. Among other factors, the selectivity of a filter is determined by the order of the filter, which in turn partly depend on the number of samples available for a particular dimension. For example, in our case, the orders of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ corresponding to the x and y dimensions, respectively, cannot exceed 8 because the number of samples available for those two dimensions (or the number of sub-apertures of the available Lytro LF camera) is 9. On the other hand, the order of $H_{uvt}(\mathbf{z})$ can be as high as $40 \times 40 \times 40$

or even higher. Consequently, if the number of samples available for the x and y dimensions are comparatively low, the performance of 5-D FIR depth-velocity filters on the enhancement of moving objects based on depth may be poor. This limitation on the highest possible order corresponding to the x and y dimensions are especially critical on the performance of FIR filters, for which substantially high orders are necessary to achieve substantial selectivity and stopband attenuation. Furthermore, the depth resolution is affected by the aperture size of the LFV camera. Consequently, in addition to the number of samples available for the x and y dimensions, a sufficiently large aperture is required for better enhancement based on depth. It is expected that the future generations of LFV cameras will offer much increased number of sub-apertures and sufficiently large apertures, which will eliminate the above mentioned drawback.

3.4 Effects of Non-Lambertian Reflections and Occlusion

The effects of non-Lambertian (or specular) reflections and occlusion on the performance of enhancing moving objects in LFVs using the 5-D depth-velocity filters are briefly discussed in this section. Recall that moving objects were assumed to be purely Lambertian in the spectral analysis presented in Chapter 2. This assumption is generally valid for most of natural scenes because almost all of natural materials, except water, are primarily diffusive [9](ch. 4.3), as confirmed by the measurements of their BRDFs [94]. That is, most of the energy of light rays emanating from the object of interest in a natural scene is confined to the ROS obtained for a purely Lambertian moving object in Chapter 2. Furthermore, no occlusion was assumed in deriving the ROS of the spectrum despite the fact that, except some trivial cases, almost all real world scenes suffer from occluded objects. However, as confirmed through the experimental results in Section 3.3, the effects of occlusion on the performance of the 5-D depth-velocity filters in enhancing moving objects in LFVs can be neglected as long as the object of interest is occluded only in a few frames. That is, in such a case, most of the energy of light rays emanating from the object of interest is confined to the ROS of the spectrum obtained for a moving object with no occlusion.

Now let us pay the attention to the case where the object of interest has considerable non-Lambertian reflections or is occluded in most of the frames. In this case,

substantial energy of light rays emanating from the object of interest may reside outside the ROS of the spectrum obtained for a purely Lambertian moving object with no occlusion. Consequently, if the passband of a 5-D depth-velocity filter only encompasses the ROS of the spectrum corresponding to a purely Lambertian moving object with no occlusion, there may appear undesired artifacts. As pointed out in [46] and [47], such artifacts may include the disappearance of non-Lambertian reflections for the object of interest, the appearance of such reflections for interfering objects and the appearance of attenuated interfering objects in place of occluded parts of the object of interest. It is shown in [65] that non-Lambertian reflections and occlusion broaden the ROS of the spectrum of a Lambertian scene[1]. Therefore, such artifacts may be alleviated by appropriately widening the passband of a 5-D depth-velocity filter so that most of the energy of the broadened spectrum resides inside the passband, however, at the cost of reduced selectivity.

3.5 Summary

A novel 5-D linear-phase FIR depth-velocity filter is proposed in this chapter to enhance objects moving with constant velocity and at constant depth in LFVs. The 5-D FIR depth-velocity filter $H(\mathbf{z})$ is designed as a cascade of three 5-D FIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$, each having a 4-D hyperplanar passband of appropriate orientation inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5\mathrm{D}}$. The ideal infinite-extent impulse responses of the three 5-D FIR filters are derived in closed form. The finite-extent impulse responses of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ are obtained by multiplying the ideal infinite-extent impulse responses with a 2-D separable or circular window. Similarly, the finite-extent impulse response of $H_{uvt}(\mathbf{z})$ is obtained by multiplying the ideal infinite-extent impulse response with a 3-D separable or spherical window. For all the possible specifications of the 4-D hyperplanar passbands, $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ can be designed so that the passbands do not have undesired ROSs. However, in the design of $H_{uvt}(\mathbf{z})$, the passband suffers from undesired ROSs for some passband specifications. Nevertheless, thanks to the fact that the undesired ROSs of the passband act like a highpass filter and most of natural scenes are lowpass signals, the effects of the undesired ROSs of the passband may be negligible for most of natural scenes.

The concept of the 5-D depth-velocity filtering of LFVs to enhance moving objects

[1]The reader is referred to [65], [95] and [96] for a more detailed analysis of the effects of non-Lambertian reflections and occlusion on the spectrum of a LF.

is experimentally proved by employing a numerically generated LFV and an LFV of a real scene, generated using a Lytro LF camera. The experimental results obtained using the numerically generated LFV confirm the effectiveness of the proposed 5-D FIR depth-velocity filters, compared to the 4-D FIR depth filters and the 3-D FIR velocity filters, in enhancing objects moving with constant velocity and at constant depth in LFVs. The experimental results obtained using the Lytro-LF-camera-based LFV indicate that the proposed 5-D FIR depth-velocity filter is capable of exposing heavily occluded parts of a scene and of attenuating noise significantly. The very high noise attenuation is obtained as a consequence of most of the noise energy reside in the 5-D hypervolume corresponding to the stopband of the 5-D FIR depth-velocity filter.

The number of sub-apertures available with the current generation of LF and LFV cameras is limited to a few. As a result, the maximum achievable orders of $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ for the x and y dimensions, respectively, are substantially lower compared to the order of $H_{uvt}(\mathbf{z})$. Therefore, $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ have relatively poor selectivity compared to that of $H_{uvt}(\mathbf{z})$. As a consequence, the performance of the proposed 5-D FIR depth-velocity filter on the enhancement of moving objects based on depth may be poor. It is expected that the future generations of LF and LFV cameras will offer much increased number of sub-apertures, which will eliminate the above mentioned drawback.

Chapter 4

Novel Ultra-Low Complexity 5-D IIR Depth-Velocity Filter

4.1 Introduction

The computational complexity[1] of a multi-dimensional filter increases with its order. In general, the order of a multi-dimensional *FIR* filter needs to be considerably higher compared to that of a multi-dimensional *IIR* filter in order to have a good frequency response, i.e., a frequency response having a less passband ripple, a high stopband attenuation and narrow transition bands. For example, as pointed out in [14], the performance of a 3-D FIR beam filter of order $20 \times 20 \times 20$ and the performance of a 3-D IIR beam filter of order $2 \times 2 \times 2$ are similar in broadband beamforming applications. In this example, the 3-D IIR beam filter provides a reduction in computational complexity that exceeds *two orders of magnitude* [14]. Because of the low computational complexity, multi-dimensional IIR filters are of significant interest for many *real-time* applications [14] [24] [29] [50] [97].

In this chapter, a novel ultra-low complexity 5-D IIR depth-velocity filter is proposed for enhancing objects moving with constant velocity and at constant depth in LFVs. The proposed 5-D IIR depth-velocity filter is realized by cascading three first-order 5-D IIR filters having 4-D hyperplanar passbands of appropriate orientations. The first-order 5-D IIR filters are designed by appropriately extending the first-order 3-D IIR planar filter design method proposed in [24]. Numerical simula-

[1]The computational complexity of a 1-D or a multi-dimensional filter means the number of multiplications and additions required by the filter to process a sample.

tion results indicate that the proposed 5-D IIR depth-velocity filter outperforms the 3-D IIR velocity filters [24] and the 4-D IIR depth filters [44] in enhancing moving objects in LFVs. Furthermore, by employing the LFV generated using the commercially available Lytro LF camera, it is shown that the performance of the proposed 5-D IIR depth-velocity filter is comparable to that of the 5-D FIR depth-velocity filter presented in Chapter 3. *Most importantly, the proposed 5-D IIR depth-velocity filter requires less than* 1% *of the arithmetic operations required by the 5-D FIR depth-velocity filter to process a sample.* Considering the ultra-low complexity, the proposed 5-D IIR depth-velocity filter has a significant potential to be employed in *real-time* applications.

The rest of the chapter is organized as follows. In Section 4.2, the design of the 3-D IIR planar filters is briefly reviewed. In Section 4.3, the proposed 5-D IIR depth-velocity filter is described in detail. Experimental results obtained for the numerically generated LFV and the Lytro-LF-camera-based LFV are presented in Section 4.4. Furthermore, the performance of the proposed 5-D IIR depth-velocity filter is compared with that of the 5-D FIR depth-velocity filter presented in Chapter 3. Finally, in Section 4.5, a summary of the chapter is presented.

4.2 Review of Design of the 3-D IIR Planar Filters

The design of the first-order 3-D IIR planar filters is briefly reviewed in this section. The design method is based on the *resonance* of a first-order 3-D pseudo-passive inductor-resistor (RL) network [24] shown in Figure 4.1. Note that $(x, y, t) \in \mathbb{R}^3$ denotes the 3-D spatio-temporal domain, and $(s_x, s_y, s_t) \in \mathbb{C}^3$ is the corresponding 3-D Laplace domain. Furthermore, $L_x, L_y, L_t \geq 0$ are the values of the inductors corresponding to the x, y and t dimensions, respectively, and $R > 0$ is the value of the resistor. In order to describe salient properties of the frequency response of such a 3-D pseudo-passive RL network, we obtain the transfer function as [24]

$$H(s_x, s_y, s_t) = \frac{V_{out}(s_x, s_y, s_t)}{V_{in}(s_x, s_y, s_t)} = \frac{R}{R + L_x s_x + L_y s_y + L_t s_t}. \tag{4.1}$$

Note that the denominator of $H(s_x, s_y, s_t)$ is a strictly Hurwitz polynomial, which ensures that $H(s_x, s_y, s_t)$ is BIBO stable [98]. From (4.1), the frequency response of the 3-D pseudo-passive RL network can be obtained, by evaluating the transfer

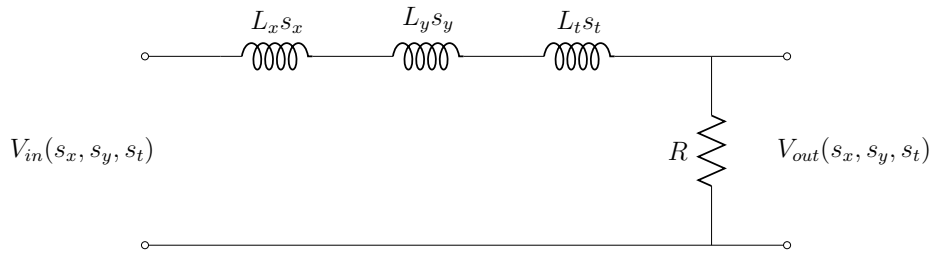


Figure 4.1: First-order 3-D pseudo-passive RL network.

function $H(s_x, s_y, s_t)$ at $(s_x, s_y, s_t) = (j\Omega_x, j\Omega_y, j\Omega_t)$, as

$$H(j\Omega_x, j\Omega_y, j\Omega_t) = \frac{R}{R + j(L_x\Omega_x + L_y\Omega_y + L_t\Omega_t)}, \tag{4.2}$$

where $(\Omega_x, \Omega_y, \Omega_t) \in \mathbb{R}^3$ is the 3-D continuous frequency domain. Note that the magnitude response $|H(j\Omega_x, j\Omega_y, j\Omega_t)|$ is given by [24]

$$|H(j\Omega_x, j\Omega_y, j\Omega_t)| = \frac{R}{[R^2 + (L_x\Omega_x + L_y\Omega_y + L_t\Omega_t)^2]^{1/2}}. \tag{4.3}$$

According to (4.3), the magnitude response has a maximum value of *unity* at the *resonant plane*, given by [24]

$$L_x\Omega_x + L_y\Omega_y + L_t\Omega_t = 0. \tag{4.4}$$

The normal vector associated with the resonant plane is $\mathbf{L} = [L_x, L_y, L_t]^{\mathrm{T}}$. Furthermore, $|H(j\Omega_x, j\Omega_y, j\Omega_{ct})| = 1/\sqrt{2}$ when

$$L_x\Omega_x + L_y\Omega_y + L_t\Omega_t = \pm R, \tag{4.5}$$

and the corresponding two planes are called -3 dB planes [24]. Note that the -3 dB bandwidth B of the magnitude response is given by

$$B = \frac{R}{||\mathbf{L}||_2}, \tag{4.6}$$

where $||\cdot||_2$ is the 2-norm of a vector. When $||\mathbf{L}||_2 \gg R$, the magnitude response $|H(j\Omega_x, j\Omega_y, j\Omega_t)|$ rapidly decreases outside the resonant plane, approximating a *highly-selective planar passband* with a -3 dB bandwidth $B \ll 1$. Hence, the resulting filter is called a 3-D *planar filter*.

In order to obtain the 3-D IIR planar filter, we first express the transfer function $H(s_x, s_y, s_t)$ in a more compact form as [97]

$$H(s_x, s_y, s_t) = \frac{1}{1 + \frac{1}{B}(d_x s_x + d_y s_y + d_t s_t)}, \tag{4.7}$$

where $\widehat{\mathbf{d}} = [d_x, d_y, d_t]^{\mathrm{T}} = \mathbf{L}/||\mathbf{L}||_2$ is the *unit* normal vector of the resonant plane, and then apply the triple bilinear transformation

$$s_i = \frac{z_i - 1}{z_i + 1}, \qquad i = x, y, t, \tag{4.8}$$

where $(z_x, z_y, z_t) \in \mathbb{C}^3$ is the 3-D z domain. This yields the first-order 3-D discrete-domain transfer function $H(z_x, z_y, z_t)$ [24]

$$H(z_x, z_y, z_t) = \frac{\sum_{i_x=0}^{1} \sum_{i_y=0}^{1} \sum_{i_t=0}^{1} z_x^{-i_x} z_y^{-i_y} z_t^{-i_t}}{\sum_{i_x=0}^{1} \sum_{i_y=0}^{1} \sum_{i_t=0}^{1} b_{i_x i_y i_t} z_x^{-i_x} z_y^{-i_y} z_t^{-i_t}}, \tag{4.9}$$

where the denominator coefficients are given by

$$b_{i_x i_y i_t} = 1 + \frac{(-1)^{i_x} d_x + (-1)^{i_y} d_y + (-1)^{i_t} d_t}{B}, \qquad i_x, i_y, i_t = 0, 1. \tag{4.10}$$

It is worthwhile to note that $H(z_x, z_y, z_t)$ may not be BIBO stable because the application of the triple bilinear transformation leads to a 3-D discrete-domain transfer function having a nonessential singularity of the second kind on the unit polydisc at $(z_x, z_y, z_t) = (-1, -1, -1)$ [97] [99] [100]. However, $H(z_x, z_y, z_t)$ is *practical-BIBO stable* [101] [102] for all 3-D amplitude-bounded input signals having finite and sufficiently small ROSs in the 3-D spatio-temporal domain (x, y, t) with respect to at least two of the three dimensions. This condition is satisfied by most of *practical* 3-D signals since the number of samples available for the two spatial dimensions is always finite and mostly small compared to the number of samples available for the temporal dimension.

Another fact of importance is the restriction of the unit normal vector $\widehat{\mathbf{d}}$ to the first closed octant of the 3-D frequency domain (i.e., all the three elements of the unit normal vector $\widehat{\mathbf{d}}$ should be non-negative) as a consequence of the non-negativity of the values (L_x, L_y and L_t) of the inductors employed in the 3-D pseudo-passive RL

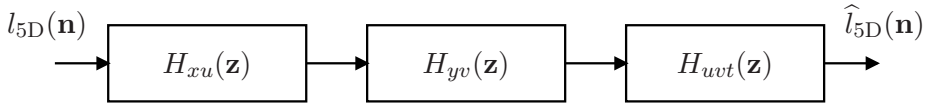


Figure 4.2: Structure of the 5-D IIR depth-velocity filter.

network. Consequently, for the applications where the required planar passband has a unit normal vector $\widehat{\mathbf{d}}$ that lies outside the first closed octant of the 3-D frequency domain, the denominator coefficients of the 3-D IIR planar filter should be calculated (using (4.10)) by employing the absolute values of the negative elements of the unit normal vector $\widehat{\mathbf{d}}$. Furthermore, in the filter implementation, the direction of recursion should be reversed for all the dimensions corresponding to the negative elements [24] [98].

4.3 Design of the 5-D IIR Depth-Velocity Filter

The structure of the proposed 5-D IIR depth-velocity filter $H(\mathbf{z})$ is shown in Figure 4.2. Similar to the design of the 5-D FIR depth-velocity filter presented in Chapter 3, the 5-D IIR depth-velocity filter is designed as a cascade of three 5-D IIR filters, $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$. The transfer function $H(\mathbf{z})$ of the 5-D IIR depth-velocity filter is given by

$$H(\mathbf{z}) = H_{xu}(\mathbf{z})\, H_{yv}(\mathbf{z})\, H_{uvt}(\mathbf{z}). \tag{4.11}$$

Note that the passbands of the 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ encompass the 5-D hypervolumes given by

$$\mathcal{B}_{5D,xu} = \bigcup_{z_0} \mathcal{H}_{5D,xu} \tag{4.12a}$$

$$\mathcal{B}_{5D,yv} = \bigcup_{z_0} \mathcal{H}_{5D,yv} \tag{4.12b}$$

$$\mathcal{B}_{5D,uvt} = \bigcup_{z_0} \mathcal{H}_{5D,uvt} \tag{4.12c}$$

respectively, inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5D}$ as illustrated in Figure 4.3. Here, $\mathcal{H}_{5D,xu}$, $\mathcal{H}_{5D,yv}$ and $\mathcal{H}_{5D,uvt}$ are the 4-D hyperplanes corresponding to the ROS of the spectrum of the object of interest, which are given by (2.27a), (2.27b) and (2.27c), respectively. The overall passband $\mathcal{B}_{5D}$ of $H(\mathbf{z})$, given by the intersection

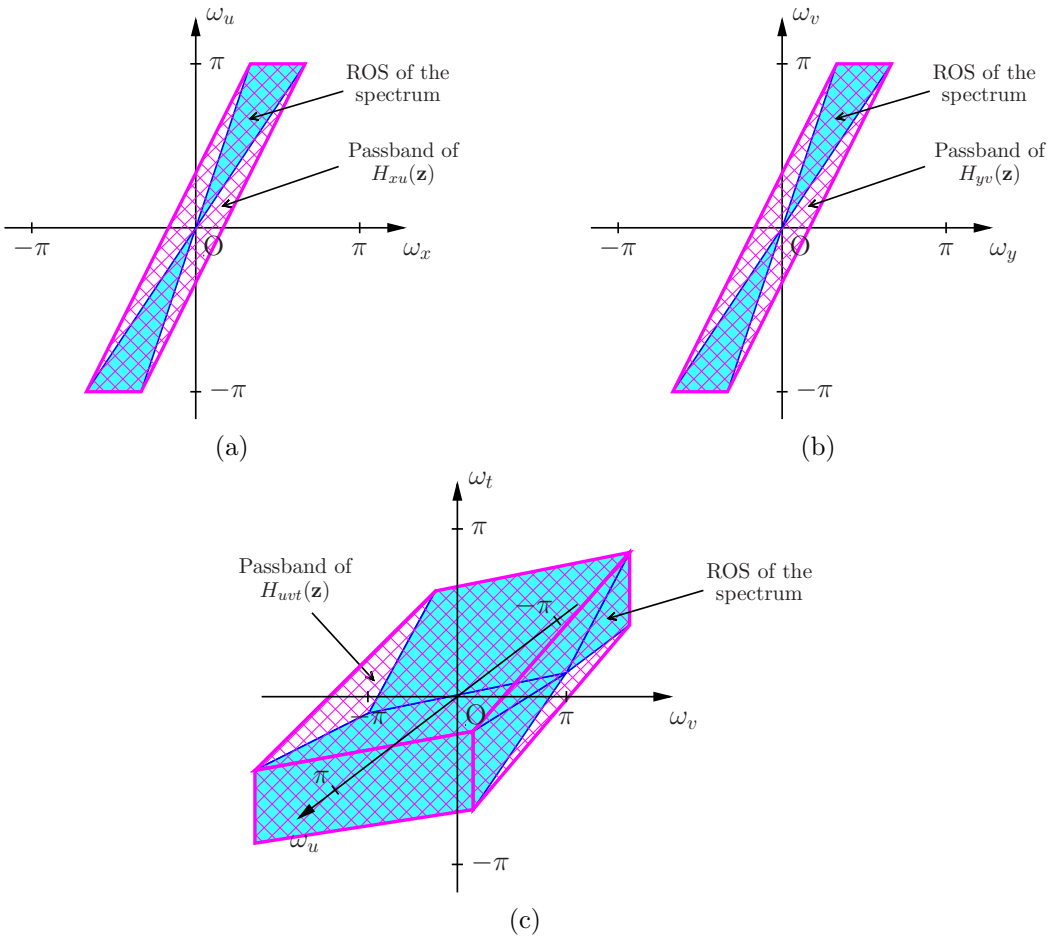


Figure 4.3: The ROS of the spectrum of the object of interest (solid) and the passbands of the 5-D IIR filters (cross-hatched); (a) $H_{xu}(\mathbf{z})$ on the $\omega_x\omega_u$ plane; (b) $H_{yv}(\mathbf{z})$ on the $\omega_y\omega_v$ plane; (c) $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space.

of $\mathcal{B}_{5\mathrm{D},xu}$, $\mathcal{B}_{5\mathrm{D},yv}$ and $\mathcal{B}_{5\mathrm{D},uvt}$, completely encompasses the 3-D hyperfan corresponding to the spectral ROS of the object of interest.

The 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ are designed by appropriately extending the first-order 3-D IIR planar filter design method [24] reviewed in Section 4.2. In particular, the transfer functions of the three filters are obtained as

$$H_{xu}(\mathbf{z}) = \frac{\sum_{i_x=0}^{1}\sum_{i_u=0}^{1} z_x^{-i_x} z_u^{-i_u}}{\sum_{i_x=0}^{1}\sum_{i_u=0}^{1} b_{i_x i_u} z_x^{-i_x} z_u^{-i_u}} \tag{4.13a}$$

$$H_{yv}(\mathbf{z}) = \frac{\sum_{i_y=0}^{1}\sum_{i_v=0}^{1} z_y^{-i_y} z_v^{-i_v}}{\sum_{i_x=0}^{1}\sum_{i_u=0}^{1} b_{i_y i_v} z_y^{-i_y} z_v^{-i_v}} \tag{4.13b}$$

$$H_{uvt}(\mathbf{z}) = \frac{\sum_{i_u=0}^{1}\sum_{i_v=0}^{1}\sum_{i_t=0}^{1} z_u^{-i_u} z_v^{-i_v} z_t^{-i_t}}{\sum_{i_u=0}^{1}\sum_{i_v=0}^{1}\sum_{i_t=0}^{1} b_{i_u i_v i_t} z_u^{-i_u} z_v^{-i_v} z_t^{-i_t}}. \tag{4.13c}$$

Here, the denominator coefficients of $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ are, respectively, obtained as [24]

$$b_{i_x i_u} = 1 + \frac{(-1)^{i_x} d_{xu}^{x} + (-1)^{i_u} d_{xu}^{u}}{B_{xu}}, \qquad i_x, i_u = 0, 1 \tag{4.14a}$$

$$b_{i_y i_v} = 1 + \frac{(-1)^{i_y} d_{yv}^{y} + (-1)^{i_v} d_{yv}^{v}}{B_{yv}}, \qquad i_y, i_v = 0, 1 \tag{4.14b}$$

$$b_{i_u i_v i_t} = 1 + \frac{(-1)^{i_u} d_{uvt}^{u} + (-1)^{i_v} d_{uvt}^{v} + (-1)^{i_t} d_{uvt}^{t}}{B_{uvt}}, \qquad i_u, i_v, i_t = 0, 1, \tag{4.14c}$$

where B_{xu}, B_{yv} and B_{uvt} are the parameters that determine the -3 dB bandwidths of the three filters, respectively, and $\widehat{\mathbf{d}}_{xu} = [d_{xu}^{x}, 0, d_{xu}^{u}, 0, 0]^{\mathrm{T}}$, $\widehat{\mathbf{d}}_{yv} = [0, d_{yv}^{y}, 0, d_{yv}^{v}, 0]^{\mathrm{T}}$ and $\widehat{\mathbf{d}}_{uvt} = [0, 0, d_{uvt}^{u}, d_{uvt}^{v}, d_{uvt}^{t}]^{\mathrm{T}}$ are the unit normal vectors corresponding to the 4-D hyperplanar passbands of the three filters, respectively. Furthermore, the partial difference equations for $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ are, respectively, given by

$$l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}) = \frac{1}{b_{00}}\left[\sum_{i_x=0}^{1}\sum_{i_u=0}^{1} l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n} - \mathbf{i}_{xu}) - \sum_{\substack{i_x=0 \\ i_x+i_u\neq 0}}^{1}\sum_{i_u=0}^{1} b_{i_x i_u} l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n} - \mathbf{i}_{xu})\right] \tag{4.15a}$$

$$l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}) = \frac{1}{b_{00}}\left[\sum_{i_y=0}^{1}\sum_{i_v=0}^{1} l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n} - \mathbf{i}_{yv}) - \sum_{\substack{i_x=0 \\ i_y+i_v\neq 0}}^{1}\sum_{i_u=0}^{1} b_{i_y i_v} l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n} - \mathbf{i}_{yv})\right] \tag{4.15b}$$

$$l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}) = \frac{1}{b_{000}}\left[\sum_{i_u=0}^{1}\sum_{i_v=0}^{1}\sum_{i_t=0}^{1} l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n} - \mathbf{i}_{uvt}) - \sum_{\substack{i_u=0 \\ i_u+i_v+i_t\neq 0}}^{1}\sum_{i_v=0}^{1}\sum_{i_t=0}^{1} b_{i_u i_v i_t} l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n} - \mathbf{i}_{uvt})\right], \tag{4.15c}$$

where $l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n})$ and $l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n})$ are the input and the output signals of each filter, respectively. Further, $\mathbf{i}_{xu} = [i_x, 0, i_u, 0, 0]^{\mathrm{T}}$, $\mathbf{i}_{yv} = [0, i_y, 0, i_v, 0]^{\mathrm{T}}$ and $\mathbf{i}_{uvt} = [0, 0, i_u, i_v, i_t]^{\mathrm{T}}$. Note that the order of the 5-D IIR depth-velocity filter $H(\mathbf{z})$ is $1 \times 1 \times 2 \times 2 \times 1$.

In order to satisfy the conditions for the practical-BIBO stability [101], the unit normal vectors $\widehat{\mathbf{d}}_{xu}$, $\widehat{\mathbf{d}}_{yv}$ and $\widehat{\mathbf{d}}_{uvt}$ corresponding to the 4-D hyperplanar passbands of the 5-D IIR filters $H_{xu}(z_x, z_u)$, $H_{yv}(z_y, z_v)$ and $H_{uvt}(z_u, z_v, z_t)$, respectively, should always lie in the first closed orthant of the 5-D frequency domain, i.e., all five elements of a vector should be non-negative. If a unit normal vector lies outside the first closed orthant, the direction of recursion should be reversed for all dimensions having negative elements so that the new unit normal vector lies in the first closed orthant [24].

4.4 Experimental Results

In this section, experimental results are presented to confirm the effectiveness of the proposed 5-D IIR depth-velocity filter in enhancing objects moving with constant velocity and at constant depth in LFVs. First, by employing the numerically generated LFV, it is shown that the proposed 5-D IIR depth-velocity filter outperforms the counterpart 3-D IIR velocity and 4-D IIR depth filters reported in [24] and [44], respectively. Next, the LFV generated using the Lytro LF camera is employed to compare the performance of the proposed 5-D IIR depth-velocity filter with that of the 5-D FIR depth-velocity filter presented in Chapter 3.

4.4.1 Numerically Generated LFV

Recall that the numerically generated LFV contains three Lambertian planar objects moving with constant velocity and at constant depth, of which the specifications are given in Table 3.1. Note that the object of interest and the interfering object 1 move at the same depth so that the interfering object 1 can not be attenuated using 4-D depth filters. On the other hand, the object of interest and the interfering object 2 move with the same apparent velocity so that the interfering object 2 can not be attenuated using 3-D velocity filters.

The unit normal vectors corresponding to the 4-D hyperplanar passbands of the 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ are obtained as $\widehat{\mathbf{d}}_{xu} = [0.71, 0, -0.71, 0, 0]^{\mathrm{T}}$, $\widehat{\mathbf{d}}_{yv} = [0, 0.71, 0, -0.71, 0]^{\mathrm{T}}$ and $\widehat{\mathbf{d}}_{uvt} = [0, 0, 0.58, 0.58, 0.58]^{\mathrm{T}}$, respectively. The filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ are designed with the parameters $B_{xu} = B_{yv} = 0.06\pi$ rad/sample. The magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 4.4. Note that the *narrower transition bands* of the magnitude response of the 5-D IIR filter $H_{xu}(\mathbf{z})$ compared to

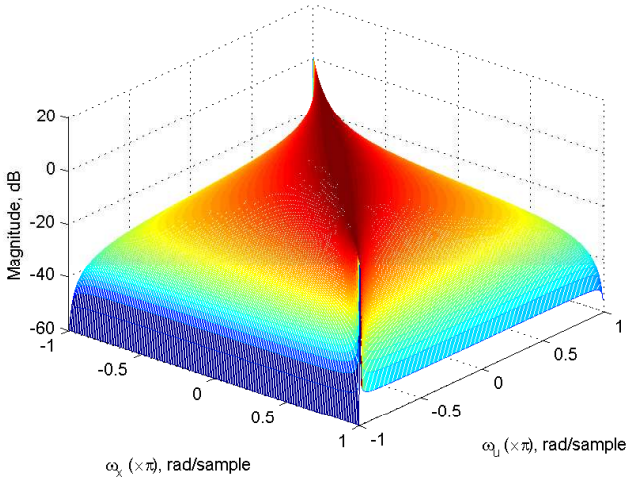


Figure 4.4: Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane.

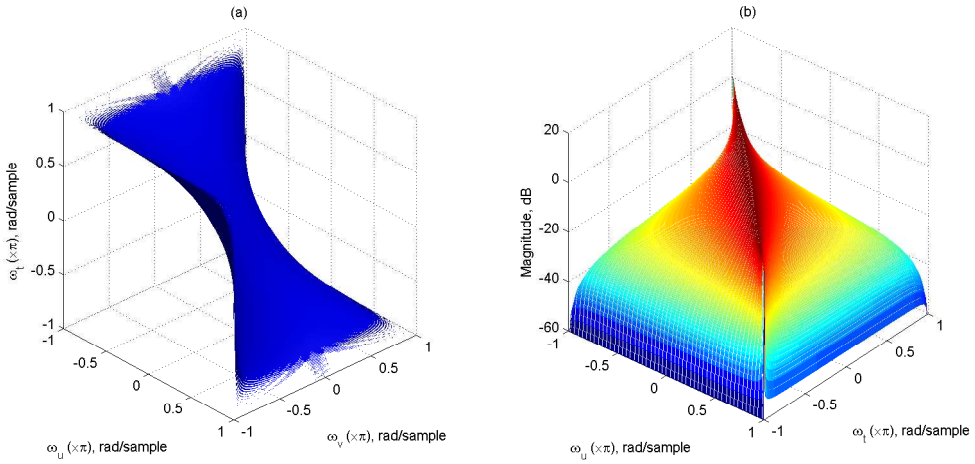


Figure 4.5: Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$.

that of the FIR counterpart, shown in Figure 3.7. The magnitude response of $H_{yv}(\mathbf{z})$ is similar to that of $H_{xu}(\mathbf{z})$. The filter $H_{uvt}(\mathbf{z})$ is designed with the parameter $B_{uvt} = 0.03\pi$ rad/sample. The -3 dB iso-surface of the magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 4.5. It is observed that the passband of $H_{uvt}(\mathbf{z})$ deviates from the desired 4-D hyperplane at higher frequencies as a consequence of the so-called frequency warping of the bilinear transform.

The numerically generated LFV is filtered with the proposed 5-D IIR depth-velocity filter, the 4-D IIR depth filter [44] and the 3-D IIR velocity filter [24]. The

5-D IIR depth-velocity filter is implemented in the spatio-temporal domain using the partial difference equations given in (4.15a)–(4.15c). The 15th frame of the original, the 4-D IIR depth filtered, the 3-D IIR velocity filtered and the 5-D IIR depth-velocity filtered LFVs corresponding to the central sub-aperture $(n_x, n_y) = (5, 5)$ are shown in Figure 4.6. Note that the interfering object 1 (bottom object) moving at the same depth of the object of interest (middle object) is not attenuated by the 4-D IIR depth filter. Similarly, the interfering object 2 (top object) moving with the same apparent velocity of the object of interest is not attenuated by the 3-D IIR velocity filter. On the other hand, *both interfering objects are attenuated by the 5-D IIR depth-velocity filter*. This confirms the effectiveness of the proposed 5-D IIR depth-velocity filter, compared to the 3-D IIR velocity filter [24] and the 4-D IIR depth filter [44], in enhancing objects moving with constant velocity and at constant depth in LFVs.

As a side note, it is observed that the achieved attenuation for the interfering object 1 is slightly poor compared to that for the interfering object 2 by the proposed 5-D IIR depth-velocity filter. This results as a consequence of overlapping a part of the spectral ROS of the interfering object 1 with the passband of $H_{uvt}(\mathbf{z})$, which deviates from the desired 4-D hyperplane at higher frequencies.

4.4.2 Lytro-LF-Camera-Based LFV

Recall that the Lytro-LF-camera-based LFV comprised of a white truck (object of interest), a red truck (moving interfering object) and a fence (static interfering object). Note that the depth ranges occupied by the white truck and the red truck are $[160, 163]$ cm and $[157, 160]$ cm, respectively, whereas the fence is located at 8 cm deep from the camera plane. Furthermore, the white truck was manually moved from right to left with the speed of 1 cm per frame whereas the red truck was manually moved from left to right with the speed of 1 cm per frame.

The unit normal vectors corresponding to the 4-D hyperplanar passbands of the 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ are derived based on the parameterization of the decoded LFV [93] which is a two-plane parameterization different than the one used for the analysis presented in Chapter 2. This leads to the unit normal vectors $\widehat{\mathbf{d}}_{xu} = [0.95, 0, 0.32, 0, 0]^{\mathrm{T}}$, $\widehat{\mathbf{d}}_{yv} = [0, 0.95, 0, 0.32, 0]^{\mathrm{T}}$ and $\widehat{\mathbf{d}}_{uvt} = [0, 0, -0.97, 0, 0.24]^{\mathrm{T}}$. Note that the velocity V_y of the object of interest employed in the experiment is zero, hence, $k_y = 0$. In this special case, $H_{uvt}(\mathbf{z})$ is independent of z_v, and the order of $H(\mathbf{z})$ is $1 \times 1 \times 2 \times 1 \times 1$. The filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ are designed with the

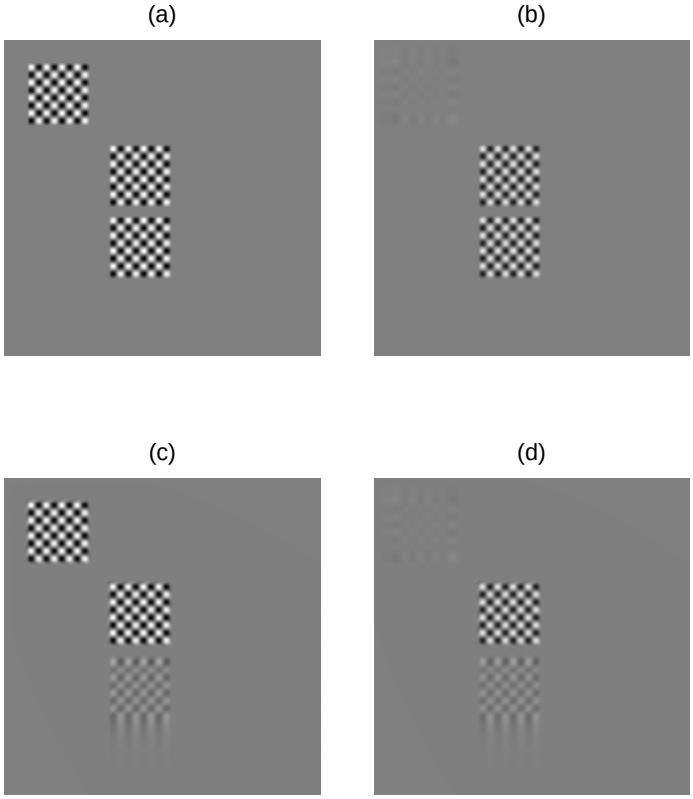


Figure 4.6: The 15th frame of (a) the original, (b) the 4-D IIR depth filtered [44], (c) the 3-D IIR velocity filtered [24] and (d) the 5-D IIR depth-velocity filtered (proposed) LFVs corresponding to the central sub-aperture $(n_x, n_y) = (5, 5)$. The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with a different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at a different depth) of the object of interest.

parameters $B_{xu} = B_{yv} = 0.04\pi$ rad/sample. The magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 4.7. The magnitude response of $H_{yv}(\mathbf{z})$ is similar to that of $H_{xu}(\mathbf{z})$. The filter $H_{uvt}(\mathbf{z})$ is designed with the parameter $B_{uvt} = 0.006\pi$ rad/sample. The magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ is shown in Figure 4.8. In this case, it is observed that the passbands of all the three filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ deviate from the corresponding desired 4-D hyperplanes at higher frequencies due to the frequency warping of the bilinear transform.

The LFV is filtered with the proposed 5-D IIR depth-velocity filter, and three frames of the original and the filtered LFVs, corresponding to the sub-aperture $(n_x, n_y) = (7, 7)$, are shown in Figure 4.9. Similar to the 5-D FIR depth-velocity filter presented in Chapter 3, the proposed 5-D IIR depth-velocity filter enhances the

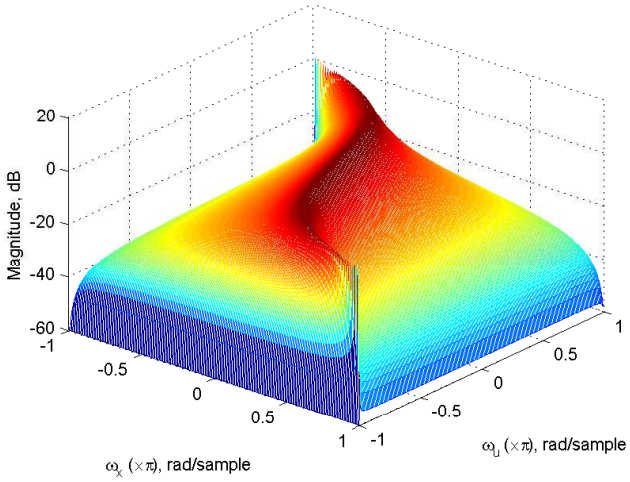


Figure 4.7: Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{xu}(\mathbf{z})$ in the $\omega_x\omega_u$ plane.

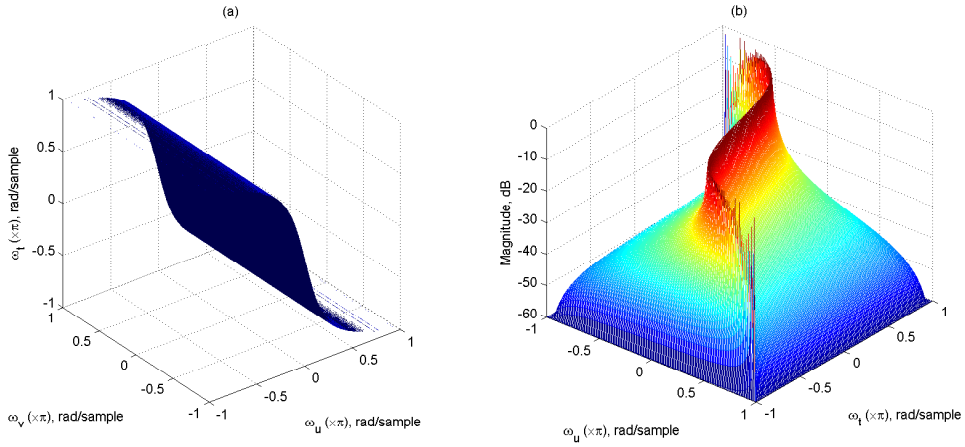


Figure 4.8: Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})|$ of the 5-D IIR filter $H_{uvt}(\mathbf{z})$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$.

object of interest (white truck) with slight distortion while substantially attenuating the two interfering objects (red truck and the fence). Furthermore, it can be seen in the 29th and 37th frames that the object of interest, which is considerably occluded in the original LFV, is substantially exposed in the filtered LFVs. *These results confirm that the performance of the proposed 5-D IIR depth-velocity filter is comparable to that of the 5-D FIR depth-velocity filter.*

Next, we evaluate the *de-noising* performance of the proposed 5-D IIR depth-velocity filter. To this end, the heavily corrupted LFV, employed in Chapter 3, is

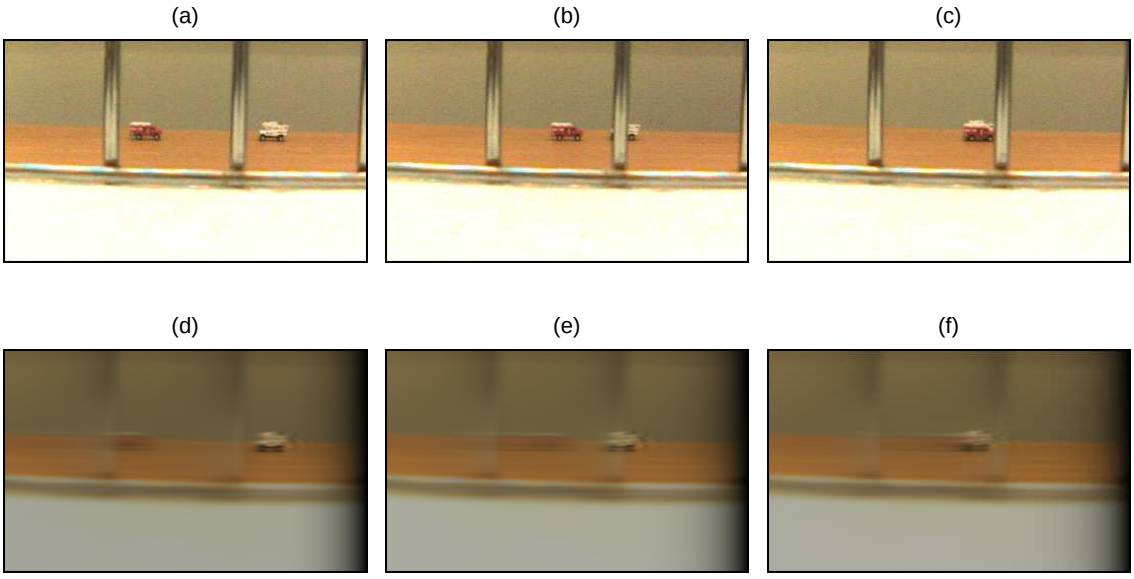


Figure 4.9: Three frames of the original LFV (top row) and the 5-D IIR depth-velocity filtered LFV (bottom row), $(n_x, n_y) = (7, 7)$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame.

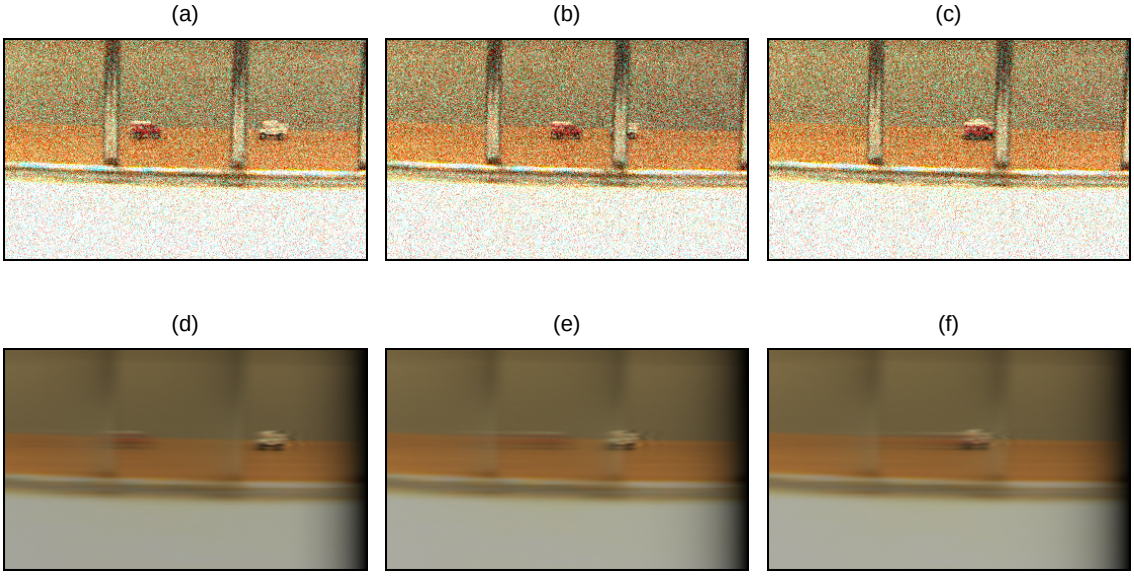


Figure 4.10: Three frames of the corrupted LFV (top row) and the 5-D IIR depth-velocity filtered LFV (bottom row), $(n_x, n_y) = (7, 7)$; (a) and (d) 20th frame; (b) and (e) 29th frame; (c) and (f) 37th frame.

filtered with the proposed 5-D IIR depth-velocity filter to enhance the white truck. Note that the standard deviation of the AWGN (σ_{in}) is 40 units, and the average PSNR of the input LFV is 16.09 dB/frame. Three frames of the corrupted and the

filtered LFVs, corresponding to the sub-aperture $(n_x, n_y) = (7, 7)$, are shown in Figure 4.10. Similar to the 5-D FIR depth-velocity filter presented in Chapter 3, it is observed that the 5-D IIR depth-velocity filter has attenuated the AWGN significantly so that there is *no visual difference* between the outputs of the noise-free (see Figure 4.9) and the noisy LFVs. The standard deviation of the AWGN (σ_{out}) and the average PSNR of the output LFV are found to be 0.44 units and 55.33 dB/frame, respectively. Similar to the case of the 5-D FIR depth-velocity filter, in calculating the average PSNR of the output LFV, the output of the noise-free LFV is considered as the ground truth. The AWGN attenuation achieved with the 5-D IIR depth-velocity filter is 39.24 dB. Recall that the AWGN attenuation achieved with the 5-D FIR depth-velocity filter is 34.73 dB. *Accordingly, the proposed 5-D IIR depth-velocity filter provides an improvement of* 4.51 *dB in attenuating AWGN, compared to the 5-D FIR depth-velocity filter.* This results as a consequence of *the higher stopband attenuation and the narrower transition bands* of the 5-D IIR depth-velocity filter compared to the FIR counterpart.

Now, we compare the computational complexity of the proposed 5-D IIR depth-velocity filter with that of the 5-D FIR depth-velocity filter presented in Chapter 3. To this end, the number of multiplications (excluding the trivial multiplications with 1) and additions required to process a sample of the Lytro-LF-camera-based LFV is calculated. *The proposed 5-D IIR depth-velocity filter (of order* $1 \times 1 \times 2 \times 1 \times 1$*) requires only* 12 *multiplications and* 18 *additions whereas the 5-D FIR depth-velocity filter (of order* $8 \times 8 \times 80 \times 40 \times 40$*) requires* 1211 *multiplications and* 2416 *additions.* Here, we assume that both filters are implemented using direct form structures, and the symmetry of the coefficients is utilized to reduce the number of multiplications required by the 5-D FIR depth-velocity filter. Accordingly, *the proposed 5-D IIR depth-velocity filter requires only* 0.99% *of the multiplications and* 0.75% *of the additions required by the 5-D FIR depth-velocity filter to process a sample.* In other words, the implementation of the 5-D IIR depth-velocity filter requires less than 1% of the resources required by that of the 5-D FIR depth-velocity filter.

4.4.3 Performance Limitations of Proposed 5-D IIR Depth-Velocity Filters

Even though the proposed 5-D IIR depth-velocity filter eliminates the main drawback of the 5-D FIR depth-velocity filter presented in Chapter 3, i.e. the poor selectivity of

$H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ due to the limited number of samples (or sub-apertures) available for the n_x and n_y dimensions, it has its own drawbacks. Note that, in the case of the Lytro-LF-camera-based LFV, the output LFV of the 5-D IIR depth-velocity filter has less brightness compared to the original LFV (see Figure 4.9) and the output LFV of the 5-D FIR depth-velocity filter (see Figure 3.14). This results as a consequence of the substantially long transient response of the 5-D IIR depth-velocity filter and the limited number of samples available for the n_x and n_y dimensions. In other words, the output LFV of the 5-D IIR depth-velocity filter has not reached to the *steady state* with respect to the n_x and n_y dimensions. This effect may be overcome by employing a post-processing technique such as histogram equalization [103] [104] [105] [106] [107]. On the other hand, we expect that the future generations of LFV cameras will offer much increased number of sub-apertures, which will eliminate this drawback.

Furthermore, the passbands of the 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ deviate from the corresponding desired 4-D hyperplanes at higher frequencies due to the *frequency warping* of the bilinear transform. Because most of natural scenes are low-pass signals, the effect of this drawback may be negligible for most of natural scenes. However, as observed in the case of the numerically generated LFV, in some LFVs, the attenuation of some interfering objects may be poor as a result of overlapping of parts of the spectral ROSs of those interfering objects with the distorted passbands of the 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$.

4.5 Summary

In this chapter, a novel ultra-low complexity 5-D IIR depth-velocity filter is proposed for enhancing objects moving with constant velocity and at constant depth in LFVs. The proposed 5-D IIR depth-velocity filter is realized by cascading three first-order 5-D IIR filters, $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$, each having a 4-D hyperplanar passband of appropriate orientation inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5\mathrm{D}}$. The first-order 5-D IIR filters are designed by appropriately extending the first-order 3-D IIR planar filter design method proposed in [24]. The 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ are practical-BIBO stable as long as the corresponding unit normal vectors lie inside the first closed orthant of the 5-D frequency domain.

The experimental results obtained for the numerical generated LFV confirm that the proposed 5-D IIR depth-velocity filter outperforms the 3-D IIR velocity filters [24] and the 4-D IIR depth filters [44] in enhancing moving objects in LFVs. Furthermore,

the experimental results obtained for the Lytro-LF-camera-based LFV indicate that the performance of the proposed 5-D IIR depth-velocity filter is comparable to that of the 5-D FIR depth-velocity filter presented in Chapter 3 when noise is negligible. On the other hand, when the LFV is heavily corrupted by noise, the 5-D IIR depth-velocity filter provides a better performance because it provides a higher noise attenuation, as a results of the higher stopband attenuation and the narrower transition bands, compared to the FIR counterpart. Most importantly, the proposed 5-D IIR depth-velocity filter requires *less than* 1% *of the arithmetic operations* required by the 5-D FIR depth-velocity filter to process a sample. Considering the ultra-low complexity, the proposed 5-D IIR depth-velocity filter has a significant potential to be employed in *real-time* applications.

In the case of the Lytro-LF-camera-based LFV, the output LFV of the 5-D IIR depth-velocity filter has less brightness compared to the input LFV because the output LFV has not reached to the steady state with respect to the n_x and n_y dimensions. This effect may be overcome by employing a post-processing technique such as histogram equalization or by increasing the number of samples (or sub-apertures) available for the n_x and n_y dimensions, a feature which is expected to be available with the future generations of LFV cameras. Furthermore, the passbands of the 5-D IIR filters $H_{xu}(\mathbf{z})$, $H_{yv}(\mathbf{z})$ and $H_{uvt}(\mathbf{z})$ deviate from the corresponding desired 4-D hyperplanes at higher frequencies due to the *frequency warping* of the bilinear transform. Even though the effect of this drawback may be negligible for most of natural scenes, in some cases, the attenuation of some interfering objects may be poor.

Chapter 5

Novel Ultra-Low Complexity 5-D IIR Adaptive Depth-Velocity Filter

5.1 Introduction

The spectrum analysis presented in Chapter 2 and the 5-D FIR and IIR depth-velocity filters presented in Chapter 3 and Chapter 4, respectively, considered a *special case* of the enhancement of moving objects in LFVs, i.e., the objects moving with constant velocity and at constant depth. In this chapter, we extend the 5-D depth-velocity filtering technique to a more general case, where objects moving with constant velocity but at *non-constant depth* are enhanced.

First, the spectrum of a Lambertian object moving in an LFV with nonzero V_z is analyzed, and it is shown that the ROS of the spectrum can be approximated as a sequence of ROSs, each of which is a skewed 3-D hyperfan, in the 5-D continuous frequency domain. Based on this analysis, a novel ultra-low complexity 5-D IIR *adaptive* depth-velocity filter is proposed for enhancing such moving objects. The proposed 5-D IIR adaptive depth-velocity filter is realized by cascading three first-order 5-D IIR adaptive filters having *time-variant* 4-D hyperplanar passbands of appropriate orientations. The first-order 5-D IIR adaptive filters are designed by appropriately extending the first-order 3-D IIR adaptive planar filter design method proposed in [25]. The time-variant coefficients of the three first-order 5-D IIR adaptive filters are derived in *closed form*. The performance of the proposed 5-D IIR adaptive depth-velocity filter is tested by employing a numerically generated LFV and an LFV of a real scene, generated by means of a commercially available Lytro LF camera.

Experimental results indicate that the proposed 5-D IIR adaptive depth-velocity filter outperforms the 3-D IIR adaptive velocity filters [25] and the 4-D IIR adaptive depth filters in enhancing moving objects in LFVs. Considering the ultra-low complexity and the availability of the closed-form expressions for the time-variant coefficients, the proposed 5-D IIR adaptive depth-velocity filter has a significant potential to be employed in *real-time* applications.

The rest of the chapter is organized as follows. In Section 5.2, the analysis of the spectrum of a Lambertian object moving in an LFV at non-constant depth is presented. The 5-D IIR adaptive depth-velocity filter is described in Section 5.3. In Section 5.4, numerical simulation results are presented. Finally, a summary of the chapter is presented in Section 5.5.

5.2 Analysis of the Spectrum of a Lambertian Object Moving at Non-Constant Depth

The spectrum of a Lambertian object moving in an LFV at non-constant depth (or with nonzero V_z) is analyzed in this section. To this end, an approximate representation of a Lambertian point source moving at non-constant depth in an LFV is proposed in the next subsection.

5.2.1 Approximate Representation of a Lambertian Point Source Moving at Non-Constant Depth Using Time-Variant 3-D Hyperplanes

Recall that a Lambertian point source of intensity l_0 moving with the constant velocity $\mathbf{V} = [V_x, V_y, V_z]^{\mathrm{T}}$ is represented in a 5-D continuous-domain LFV $l_{5\mathrm{C}}(\mathbf{x})$ as a 3-D hypersurface of constant value l_0, which is given by the intersection of the two 4-D hypersurfaces

$$mx + u + k_x t - k_z u t + c_x = 0 \tag{5.1a}$$

$$my + v + k_y t - k_z v t + c_y = 0, \tag{5.1b}$$

where m, c_x and c_y are given in (2.2a), (2.2b) and (2.2c), respectively, and k_x, k_y and k_z are given in (2.13a), (2.13b) and (2.13c), respectively. In this case, $l_{5\mathrm{C}}(\mathbf{x})$ is given

by

$$l_{5\mathrm{C}}(\mathbf{x}) = l_0\,\delta(mx + u + k_x t - k_z u t + c_x) \times \delta(my + v + k_y t - k_z v t + c_y). \tag{5.2}$$

Obtaining a closed-form expression for the spectrum of $l_{5\mathrm{C}}(\mathbf{x})$, when $k_z \neq 0$ (or $V_z \neq 0$), is highly complicated due to the nonlinear nature of the arguments of the $\delta(\cdot)$ function. *Motivated from the LFV representation of Lambertian point sources moving at constant depth (with $V_z = 0$)* presented in Chapter 2, where a *closed-form* expression for the spectrum can be obtained, we propose an approximate representation for a Lambertian point source moving with nonzero V_z in this subsection. More specifically, the two 4-D hypersurfaces given by (5.1a) and (5.1b) are approximated with the two *time-variant* 4-D hyperplanes

$$m(t)x + u + k_x(t)t + c_x(t) = 0 \tag{5.3a}$$

$$m(t)y + v + k_y(t)t + c_y(t) = 0, \tag{5.3b}$$

respectively, where the parameters $m(t)$, $k_x(t)$, $k_y(t)$, $c_x(t)$ and $c_y(t)$ are restricted to be *piece-wise constant* with respect to the time t. For a short time interval $t_0 \in [t_i, t_f]$, for which the parameters $m(t)$, $k_x(t)$, $k_y(t)$, $c_x(t)$ and $c_y(t)$ are constant, they can be obtained as

$$m(t_0) = \frac{D}{z_0(t_i)} \tag{5.4a}$$

$$k_x(t_0) = \frac{-DV_x}{z_0(t_i)} \tag{5.4b}$$

$$k_y(t_0) = \frac{-DV_y}{z_0(t_i)} \tag{5.4c}$$

$$c_x(t_0) = \frac{-Dx_0(t_i)}{z_0(t_i)} \tag{5.4d}$$

$$c_y(t_0) = \frac{-Dy_0(t_i)}{z_0(t_i)}, \tag{5.4e}$$

respectively, where $(x_0(t_i), y_0(t_i), z_0(t_i))$ is the position of the Lambertian point source at $t_0 = t_i$. In fact, in this case, the LFV $l_{5\mathrm{C}}(\mathbf{x})$ is *locally* approximated for the short time interval t_0 with the LFV corresponding to a Lambertian point source moving at

the constant depth $z_0(t_i)$, i.e.,

$$l_{5\mathrm{C}}(\mathbf{x}) \approx l_0\,\delta(m(t_0)x + u + k_x(t_0)t + c_x(t_0)) \times \delta(m(t_0)y + v + k_y(t_0)t + c_y(t_0)). \tag{5.5}$$

Furthermore, for $t \in \mathbb{R}$, the LFV $l_{5\mathrm{C}}(\mathbf{x})$ is approximated with a *sequence* of LFVs, each corresponding to a Lambertian point source moving at a constant depth for a short time interval.

5.2.2 Approximate Spectrum of a Lambertian Object Moving at Non-Constant Depth and Its ROS

Following the spectral analysis presented in Chapter 2, the time-variant spectrum $L_{5\mathrm{C}}(\mathbf{\Omega}, t)$ can be approximated, for the short time interval t_0, as

$$L_{5\mathrm{C}}(\mathbf{\Omega}, t_0) \approx 8\pi^3 l_0\,\delta(\Omega_x - m(t_0)\Omega_u)\,\delta(\Omega_y - m(t_0)\Omega_v) \times \delta(\Omega_t - k_x(t_0)\Omega_u - k_y(t_0)\Omega_v)\,\mathrm{e}^{j(\Omega_u c_x(t_0) + \Omega_v c_y(t_0))}. \tag{5.6}$$

Furthermore, the time-variant ROS of the spectrum, $\mathcal{P}_{5\mathrm{C}}(t)$, can be obtained as

$$\mathcal{P}_{5\mathrm{C}}(t_0) \approx \mathcal{H}_{5\mathrm{C},xu}(t_0) \cap \mathcal{H}_{5\mathrm{C},yv}(t_0) \cap \mathcal{H}_{5\mathrm{C},uvt}(t_0), \tag{5.7}$$

where

$$\mathcal{H}_{5\mathrm{C},xu}(t_0) = \{\mathbf{\Omega} \in \mathbb{R}^5 \mid \Omega_x - m(t_0)\Omega_u = 0\} \tag{5.8a}$$

$$\mathcal{H}_{5\mathrm{C},yv}(t_0) = \{\mathbf{\Omega} \in \mathbb{R}^5 \mid \Omega_y - m(t_0)\Omega_v = 0\} \tag{5.8b}$$

$$\mathcal{H}_{5\mathrm{C},uvt}(t_0) = \{\mathbf{\Omega} \in \mathbb{R}^5 \mid \Omega_t - k_x(t_0)\Omega_u - k_y(t_0)\Omega_v = 0\}. \tag{5.8c}$$

Note that the ROS $\mathcal{P}_{5\mathrm{C}}(t)$ is a time-variant plane through the origin in the 5-D continuous frequency domain, which is given by the intersection of the three time-variant 4-D hyperplanes

$$\Omega_x - m(t)\Omega_u = 0 \tag{5.9a}$$

$$\Omega_y - m(t)\Omega_v = 0 \tag{5.9b}$$

$$\Omega_t - k_x(t)\Omega_u - k_y(t)\Omega_v = 0 \tag{5.9c}$$

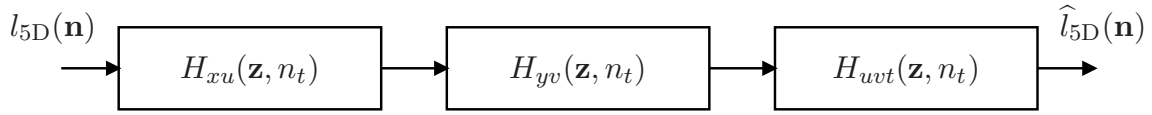


Figure 5.1: Structure of the proposed 5D IIR adaptive depth-velocity filter.

having normal vectors $\mathbf{d}_{5\mathrm{C},xu}(t) = [1, 0, -m(t), 0, 0]^{\mathrm{T}}$, $\mathbf{d}_{5\mathrm{C},yv}(t) = [0, 1, 0, -m(t), 0]^{\mathrm{T}}$ and $\mathbf{d}_{5\mathrm{C},uvt}(t) = [0, 0, -k_x(t), -k_y(t), 1]^{\mathrm{T}}$, respectively, which are piece-wise constant with respect to the time t.

Similar to Chapter 2, a Lambertian object moving with the constant velocity $\mathbf{V} = [V_x, V_y, V_z]^{\mathrm{T}}$ is considered as a collection of Lambertian point sources moving with the same velocity. Consequently, the LFV of the Lambertian object is given by the superposition of the LFVs of the corresponding Lambertian point sources. In this case, the LFV of the Lambertian object is approximated with a sequence of LFVs, each corresponding to the Lambertian object moving at a constant depth for a short time interval. Therefore, the time-variant ROS of the spectrum, $\mathcal{O}_{5\mathrm{C}}(t)$, is obtained as a sequence of ROSs. For the short time interval t_0, $\mathcal{O}_{5\mathrm{C}}(t_0)$ is given by

$$\begin{aligned}\mathcal{O}_{5\mathrm{C}}(t_0) &\approx \bigcup_{z_0(t_i)} \mathcal{P}_{5\mathrm{C}}(t_0) \\ &\approx \bigcup_{z_0(t_i)} \left(\mathcal{H}_{5\mathrm{C},xu}(t_0) \cap \mathcal{H}_{5\mathrm{C},yv}(t_0) \cap \mathcal{H}_{5\mathrm{C},uvt}(t_0)\right), \end{aligned} \tag{5.10}$$

where $z_0(t_i)$ is the depth range occupied by the Lambertian object at $t_0 = t_i$. Due to the time-variant nature of the ROS of the spectrum, in this case, the passband (or stopband) of a 5-D depth-velocity filter needs to be appropriately tuned with respect to time in order to enhance (or attenuate) a Lambertian object, i.e, the 5-D depth-velocity filter should be *adaptive*.

5.3 Design of the 5-D IIR Adaptive Depth-Velocity Filter

The design of the 5-D IIR adaptive depth-velocity filter is described in this section. The structure of the proposed 5-D IIR adaptive depth-velocity filter $H(\mathbf{z}, n_t)$ is shown in Figure 5.1. Similar to the design of the 5-D IIR depth-velocity filter presented in Chapter 4, the 5-D IIR adaptive depth-velocity filter is designed as a cascade of three

5-D IIR adaptive filters; $H_{xu}(\mathbf{z}, n_t)$, $H_{yv}(\mathbf{z}, n_t)$ and $H_{uvt}(\mathbf{z}, n_t)$. The transfer function $H(\mathbf{z}, n_t)$ of the 5-D IIR adaptive depth-velocity filter is given by

$$H(\mathbf{z}, n_t) = H_{xu}(\mathbf{z}, n_t)\, H_{yv}(\mathbf{z}, n_t)\, H_{uvt}(\mathbf{z}, n_t). \tag{5.11}$$

The coefficients of each 5-D IIR adaptive filter are adapted with respect to the time n_t such that the passband of $H(\mathbf{z}, n_t)$ encompasses the instantaneous (or local) spectral ROS of the object of interest. Note that, for a time instant n_{t0}, the passbands of the 5-D IIR adaptive filters $H_{xu}(\mathbf{z}, n_t)$, $H_{yv}(\mathbf{z}, n_t)$ and $H_{uvt}(\mathbf{z}, n_t)$ encompass the 5-D hypervolumes given by

$$\mathcal{B}_{5\mathrm{D},xu}(n_{t0}) = \bigcup_{z_0(n_{t0})} \mathcal{H}_{5\mathrm{D},xu}(n_{t0}) \tag{5.12a}$$

$$\mathcal{B}_{5\mathrm{D},yv}(n_{t0}) = \bigcup_{z_0(n_{t0})} \mathcal{H}_{5\mathrm{D},yv}(n_{t0}) \tag{5.12b}$$

$$\mathcal{B}_{5\mathrm{D},uvt}(n_{t0}) = \bigcup_{z_0(n_{t0})} \mathcal{H}_{5\mathrm{D},uvt}(n_{t0}), \tag{5.12c}$$

respectively, inside the 5-D principal Nyquist hypercube $\mathcal{N}_{5\mathrm{D}}$, where $z_0(n_{t0})$ is the depth range occupied by the Lambertian object at $n_t = n_{t0}$. Moreover, $\mathcal{H}_{5\mathrm{D},xu}(n_{t0})$, $\mathcal{H}_{5\mathrm{D},yv}(n_{t0})$ and $\mathcal{H}_{5\mathrm{D},uvt}(n_{t0})$ are the discrete-domain counterparts of the 4-D hyperplanes given by (5.8a), (5.8b) and (5.8c), respectively. Furthermore, in this case, the overall passband $\mathcal{B}_{5\mathrm{D}}(n_{t0})$ of $H(\mathbf{z}, n_t)$, given by the intersection of $\mathcal{B}_{5\mathrm{D},xu}(n_{t0})$, $\mathcal{B}_{5\mathrm{D},yv}(n_{t0})$ and $\mathcal{B}_{5\mathrm{D},uvt}(n_{t0})$, completely encompasses the approximated 3-D hyperfan-shaped spectral ROS of the object of interest, which corresponds to the time instant n_{t0}.

The design of the 5-D IIR adaptive filters $H_{xu}(\mathbf{z}, n_t)$, $H_{yv}(\mathbf{z}, n_t)$ and $H_{uvt}(\mathbf{z}, n_t)$ are carried out by appropriately extending the design method reported for the first-order 3-D IIR adaptive planar filters [25]. The time-variant difference equations for $H_{xu}(\mathbf{z}, n_t)$, $H_{yv}(\mathbf{z}, n_t)$ and $H_{uvt}(\mathbf{z}t, n_t)$ are, respectively, obtained as

$$l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}) = \frac{1}{b_{00}(n_t)} \left[\sum_{i_x=0}^{1} \sum_{i_u=0}^{1} l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n} - \mathbf{i}_{xu}) - \sum_{\substack{i_x=0 \\ i_x+i_u\neq 0}}^{1} \sum_{i_u=0}^{1} b_{i_x i_u}(n_t)\, l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n} - \mathbf{i}_{xu}) \right] \tag{5.13a}$$

$$l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}) = \frac{1}{b_{00}(n_t)} \left[\sum_{i_y=0}^{1} \sum_{i_v=0}^{1} l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n} - \mathbf{i}_{yv}) - \sum_{\substack{i_x=0 \\ i_y+i_v\neq 0}}^{1} \sum_{i_u=0}^{1} b_{i_y i_v}(n_t)\, l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n} - \mathbf{i}_{yv}) \right] \tag{5.13b}$$

$$l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}) = \frac{1}{b_{000}(n_t)}\left[\sum_{i_u=0}^{1}\sum_{i_v=0}^{1}\sum_{i_t=0}^{1} l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n}-\mathbf{i}_{uvt}) - \sum_{\substack{i_u=0\\ i_u+i_v+i_t\neq 0}}^{1}\sum_{i_v=0}^{1}\sum_{i_t=0}^{1} b_{i_u i_v i_t}(n_t)\, l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n}-\mathbf{i}_{uvt})\right], \tag{5.13c}$$

where $l_{5\mathrm{D}}^{\mathrm{in}}(\mathbf{n})$ and $l_{5\mathrm{D}}^{\mathrm{out}}(\mathbf{n})$ are the input and the output signals of each filter, respectively. Further, $\mathbf{i}_{xu} = [i_x, 0, i_u, 0, 0]^{\mathrm{T}}$, $\mathbf{i}_{yv} = [0, i_y, 0, i_v, 0]^{\mathrm{T}}$ and $\mathbf{i}_{uvt} = [0, 0, i_u, i_v, i_t]^{\mathrm{T}}$. The time-variant denominator coefficients of the three adaptive filters $H_{xu}(\mathbf{z}, n_t)$, $H_{yv}(\mathbf{z}, n_t)$ and $H_{uvt}(\mathbf{z}, n_t)$ are, respectively, obtained as [24] [25]

$$b_{i_x i_u}(n_t) = 1 + \frac{(-1)^{i_x} d_{xu}^{x}(n_t) + (-1)^{i_u} d_{xu}^{u}(n_t)}{B_{xu}}, \qquad i_x, i_u = 0, 1 \tag{5.14a}$$

$$b_{i_y i_v}(n_t) = 1 + \frac{(-1)^{i_y} d_{yv}^{y}(n_t) + (-1)^{i_v} d_{yv}^{v}(n_t)}{B_{yv}}, \qquad i_y, i_v = 0, 1 \tag{5.14b}$$

$$b_{i_u i_v i_t}(n_t) = 1 + \frac{(-1)^{i_u} d_{uvt}^{u}(n_t) + (-1)^{i_v} d_{uvt}^{v}(n_t) + (-1)^{i_t} d_{uvt}^{t}(n_t)}{B_{uvt}}, \quad i_u, i_v, i_t = 0, 1, \tag{5.14c}$$

where B_{xu}, B_{yv} and B_{uvt} are the *time-invariant* -3 dB bandwidths of the three filters, respectively. Furthermore, $\widehat{\mathbf{d}}_{xu}(n_t) = [d_{xu}^{x}(n_t), 0, d_{xu}^{u}(n_t), 0, 0]^{\mathrm{T}}$, $\widehat{\mathbf{d}}_{yv}(n_t) = [0, d_{yv}^{y}(n_t), 0, d_{yv}^{v}(n_t), 0]^{\mathrm{T}}$ and $\widehat{\mathbf{d}}_{uvt}(n_t) = [0, 0, d_{uvt}^{u}(n_t), d_{uvt}^{v}(n_t), d_{uvt}^{t}(n_t)]^{\mathrm{T}}$ are the time-variant unit normal vectors corresponding to the time-variant 4-D hyperplanar passbands, respectively.

Because the depth of the object of interest, $z_0(n_t)$, is always positive and the velocity of the object of interest is constant, the signs of the time-variant parameters $m(n_t)$, $k_x(n_t)$ and $k_y(n_t)$, which determine the unit normal vectors $\widehat{\mathbf{d}}_{xu}(n_t)$, $\widehat{\mathbf{d}}_{yv}(n_t)$ and $\widehat{\mathbf{d}}_{uvt}(n_t)$, do not change with the time n_t, i.e., positive parameters remain positive and negative parameters remain negative for all the time. Consequently, the unit normal vectors $\widehat{\mathbf{d}}_{xu}(n_t)$, $\widehat{\mathbf{d}}_{yv}(n_t)$ and $\widehat{\mathbf{d}}_{uvt}(n_t)$ always lie inside the closed orthants corresponding to $\widehat{\mathbf{d}}_{xu}(0)$, $\widehat{\mathbf{d}}_{yv}(0)$ and $\widehat{\mathbf{d}}_{uvt}(0)$. That is, the unit normal vectors always lie inside their initial closed orthants. Therefore, the 5-D IIR adaptive filters $H_{xu}(\mathbf{z}, n_t)$, $H_{yv}(\mathbf{z}, n_t)$ and $H_{uvt}(\mathbf{z}, n_t)$ can be implemented so that the practical-BIBO stability [101] is always guaranteed.

Object type	Size height × width (cm^2)	Initial Depth $z_0(0)$ (cm)	True velocity $[V_x, V_y, V_z]$ (cm/frame)
Object of interest	50×50	50	$[1, 1, 0.3]$
Interfering object 1	50×50	50	$[1, -2, 0.3]$
Interfering object 2	10×10	10	$[0.2, 0.2, 0.06]$

Table 5.1: Specifications of the Lambertian planar objects.

5.4 Experimental Results

In this section, experimental results are presented in order to confirm the effectiveness of the proposed 5-D IIR adaptive depth-velocity filter in enhancing objects moving with constant velocity and at non-constant depth in LFVs based on their depth and velocity. First, by employing a numerically generated LFV, it is shown that the proposed 5-D IIR adaptive depth-velocity filter outperforms the 3-D IIR adaptive velocity filters [25] and the 4-D IIR depth filters. Next, the experimental results obtained using an LFV generated using the Lytro LF camera are presented.

5.4.1 Numerically Generated LFV

Similar to the numerically generated LFV employed in Chapters 3 and 4, the numerically generated LFV employed in this chapter contains three Lambertian planar objects. All the three planar objects have a Lambertian intensity pattern that resembles a checkerboard of 8×8 squares. The other specifications of the three planar objects are presented in Table 5.1. Note that the distance D between the camera plane and the image plane is selected as 50 cm. Note that the object of interest and the first interfering object always move at the same depth so that the first interfering object can not be attenuated using 4-D adaptive depth filters. On the other hand, the object of interest and the second interfering object always move with the same apparent velocity, i.e. both have the same $k_x(n_t)$ and $k_y(n_t)$, so that the second interfering object can not be attenuated using 3-D adaptive velocity filters. The generated LFV is of the size $9 \times 9 \times 256 \times 256 \times 64$, and the 10th frame of the LFV corresponding to the central 25 sub-apertures ($n_x, n_y = 3, 4, \ldots, 7$) is shown in Figure 5.2.

The 5-D IIR adaptive filters $H_{xu}(\mathbf{z}, n_t)$ and $H_{yv}(\mathbf{z}, n_t)$ are designed with the pa-

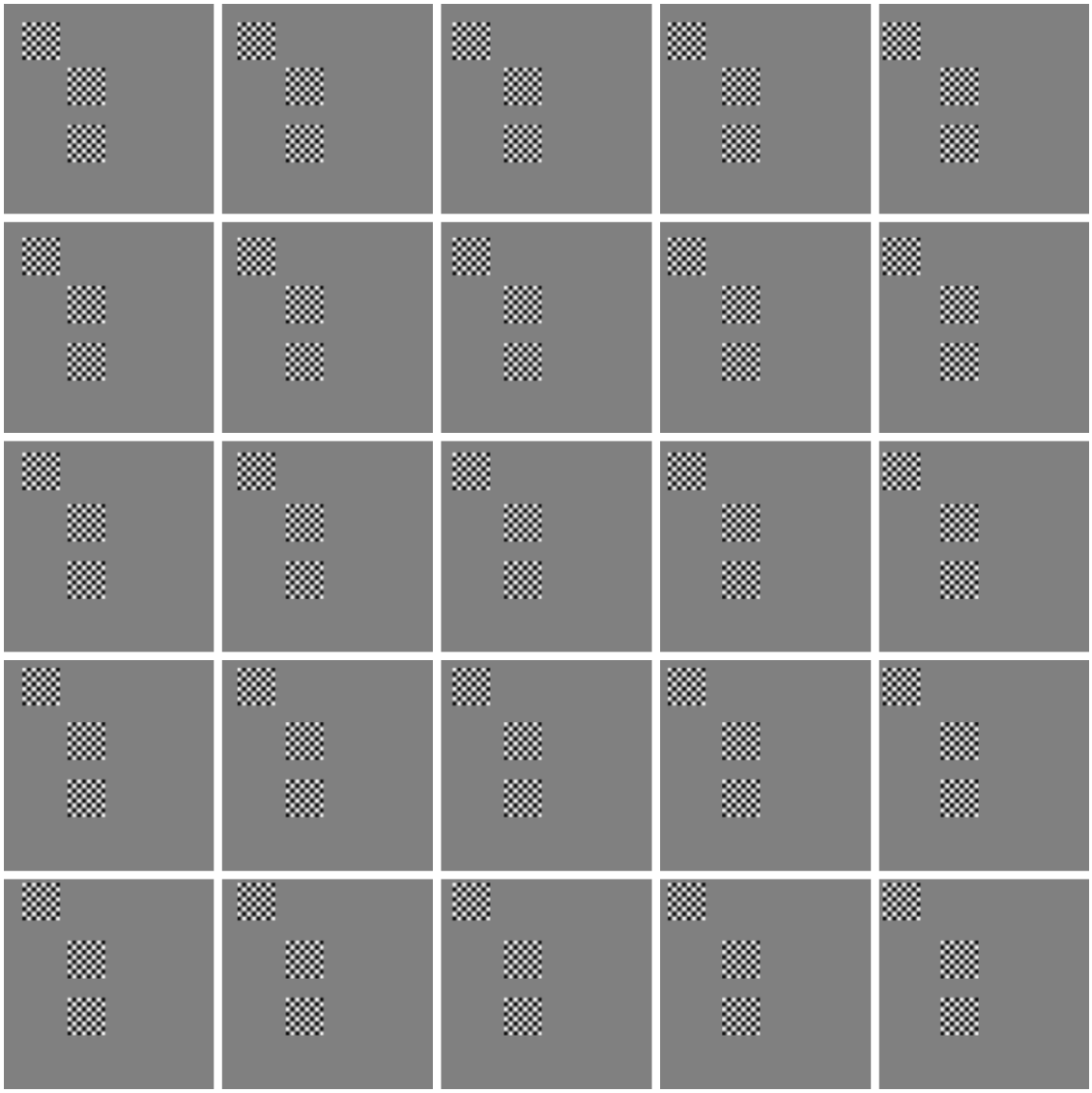


Figure 5.2: The 10th frame of the numerically generated LFV corresponding to the central 25 sub-apertures ($n_x, n_y = 3, 4, \ldots, 7$). The middle object is the object of interest; the bottom object is the interfering object 1 moving at the same depth (but with different apparent velocity) of the object of interest; the top object is the interfering object 2 moving with the same apparent velocity (but at different depth) of the object of interest.

rameters $B_{xu} = B_{yv} = 0.1\pi$ rad/sample. The magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 50$ is shown in Figure 5.3. The magnitude response of $H_{yv}(\mathbf{z}, n_t)$ at $n_t = 50$ is similar to that of $H_{xu}(\mathbf{z}, n_t)$ shown in Figure 5.3. The filter $H_{uvt}(\mathbf{z}, n_t)$ is designed with the parameter $B_{uvt} = 0.03\pi$ rad/sample. The -3 dB iso-surface of the magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 50$ is shown in Figure 5.4.

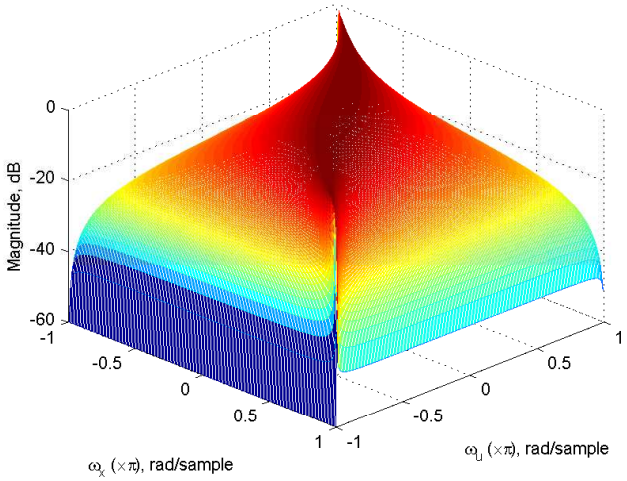


Figure 5.3: Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 50$ in the $\omega_x\omega_u$ plane.

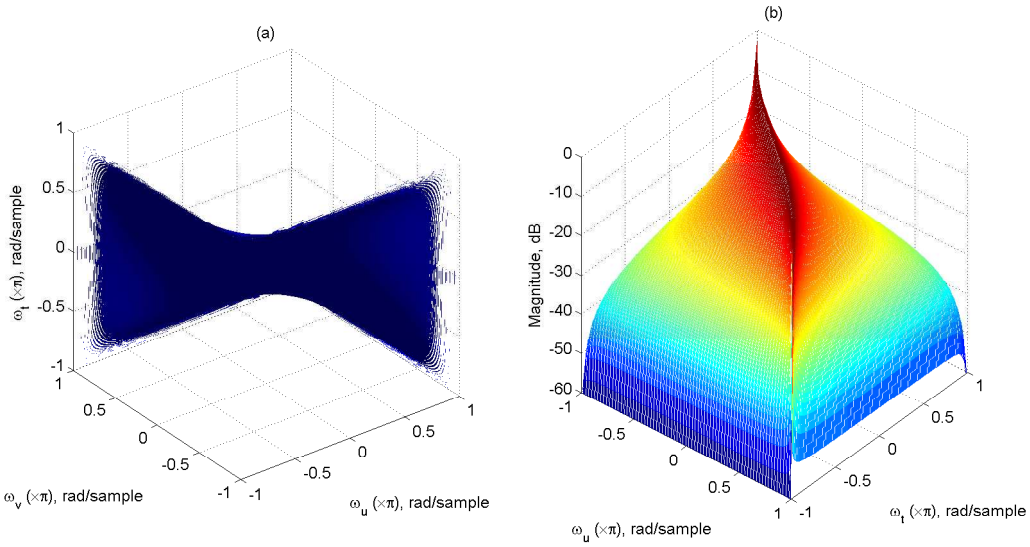


Figure 5.4: Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 50$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$.

The numerically generated LFV is filtered with the proposed 5-D IIR adaptive depth-velocity filter, a 4-D IIR adaptive depth filter and the 3-D IIR adaptive velocity filter [25]. The 4-D IIR adaptive depth filter was designed by appropriately extending the 4-D IIR time-invariant depth filter presented in [44]. Note that the coefficients of all the three filters were adapted for each temporal frame. The 5-D IIR adaptive depth-velocity filter $H(\mathbf{z}, n_t)$ was implemented using the time-variant difference equations given by (5.13a)–(5.13c). The 10th frame of the original, the 4-D

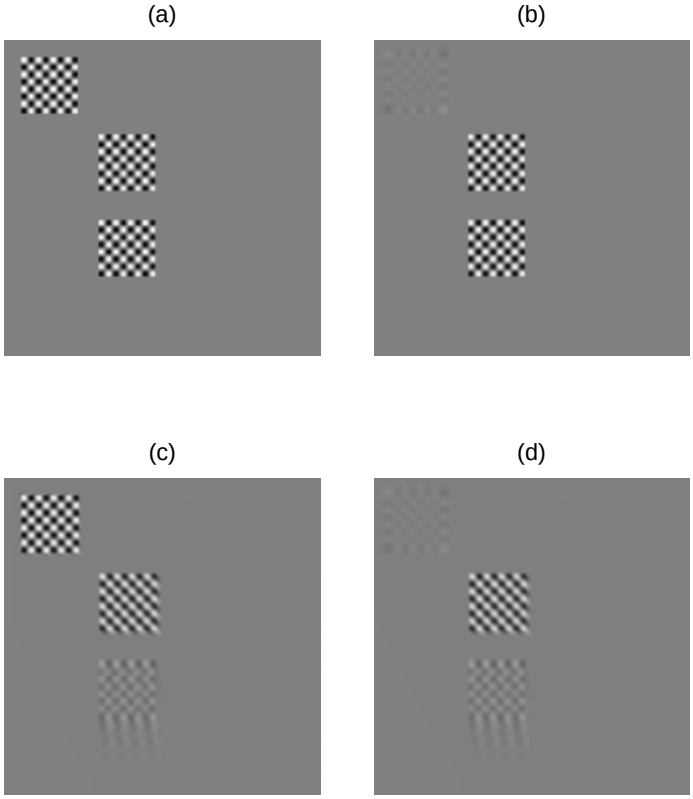


Figure 5.5: The 10th frame of (a) the original, (b) the 4-D IIR adaptive depth filtered (c) the 3-D IIR adaptive velocity filtered [25] and (d) the 5-D IIR adaptive depth-velocity filtered (proposed) LFVs corresponding to the central sub-aperture ($n_x, n_y = 5$). The middle object is the object of interest; the top object is the interfering object moving at the same apparent velocity (but at different depth) of the object of interest; the bottom object is the interfering object moving at the same depth (but with different apparent velocity) of the object of interest.

IIR adaptive depth filtered, the 3-D IIR adaptive velocity filtered [25] and the 5-D IIR adaptive depth-velocity filtered LFVs corresponding to the central sub-aperture ($n_x, n_y = 5$) are shown in Figure 5.5. It can be clearly observed that the first interfering object (bottom object) moving at the same depth of the object of interest (middle object) is not attenuated by the 4-D IIR adaptive depth filter. Similarly, the second interfering object (top object) moving with the same apparent velocity of the object of interest is not attenuated by the 3-D IIR adaptive velocity filter. On the other hand, both interfering objects are attenuated by the 5-D IIR adaptive depth-velocity filter. This confirms the effectiveness of the proposed 5-D IIR adaptive depth-velocity filters, compared to the 4-D IIR adaptive depth filters and 3-D IIR adaptive velocity filters [25], in enhancing objects moving at non-constant depth in LFVs.

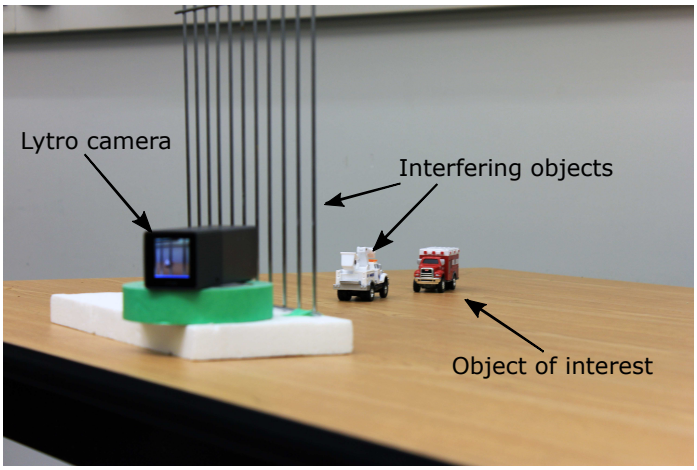


Figure 5.6: The experimental setup employed to generate the LFV of a real scene. A Lytro LF camera was employed to capture the individual frames of the scene. The red truck is the object of interest whereas the white truck and the fence are a moving interfering object and a static interfering object, respectively.

5.4.2 Lytro-LF-Camera-Based LFV

The experimental setup employed to generate the LFV is shown in Figure 5.6. It is comprised of a red truck (object of interest), a white truck (moving interfering object) and a fence (static interfering object). The red truck was manually moved towards the Lytro LF camera while the white truck was manually moved outwards from the Lytro LF camera. Fifty six LFs were captured, and each LF was decoded by using the MATLAB LFToolbox (Version 0.3) [93]. The generated LFV is of the size $9 \times 9 \times 256 \times 360 \times 56$. The 40th frame of the LFV corresponding to the central 15 sub-apertures ($n_x = 4, 5, 6$ and $n_y = 3, 4, \ldots, 7$) is shown in Figure 5.7.

The time-variant parameters $m(n_t)$, $k_x(n_t)$ and $k_y(n_t)$ corresponding to the object of interest and the interfering objects are derived based on the parameterization of the decoded LFV [93]. When n_t varies from 0 to 55, $m(n_t)$, $k_x(n_t)$ and $k_y(n_t)$ corresponding to the object of interest vary from -2.8 to -3.2, -0.3 to -1, and -0.4 to -1.2, respectively. The variations of these time-variant parameters corresponding to the moving interfering object are -3.2 to -2.8, 1 to 0.3, and 1.2 to 0.4, respectively. Furthermore, $m(n_t)$, $k_x(n_t)$ and $k_y(n_t)$ corresponding to the static interfering object are -3.3, 0, and 0, respectively. Note that the time-variant parameters corresponding to the object of interest are close to those of the interfering objects. Consequently, the spectral ROSs of the three objects lie close to each other, and the enhancement of the object of interest using the 5-D IIR adaptive depth-velocity filter is challenging in

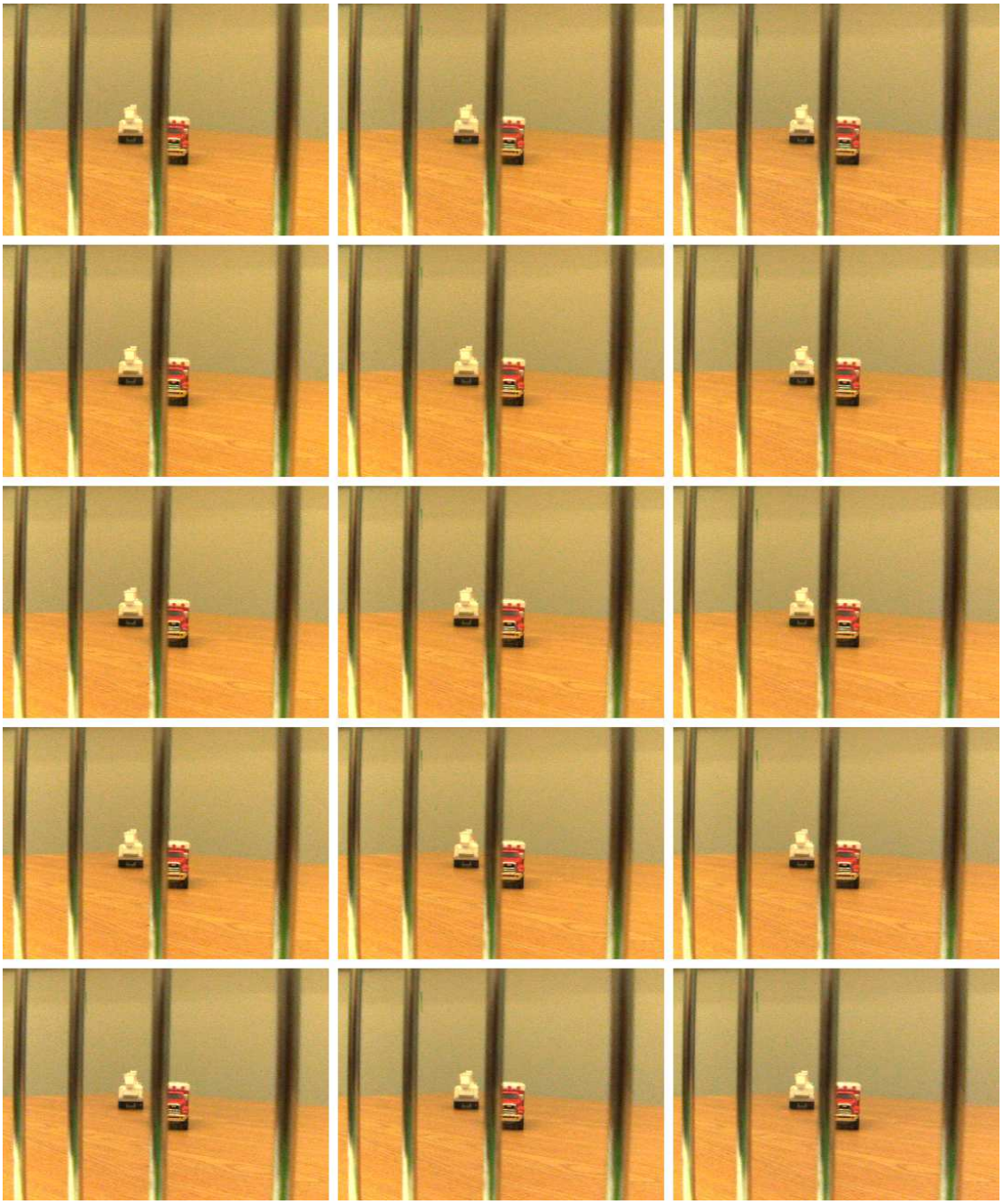


Figure 5.7: The 40th frame of the LFV corresponding to the central 15 sub-apertures ($n_x = 4, 5, 6$ and $n_y = 3, 4, \ldots, 7$). The red truck is the object of interest; the white truck is the moving interfering object; the fence is the static interfering object.

this case. The 5-D IIR adaptive filters $H_{xu}(\mathbf{z}, n_t)$ and $H_{yv}(\mathbf{z}, n_t)$ are designed with the parameters $B_{xu} = B_{yv} = 0.08\pi$ rad/sample. The magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$

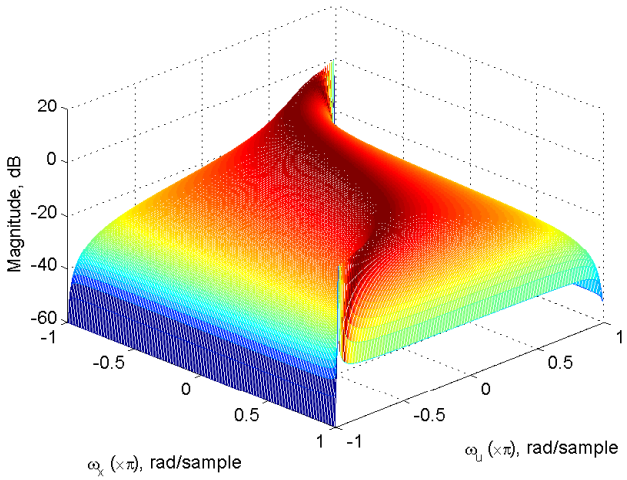


Figure 5.8: Magnitude response $|H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 30$ in the $\omega_x\omega_u$ plane.

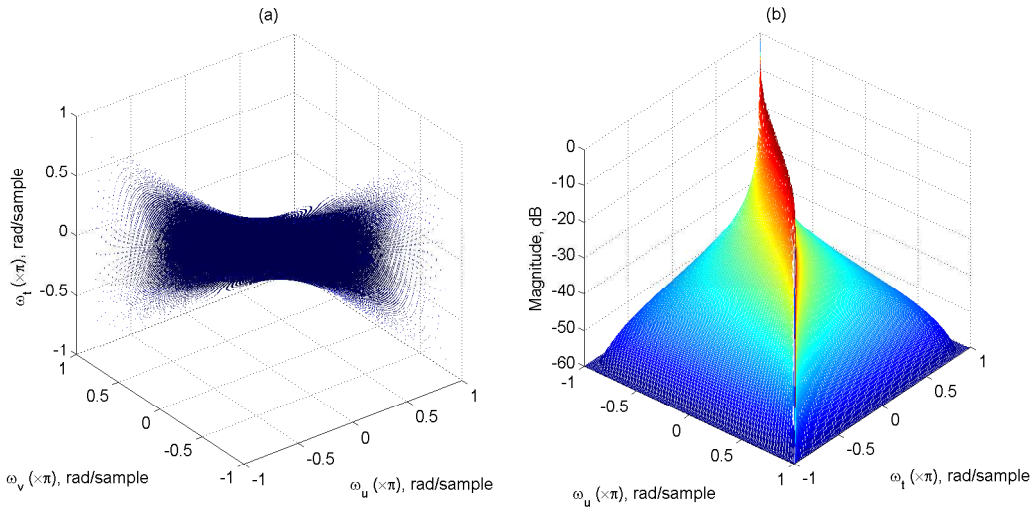


Figure 5.9: Magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 30$ (a) -3 dB iso-surface in the $\omega_u\omega_v\omega_t$ space (b) cross section obtained at $\omega_v = 0$.

obtained at $n_t = 30$ is shown in Figure 5.8. The magnitude response of $H_{yv}(\mathbf{z}, n_t)$ at $n_t = 30$ is similar to that of $H_{xu}(\mathbf{z}, n_t)$ shown in Figure 5.8. The filter $H_{uvt}(\mathbf{z}, n_t)$ is designed with the parameter $B_{uvt} = 0.003\pi$ rad/sample. The -3 dB iso-surface of the magnitude response $|H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}, n_t)|$ obtained at $n_t = 30$ is shown in Figure 5.9.

The LFV is filtered with the proposed 5-D IIR adaptive depth-velocity filter, and three frames of the original and the filtered LFVs, corresponding to the sub-aperture $(n_x, n_y) = (7, 7)$, are shown in Figure 5.10. It is observed that the proposed 5-D IIR depth-velocity filter enhances the object of interest (red truck) with slight

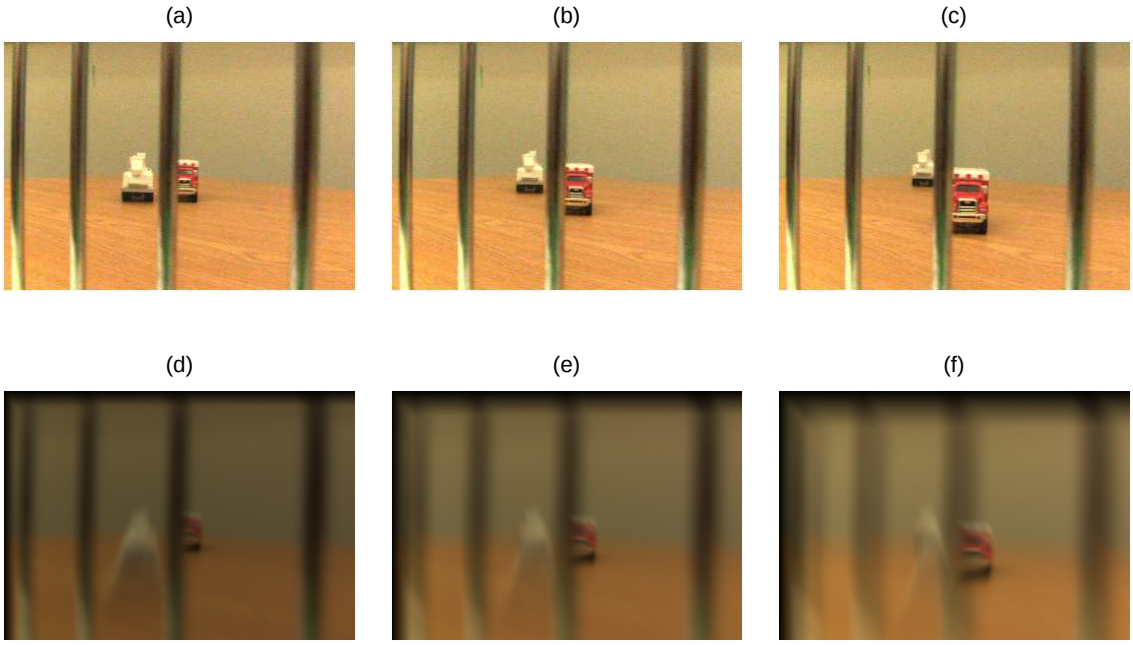


Figure 5.10: Three frames of the original LFV (top row) and the 5-D IIR adaptive depth-velocity filtered LFV (bottom row), $(n_x, n_y) = (7, 7)$; (a) and (d) 30th frame; (b) and (e) 40th frame; (c) and (f) 50th frame.

distortion while substantially attenuating the moving interfering object (white truck) and modestly attenuating the static interfering object (fence). *These results confirm the effectiveness of the proposed 5-D IIR adaptive depth-velocity filter in enhancing objects moving at non-constant depth in LFVs.*

In the filtered LFV, the attenuation of the two interfering objects is higher in the 50th frame, compared to that in the 30th frame. This results as a consequence of the increase of the difference between the apparent velocity of the object of interest and those of the two interfering objects when n_t varies from 0 to 55. Furthermore, it is worthwhile to note the attenuation of the interfering objects can be increased by reducing the parameters B_{xu} and B_{yv} employed in the design of the 5-D IIR adaptive filters $H_{xu}(\mathbf{z}, n_t)$ and $H_{yv}(\mathbf{z}, n_t)$, respectively. However, in this case, the number of samples available for the n_x and n_y should be substantially higher than the number of samples available with the LFV employed in the experiment in order to overcome the long transient responses.

5.5 Summary

In this chapter, the 5-D depth-velocity filtering technique is extended to the more general case, where objects moving with constant velocity but at non-constant depth are enhanced. First, the spectrum of a Lambertian object moving in an LFV at non-constant depth is analyzed. In particular, an approximate representation for a Lambertian point source moving at non-constant depth is presented. In this case, the two 4-D hypersurfaces corresponding to the LFV representation of the Lambertian point source are approximated with two 4-D time-variant hyperplanes. Furthermore, the ROS of the spectrum of a Lambertian object moving at non-constant depth is approximated as a sequence of ROSs, each of which is a skewed 3-D hyperfan, in the 5-D continuous frequency domain.

Based on the spectral analysis, a novel ultra-low complexity 5-D IIR adaptive depth-velocity filter is proposed for enhancing objects moving at non-constant depth. The proposed 5-D IIR adaptive depth-velocity filter is realized by cascading three first-order 5-D IIR adaptive filters having time-variant 4-D hyperplanar passbands of appropriate orientations. The three first-order 5-D IIR adaptive filters are designed by appropriately extending the first-order 3-D IIR adaptive planar filter design method presented in [25]. The time-variant coefficients of the three first-order 5-D IIR adaptive filters are derived in closed form. The proposed 5-D IIR adaptive depth-velocity filter is practical-BIBO stable.

The performance of the proposed 5-D IIR adaptive depth-velocity filter is confirmed by employing a numerically generated LFV and an LFV of a real scene, generated by means of a commercially available Lytro LF camera. Experimental results indicate that the proposed 5-D IIR adaptive depth-velocity filter outperforms the 3-D IIR adaptive velocity filters [25] and the 4-D IIR adaptive depth filters in enhancing moving objects in LFVs. The performance of the 5-D IIR adaptive depth-velocity filter can be increased if a few tens of samples along the n_x and n_y dimensions are available with an LFV, a feature that is expected to be available with the future generations of LFV cameras. Considering the ultra-low complexity and the availability of the closed-form expressions for the time-variant coefficients, the proposed 5-D IIR adaptive depth-velocity filter has a significant potential to be employed in real-time applications.

Chapter 6

Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter

6.1 Introduction

The 5-D FIR and IIR depth-velocity filters presented in Chapter 3 and Chapter 4, respectively, employed 5-D FIR and IIR filters having 4-D hyperplanar passbands. One disadvantage of such a 4-D hyperplanar passband is that it provides poor selectivity at frequencies near the origin inside the 5-D principal Nyquist hypercube. This disadvantage can be overcome by employing 5-D FIR or IIR filters having 5-D hyperfan-shaped passbands instead of 4-D hyperplanar passbands. For example, the required 5-D hyperfan-shaped passband for the 5-D filter $H_{uvt}(\mathbf{z})$ is shown in Figure 6.1. Note that the 5-D hyperfan-shaped passband encompasses the ROS of the object of interest more compactly compared to a 4-D hyperplanar passband. In this chapter, we present the design of an ultra-low complexity 3-D *linear-phase* IIR velocity filter in the context of conventional 3-D video enhancement. This filter can be directly employed as the 5-D IIR filter $H_{uvt}(\mathbf{z})$ in the design of 5-D IIR depth-velocity filters. Furthermore, the 5-D IIR filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ can be designed by slightly modifying the design method of the proposed filter.

3-D velocity filters (also known as linear trajectory (LT) filters) are employed in such applications as traffic analysis, motion detection and analysis, radar tracking and advanced driver assistance systems to selectively enhance or attenuate moving objects based on their 2-D spatial (or apparent) velocities [18] [19] [20] [21] [23] [27] [30]. Another potential application is the attenuation of moving artifacts in underwater

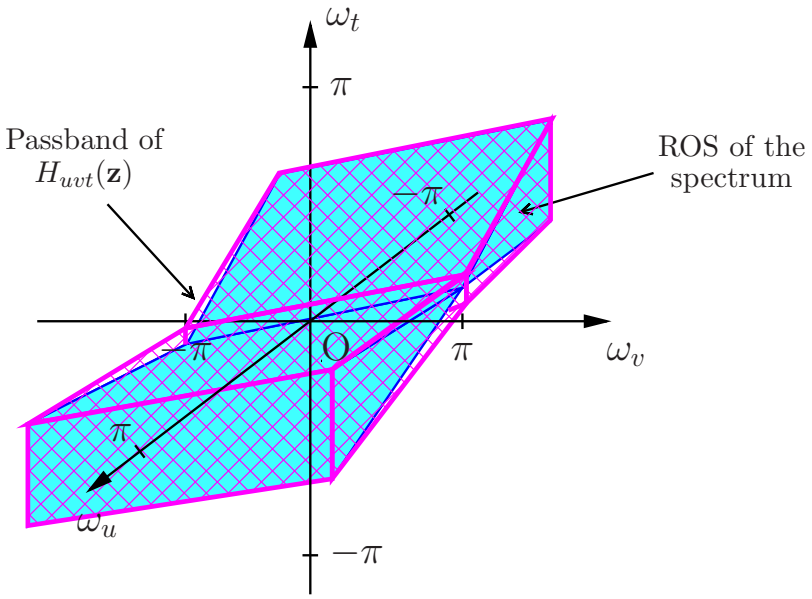


Figure 6.1: The ROS of the spectrum of the object of interest (solid) and the 5-D hyperfan-shaped passband of the 5-D filter $H_{uvt}(\mathbf{z})$ in the $\omega_u\omega_v\omega_t$ space.

videos while enhancing static objects or objects moving on smooth approximately-linear trajectories. Two premier examples are the attenuation of sunlight flicker patterns in shallow underwater videos [108] [109] [110] and the attenuation of live creatures, such as fish, in seabed modeling and coral reef monitoring [15].

A number of methods for the design of 3-D FIR and IIR velocity filters are found in the literature. In most of the design methods, the passband has been selected to be planar shaped [24] [25] [26] or wedge shaped (i.e., exterior of a wide-angle cone) [22] [27] [28] [29] [111] inside the principal Nyquist cube in the 3-D frequency domain. In general, these passbands are not symmetric with respect to the temporal frequency axis.

Recently, a low-complexity 3-D FIR velocity filter is proposed to enhance severely corrupted LT signals [77]. Motivated from *the temporal delaying in conventional delay-and-sum beamformers* [73](ch. 6) [112] [113], the low-complexity 3-D FIR velocity filter has employed a 2-D spatial variable-shift filter to spatially shift the 3-D input video signal before filtering with a low-complexity 3-D wide-angle FIR cone filter. The use of the 2-D spatial shifts has the effect of skewing (or shearing) the region of support of the spectrum of an LT signal so that it is *symmetric with respect to the temporal frequency axis*. Therefore, the wedge-shaped passband of the low-complexity 3-D wide-angle FIR cone filter can be selected to be symmetric with

respect to the temporal frequency axis and has been approximated by employing a low-complexity filter bank structure similar to those employed in 3-D FIR cone and frustum filters, which have been used for broadband beamforming [114] [115].

In this chapter, an ultra-low complexity 3-D *linear-phase* IIR velocity filter, which extends the low-complexity 3-D FIR velocity filter proposed in [77], is proposed to further reduce the computational complexity of 3-D velocity filters. The proposed 3-D linear-phase IIR velocity filter consists of an ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank between two 2-D spatial variable-shift filters. The ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank is designed by employing a 1-D temporal modified DFT filter bank and 2-D spatial allpass, IIR highpass and allstop filters. The ultra-low complexity is achieved by exploiting the maximal decimation in the temporal dimension and by realizing the *nonseparable* 2-D spatial IIR highpass filters using the 2-D spatial allpass filters and 1-D spatial IIR lowpass filters. Furthermore, the linear phase response is achieved by employing zero-phase filtering for the 1-D spatial IIR lowpass filters. A typical 3-D linear-phase IIR velocity filter of order $4 \times 4 \times 510$, applied to enhance a heavily corrupted test video signal, provides a signal-to-interference-and-noise ratio (SINR) improvement of 13.86 dB, which is almost *independent* of the scene complexity. Compared to the 3-D FIR velocity filter presented in [77], the proposed 3-D linear-phase IIR velocity filter requires 52.28% and 66.41% less real multiplications and additions, respectively, for similar SINR improvements. In order to illustrate the performance of the proposed 3-D linear-phase IIR velocity filter in enhancing videos, it is employed to attenuate sunlight flicker patterns in shallow underwater videos. Experimental results confirm the effectiveness of the 3-D linear-phase IIR velocity filter in attenuating the sunlight flicker patterns and its robustness to motion estimation errors. Note that, in this chapter, the image coordinate of a 3-D video is denoted as (x, y) instead of (u, v). Furthermore, the apparent velocities of a moving object in a 3-D video along the x and y dimensions are denoted as v_x and v_y, respectively.

The organization of the rest of the chapter is as follows. A review of LT signals and their spectra is presented in Section 6.2. In Section 6.3, the proposed 3-D linear-phase IIR filter is described in detail. In Section 6.4, the design and the efficient implementation of the 3-D wide-angle linear-phase IIR cone filter bank are presented. Numerical simulation results are presented in Section 6.5. Finally, a summary of the chapter is presented in Section 6.6.

6.2 Review of LT Signals and Their Spectra

Let us consider a sampled LT signal corresponding to an object moving with a constant 2-D spatial velocity $[v_x, v_y]^{\mathrm{T}}$ as shown in Figure 6.2 (a). Under the *constant intensity assumption* [88](ch. 2.3), i.e. assuming that the intensity of the moving object is time invariant, the 3-D intensity of the moving object, $i(n_x, n_y, n_t)$, $(n_x, n_y, n_t) \in \mathbb{Z}^3$, may be written as

$$i(n_x, n_y, n_t) = i_R(n_x - \Delta x(n_t), n_y - \Delta y(n_t), n_{t0}), \tag{6.1}$$

where $i_R(n_x, n_y, n_{t0})$ is the 2-D intensity of the moving object at $n_t = n_{t0}$. The frame obtained at $n_t = n_{t0}$ is called the *reference frame*, and $\Delta x(n_t)$ and $\Delta y(n_t)$ are, respectively, the displacements (in pixels) of the moving object along the n_x and n_y dimensions at n_t with respect to the reference frame. The displacements $\Delta x(n_t)$ and $\Delta y(n_t)$ are given by [88](ch. 2.3)

$$\Delta x(n_t) = v_x(n_t - n_{t0}) \tag{6.2a}$$

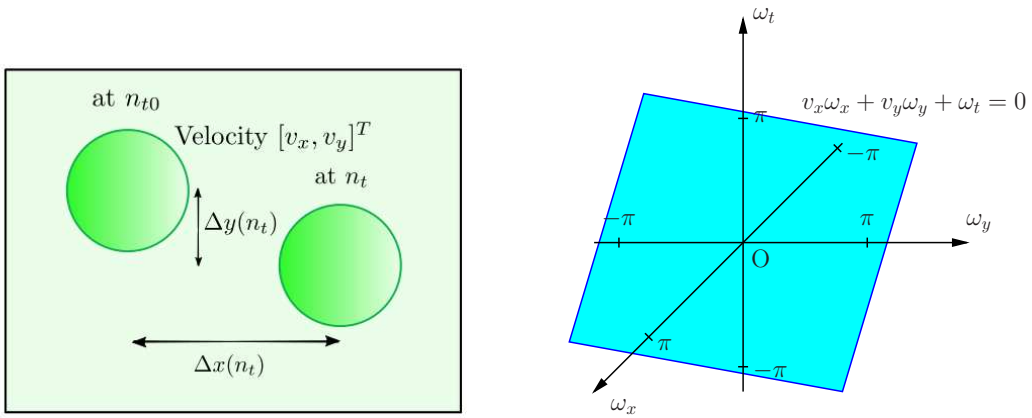
$$\Delta y(n_t) = v_y(n_t - n_{t0}), \tag{6.2b}$$

respectively.

The 3-D spectrum of the LT signal $i(n_x, n_y, n_t)$ can be obtained as [24]

$$I(\omega_x, \omega_y, \omega_t) = I_R(\omega_x, \omega_y)\, \delta(v_x\omega_x + v_y\omega_y + \omega_t), \tag{6.3}$$

where $(\omega_x, \omega_y, \omega_t) \in \mathbb{R}^3$ and $\delta(\cdot)$ is the 1-D continuous-domain impulse function. According to (6.3), the region of support of the spectrum $I(\omega_x, \omega_y, \omega_t)$ is a *plane* going through the origin inside the principal Nyquist cube $\{(\omega_x, \omega_y, \omega_t) \in \mathbb{R}^3,\ -\pi \leq \omega_x, \omega_y, \omega_t \leq \pi\}$ [24] as shown in Figure 6.2 (b). Note that the orientation of the plane is determined by the 2-D spatial velocity $[v_x, v_y]^{\mathrm{T}}$. Consequently, the regions of support of the spectra corresponding to objects moving with different 2-D spatial velocities do not ideally overlap except at the origin in the 3-D frequency domain [24], and such moving objects can be selectively enhanced or attenuated by employing 3-D velocity filters. Nevertheless, it is worthwhile to note that the aliasing components resulting from sampling rates employed in video capturing processes [88](ch. 3.3) cannot be eliminated by 3-D velocity filters. This, in general, poses a natural limit for applications of 3-D velocity filters.



(a) (b)

Figure 6.2: (a) An LT signal corresponding to an object moving with a constant 2-D spatial velocity $[v_x, v_y]^{\mathrm{T}}$; (b) the region of support of its spectrum inside the principal Nyquist cube.

6.3 Proposed Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter

The proposed ultra-low complexity 3-D linear-phase IIR velocity filter, henceforth called as the 3-D IIR velocity filter for brevity, is illustrated in Figure 6.3. It consists of a 3-D wide-angle linear-phase IIR cone filter $C(z_x, z_y, z_t)$, $(z_x, z_y, z_t) \in \mathbb{C}^3$, between two 2-D spatial variable-shift filters $S_b(z_x, z_y)$ and $S_f(z_x, z_y)$, $(z_x, z_y) \in \mathbb{C}^2$. In order to describe the operation of the proposed 3-D IIR velocity filter, let us consider a 3-D input signal $i_{in}(n_x, n_y, n_t)$ consisting of an LT signal of interest having the 3-D intensity $i(n_x, n_y, n_t)$ and moving with the 2-D spatial velocity $[v_x, v_y]^{\mathrm{T}}$, and interfering signals having 2-D spatial velocities other than $[v_x, v_y]^{\mathrm{T}}$. The first 2-D spatial variable-shift filter $S_b(z_x, z_y)$ spatially shifts $i_{in}(n_x, n_y, n_t)$ by $[-\Delta x(n_t), -\Delta y(n_t)]^{\mathrm{T}}$ (or provides the *backward* shifts) so that the 3-D intensity $i(n_x, n_y, n_t)$ of the LT signal of interest is *stationary* with respect to the temporal dimension. This is equivalent to skewing the region of support of the spectrum $I(\omega_x, \omega_y, \omega_t)$ to lie on the $\omega_x\omega_y$ plane. On the other hand, the regions of support of the spectra of interfering signals occupy regions other than the $\omega_x\omega_y$ plane (except the origin).

In real applications, the 1-D spatial shifts $\Delta x(n_t)$ and $\Delta y(n_t)$ can be estimated by means of readily available motion estimation techniques [88](ch. 6). In this case, the 1-D spatial shifts are approximations of the exact 1-D spatial shifts. As a result, after the 2-D backward spatial shifts $[-\Delta x(n_t), -\Delta y(n_t)]^{\mathrm{T}}$, the region of support of

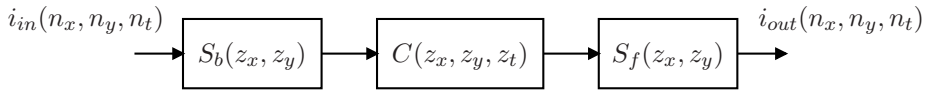


Figure 6.3: Proposed ultra-low complexity 3-D linear-phase IIR velocity filter.

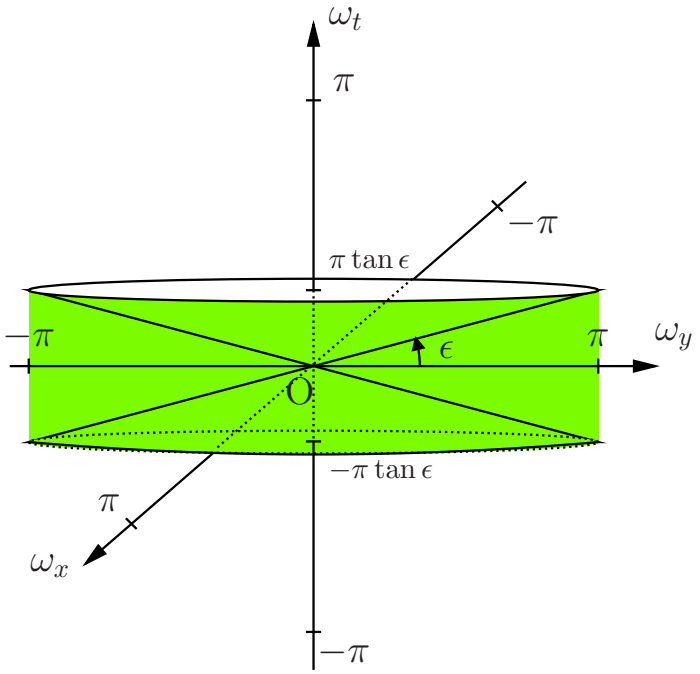


Figure 6.4: Ideal passband of the 3-D IIR filter $C(z_x, z_y, z_t)$, which is the exterior of a wide-angle cone, where ϵ is the angle between the $\omega_x\omega_y$ plane and the surface of the wide-angle cone.

the spectrum $I(\omega_x, \omega_y, \omega_t)$ may deviate slightly from the $\omega_x\omega_y$ plane. Consequently, in order to make the 3-D IIR velocity filter *robust* to spatial shift estimation errors, the passband of the 3-D IIR filter $C(z_x, z_y, z_t)$ should be the exterior of a wide-angle cone [29], as shown in Figure 6.4. Note that the passband of the 3-D IIR filter $C(z_x, z_y, z_t)$ is *symmetric with respect to the* ω_t *axis*. Therefore, it can be approximated using an ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank as described in Section 6.4. The second 2-D spatial variable-shift filter $S_f(z_x, z_y)$ is used to reverse the 2-D spatial shifts introduced by the first 2-D spatial variable-shift filter $S_b(z_x, z_y)$. In other words, $S_f(z_x, z_y)$ provides the *forward* shifts necessary to reintroduce the motion for the LT signal of interest.

In the implementation of the 2-D spatial variable-shift filters $S_b(z_x, z_y)$ and $S_f(z_x, z_y)$, the 2-D spatial shifts can be carried out as two 1-D spatial shifts for each temporal frame. For most applications, it is sufficient to estimate the 1-D spatial shifts $\Delta x(n_t)$

and $\Delta y(n_t)$ to the accuracy of one pixel, i.e. both $\Delta x(n_t)$ and $\Delta y(n_t)$ are *integer valued*. Consequently, for each temporal frame, the 1-D spatial shifts can be implemented using *circular shifts* along the n_x and n_y dimensions. *Note that this does not require any arithmetic operations, and the additional memory requirement is very low.*

6.4 Design and Efficient Implementation of the Ultra-Low Complexity 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank

The design and the efficient implementation of the ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank are presented in this section. For brevity, we henceforth call it the cone filter bank. The proposed cone filter bank, shown in Figure 6.5(a), consists of a 1-D temporal modified DFT filter bank [116] [117] and 2-D spatial filters. The 1-D temporal modified DFT filter bank employs a two-step decimation and the subsampling matrices $\mathbf{M_1}$ and $\mathbf{M_2}$ are given by [118](ch. 12)

$$\mathbf{M_1} = \text{diag}\,[1 \quad 1 \quad M/2] \tag{6.4a}$$

$$\mathbf{M_2} = \text{diag}\,[1 \quad 1 \quad 2]\,, \tag{6.4b}$$

respectively, where the term "diag" denotes a diagonal matrix. The 2-D spatial filters include a 2-D spatial allpass filter, $2M_h$ 2-D spatial IIR highpass filters and $M-(2M_h+1)$ 2-D spatial allstop filters, where M (an even integer) is the number of bands of the 1-D temporal modified DFT filter bank. For given M and ϵ, M_h is obtained as

$$M_h = \arg\,\min_k \left| \frac{2k+1}{M} - \tan\epsilon \right|, \tag{6.5}$$

where $k \in \{1, 2, \ldots, M/2-1\}$. The passband of the cone filter bank is approximated using the 2-D spatial allpass filter and the $2M_h$ 2-D spatial IIR highpass filters as illustrated in Figure 6.5(b). The 2-D spatial allstop filters are utilized for the stopband of the cone filter bank.

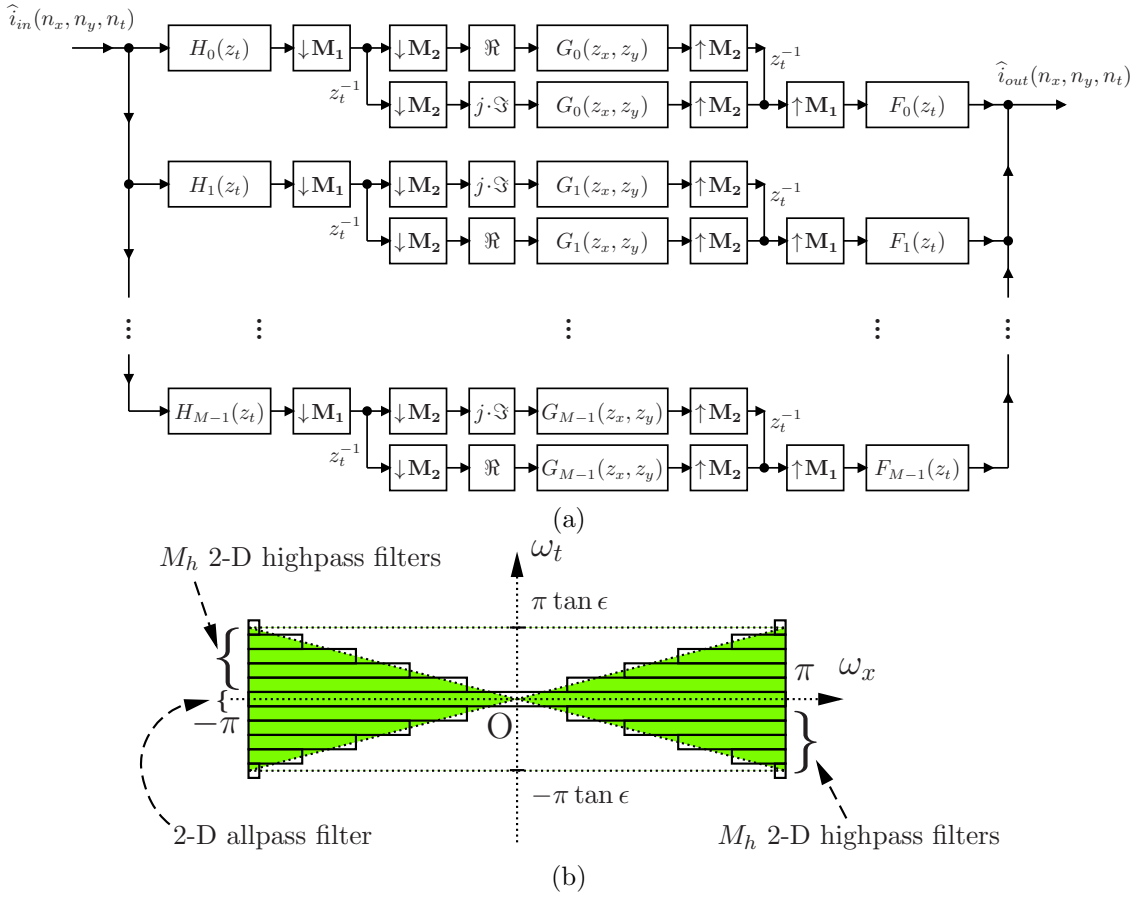


Figure 6.5: (a) The 3-D wide-angle linear-phase IIR cone filter bank. (b) Approximation of the passband using a 2-D spatial allpass filter and $2M_h$ 2-D spatial highpass filters; a planar view on the $\omega_x\omega_t$ plane.

6.4.1 Design of the 1-D Temporal Modified DFT Filter Bank

Both analysis filters $H_k(z_t)$ and synthesis filters $F_k(z_t)$, $z_t \in \mathbb{C}$ and $k = 0, 1, \ldots, M-1$, of a 1-D modified DFT filter bank are obtained by means of uniform complex modulation of a zero-phase FIR lowpass prototype filter $P(z_t)$ having a cutoff frequency π/M and a passband gain $\sqrt{M}$ as [117]

$$H_k(z_t) = F_k(z_t) = z_t^{-N_P/2} P(z_t W_M^k), \tag{6.6}$$

where $W_M = \mathrm{e}^{-j2\pi/M}$ and N_P is the order of the prototype filter. Therefore, the design of the 1-D temporal modified DFT filter bank is effectively reduced to the design of the prototype filter. The order N_P of the prototype filter is selected as $2M - 1$ and, for this case, the impulse response of the prototype filter, $p(n_t)$, can be

obtained in closed form as [117] [119]

$$p(n_t) = \frac{-1}{2\sqrt{M}} + \frac{1}{\sqrt{2M}} \cos\left(\frac{\pi}{2M}(2n_t + 2M)\right),$$
$$n_t = -M + 0.5, -M + 1.5, \ldots, M - 0.5. \tag{6.7}$$

6.4.2 Design of the 2-D Spatial Filters

The 2-D spatial filter corresponding to the 0th band, $G_0(z_x, z_y)$, $(z_x, z_y) \in \mathbb{C}^2$, is allpass and its impulse response is given by

$$g_0(n_x, n_y) = \delta_{2\mathrm{D}}(n_x, n_y), \tag{6.8}$$

where $(n_x, n_y) \in \mathbb{Z}^2$ and $\delta_{2\mathrm{D}}(\cdot, \cdot)$ is the 2-D discrete-domain impulse function. The 2-D spatial filters $G_k(z_x, z_y)$ corresponding to $k = 1, 2, \ldots, M_h$ and $k = M - M_h, M - (M_h + 1), \ldots, M - 1$ are 2-D spatial IIR highpass filters having circularly symmetric stopbands. In order to reduce the computational complexity, the frequency responses of the 2-D spatial IIR highpass filters are approximated by employing a 2-D spatial allpass filter and 1-D spatial IIR lowpass filters as

$$G_k(z_x, z_y) = AP(z_x, z_y) - G_{lp,k}(z_x)G_{lp,k}(z_x^{-1})G_{lp,k}(z_y)G_{lp,k}(z_y^{-1}), \tag{6.9}$$

where $AP(z_x, z_y)$ is the 2-D spatial allpass filter having an impulse response $ap(n_x, n_y) = \delta_{2\mathrm{D}}(n_x, n_y)$, and $G_{lp,k}(\cdot)$ are the 1-D spatial IIR lowpass filters. Note that $G_k(z_x, z_y)$ is a *zero-phase* filter [86](ch. 12.5). The cutoff frequencies of the 1-D spatial IIR lowpass filters, $\omega_{c,k}$, are selected as

$$\omega_{c,k} = \frac{2k\pi}{M} \cot \epsilon, \qquad k = 1, 2, \ldots, M_h, \tag{6.10}$$

and

$$\omega_{c,k} = \omega_{c,M-k}, \qquad k = M - M_h, M - (M_h + 1), \ldots, M - 1. \tag{6.11}$$

In general, low-order classical 1-D IIR filter types, such as Butterworth and elliptic, are good choice for the 1-D spatial IIR lowpass filters. The design of such a 1-D IIR lowpass filter can be carried out following a standard design procedure [86](ch. 12).

The impulse responses of the 2-D spatial allstop filters are simply zero, i.e.,

$$g_k(n_x, n_y) = 0, \qquad k = M_h + 1, M_h + 2, \ldots, M - (M_h - 1). \tag{6.12}$$

6.4.3 Near-Perfect Reconstruction of the 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank

The cone filter bank provides the *near-perfect reconstruction* as shown in Appendix D. That is, it approximates a linear time-invariant system, and the transfer function can be obtained as

$$\begin{aligned} C(z_x, z_y, z_t) &\approx \frac{z_t^{-M/2}}{M} \sum_{k \in \mathcal{C}} H_k(z_t) G_k(z_x, z_y) F_k(z_t) \\ &\approx \frac{z_t^{-(N_P + M/2)}}{M} \sum_{k \in \mathcal{C}} \left[P(z_t W_M^k) \right]^2 G_k(z_x, z_y), \end{aligned} \tag{6.13}$$

where $\mathcal{C} = \{0, 1, \ldots, M_h, M - M_h, M - (M_h + 1), \ldots, M - 1\}$. Furthermore, the temporal group delay $\tau(\omega_x, \omega_y, \omega_t)$ of the cone filter bank can be obtained from (6.13) as

$$\begin{aligned} \tau(\omega_x, \omega_y, \omega_t) &= N_P + \frac{M}{2} \\ &= 2M + \frac{M}{2} - 1. \end{aligned} \tag{6.14}$$

6.4.4 Efficient Implementation of the 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank

The efficient realization of the cone filter bank is shown in Figure 6.6. In this case, the 1-D temporal modified DFT filter bank is realized by employing two 1-D temporal DFT-polyphase filter banks [116] [120], and the subsampling matrix $\mathbf{M}$ is given by [118](ch. 12)

$$\mathbf{M} = \mathrm{diag}\,[1 \quad 1 \quad M]\,. \tag{6.15}$$

Note that $E_m(z_t)$ and $R_m(z_t)$, $m = 0, 1, \ldots, (M-1)$, are the type-1 polyphase components of $H_0(z_t)$ and the type-3 polyphase components of $F_0(z_t)$, respectively. Their

definitions are, respectively, given by [121](ch. 3.3)

$$H_0(z_t) = \sum_{m=0}^{M-1} z_t^{-m} E_m(z_t^M) \tag{6.16a}$$

$$F_0(z_t) = \sum_{m=0}^{M-1} z_t^{m} R_m(z_t^M). \tag{6.16b}$$

Note that $H_0(z_t) = F_0(z_t) = z_t^{-N_P/2} P(z_t)$. For more details on the polyphase decompositions, the reader is referred to a standard textbook on filter banks such as [118] or [121]. Note that, in the efficient realization of the cone filter bank, both temporal and spatial filtering operations are carried out after the temporal downsampling and before the temporal upsampling. Consequently, the computational complexity of the efficient realization is M times less than that of the original realization of the cone filter bank shown in Figure 6.5 (a).

The number of *nontrivial* real multiplications and additions, i.e. excluding the multiplications with ± 1, $\pm j$ and 0 and the additions with 0, required to process a sample by the different blocks of the efficient realization of the cone filter bank are presented in Table 6.1. In this derivation, the order of the 2-D spatial IIR highpass filters is assumed to be $2N_S \times 2N_S$, where N_S is the order of the 1-D spatial IIR lowpass filters. Furthermore, all the 1-D temporal polyphase filters and the 1-D spatial IIR lowpass filters are assumed to be implemented using direct-form structures. Note that *no nontrivial arithmetic operation* is required for the implementations of the 2-D allpass filters and the 2-D allstop filters. Because the input signals for the inverse DFT (IDFT) blocks are real-valued, a fast Fourier transform (FFT) algorithm optimized for real-valued signals [122] can be employed to implement the IDFT blocks whereas the DFT blocks can be implemented by employing a general FFT algorithm designed for complex-valued signals [123](chs. 3 and 12). In Table 6.1, $M_{\mathrm{FFT},R}$, $A_{\mathrm{FFT},R}$, $M_{\mathrm{FFT},C}$ and $A_{\mathrm{FFT},C}$, respectively, denote the number of nontrivial real multiplications and additions required by the employed real-valued and complex-valued FFT algorithms to perform an M-point DFT (or IDFT). For example, the split-radix FFT algorithm optimized for real-valued signals [122] requires 258 nontrivial real multiplications ($M_{\mathrm{FFT},R}$) and 1028 nontrivial real additions ($A_{\mathrm{FFT},R}$) to perform a 128-point DFT (or IDFT). Furthermore, in the derivation, the multiplication by M is assumed to be carried out before performing the IDFT in the IDFT blocks, and a complex multiplication is assumed to be carried out with 3 real multiplications and

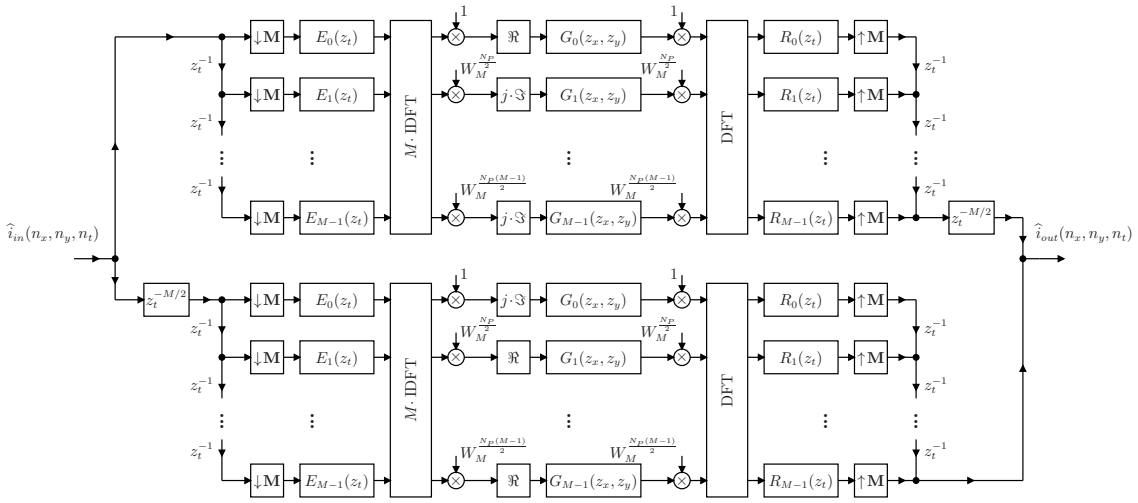


Figure 6.6: Efficient realization of the 3-D wide-angle linear-phase IIR cone filter bank. The 1-D temporal modified DFT filter bank is realized by two 1-D temporal DFT-polyphase filter banks.

Table 6.1: Nontrivial real multiplications and additions required to process a sample by the different blocks of the 3-D wide-angle linear-phase IIR cone filter bank.

Block	Real multiplications	Real additions
1-D temporal polyphase filters	8	4
IDFT and multiplication by M	$\frac{2}{M} M_{\mathrm{FFT},R} + 2$	$\frac{2}{M} A_{\mathrm{FFT},R}$
2-D spatial IIR highpass filters	$\frac{16 M_h (2N_S + 1)}{M}$	$\frac{4 M_h (8N_S + 1)}{M}$
DFT	$\frac{2}{M} M_{\mathrm{FFT},C}$	$\frac{2}{M} A_{\mathrm{FFT},C}$
$W_M^{\frac{kN_P}{2}}$ terms	$\frac{20 M_h}{M}$	$\frac{12 M_h}{M}$
Addition after upsampling	—	1

3 real additions.

6.5 Numerical Simulation Results

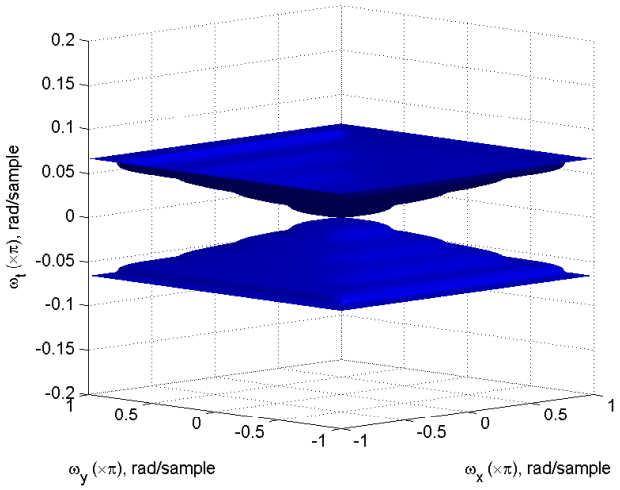
Numerical simulation results are presented in this section to confirm the effectiveness of the proposed 3-D IIR velocity filter. In particular, we consider the enhancement of a heavily corrupted test video signal and the attenuation of the sunlight flicker patterns in shallow underwater videos. Furthermore, a design example of the proposed 3-D IIR velocity filter and a brief comparison between previously reported 3-D velocity filters and the proposed 3-D IIR velocity filter are presented.

6.5.1 A Design Example of the Proposed Ultra-Low Complexity 3-D Linear-Phase IIR Velocity Filter

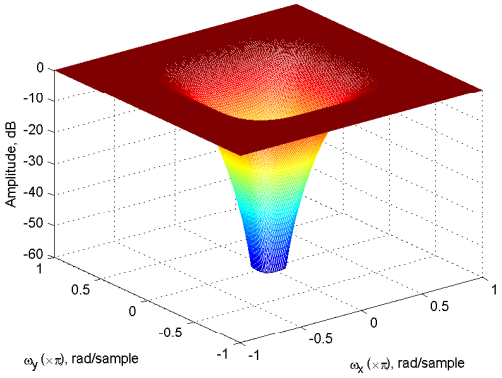
The design of the proposed 3-D IIR velocity filter mainly consists of the design of the cone filter bank, which in turn consists of the design of the 1-D FIR prototype filter $P(z_t)$ of the 1-D temporal modified DFT filter bank and the designs of the $2M_h$ 1-D spatial IIR lowpass filters $G_{lp,k}(\cdot)$, $k = \{1, 2, \ldots, M_h, M - M_h, M - (M_h + 1), \ldots, M - 1\}$. For the design example presented in this subsection, the number of bands M of the 1-D temporal modified DFT filter bank and the angle ϵ are selected as 128 and 4°, respectively. Note that the order N_P of the 1-D FIR prototype filter $P(z_t)$ is 255. For the selected values of M and ϵ, M_h is obtained as 4 from (6.5), and 2nd-order 1-D Butterworth lowpass filters are employed to realize the 2-D spatial IIR highpass filters. The order of the cone filter bank is $4 \times 4 \times 510$. The magnitude response of the cone filter bank $C(z_x, z_y, z_t)$ is shown in Figure 6.7. It is observed that the magnitude response approximates the required passband fairly well.

6.5.2 Performance in Enhancing a Heavily Corrupted Approximately LT Signal

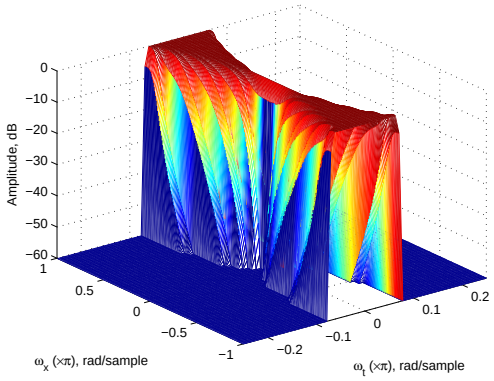
The performance of the proposed 3-D IIR velocity filter in enhancing a heavily corrupted approximately LT signal is numerically evaluated in this subsection. To this end, a test video signal of size $256 \times 256 \times 512$, which is similar to that employed in [29], is numerically generated. The test video signal contains an approximately LT signal of interest $i_s(n_x, n_y, n_t)$, an interfering signal $i_i(n_x, n_y, n_t)$ and additive white Gaussian noise. Both the LT signal of interest and the interfering signal consist of a *highly-detailed* black (intensity $= -0.5$) and white (intensity $= +0.5$) checkerboard



(a)



(b)



(c)

Figure 6.7: Magnitude response of the 3-D wide-angle linear-phase IIR cone filter bank (a) -3 dB iso-surface; (b) a slice at $\omega_t = 0.0313\pi$; (c) a slice at $\omega_y = 0$.

pattern $i_c(n_x, n_y, n_t)$ of 64×64 pixels in a gray (intensity $= 0$) background [29]. The two signals $i_s(n_x, n_y, n_t)$ and $i_i(n_x, n_y, n_t)$ are numerically generated as

$$i_s(n_x, n_y, n_t) = i_c(\lceil 5 + 0.1n_t + 0.0005n_t^2 \rceil, \lceil 100 + 0.09n_t - 7\sin(2\pi n_t/800) \rceil, n_t) \tag{6.17a}$$

$$i_i(n_x, n_y, n_t) = i_c(\lceil 180 - 0.7n_t + 0.0007n_t^2 \rceil, \lceil 50 + 0.09n_t - 7\sin(2\pi n_t/800) \rceil, n_t), \tag{6.17b}$$

respectively, where $\lceil \cdot \rceil$ is the ceiling function. The energy of the interfering signal and the noise is selected as the same and 10 times the energy of the signal of interest,

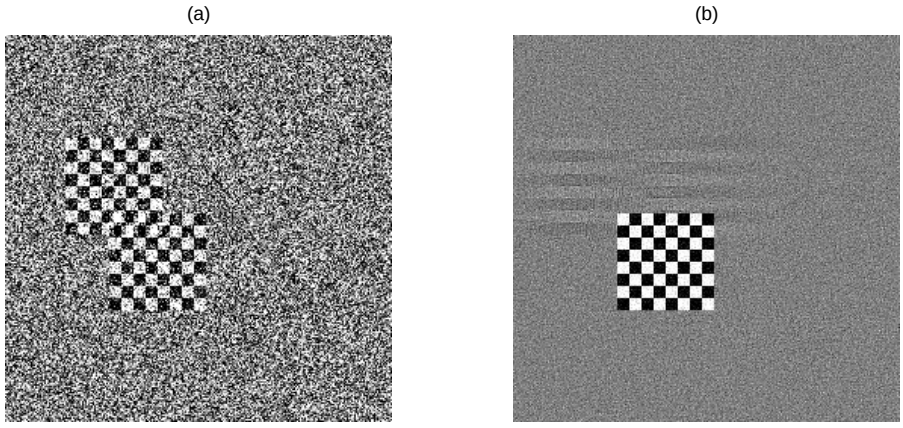


Figure 6.8: The 275th frame of (a) input video (b) output video.

respectively, in order to represent a *heavily corrupted* input signal. The 3-D IIR velocity filter designed in Section 6.5.1 is employed to enhance the signal of interest. The 275th frame of the input and the output test video signals are shown in Figure 6.8. Note that the signal of interest and the interfering signal partially overlap in the input frame, and the signal of interest is successfully enhanced by the proposed 3-D IIR velocity filter.

The SINR improvement is employed as a measure in order to numerically evaluate the performance of the proposed 3-D IIR velocity filter. The SINR of the input signal is -10.41 dB. The segment of the output signal corresponding to the temporal steady state, $\{(n_x, n_y, n_t)|1 \leq n_x, n_y \leq 256, 129 \leq n_t \leq 384\}$, is employed to estimate the SINR of the output signal. The SINR of the output signal is 3.45 dB and, consequently, the SINR improvement is 13.86 dB. For the same test video signal, the 3-D FIR velocity filter (of order $30 \times 30 \times 510$) proposed in [77] provides a 13.25 dB SINR improvement. Accordingly, the proposed 3-D IIR velocity filter (of order $4 \times 4 \times 510$) provides an additional SINR improvement of 0.61 dB because it has a slightly higher stopband attenuation compared to the 3-D FIR velocity filter of order $30 \times 30 \times 510$ [77]. *Note that the SINR improvement mainly depends on the stopband attenuation of the 3-D velocity filter and the 2-D spatial velocities of the moving objects. It is almost independent of the scene complexity.*

6.5.3 Attenuation of Sunlight Flicker Patterns in Shallow Underwater Videos

Sunlight flicker is a familiar natural phenomenon observed on shallow underwater surfaces on sunny days. The sunlight is converged by the convex regions of water surface waves creating sunlight flicker patterns. These patterns appear as spatio-temporal-varying brighter regions on underwater surfaces and considerably distort the shallow underwater images and videos. Such moving artifacts in shallow underwater videos affect not only the human perception of the scene but also the behaviour of image processing and computer vision algorithms [108] [109]. Consequently, the attenuation of the sunlight flicker patterns is of great interest. Previously proposed methods [108] [109] [110] to attenuate the sunlight flicker patterns are nonlinear and employ temporal median filtering in the log-intensity domain. In this subsection, the performance of the proposed 3-D IIR velocity filter in attenuating the sunlight flicker patterns is evaluated. Note that the proposed 3-D IIR velocity filtering method is *linear* and operates in the *linear-intensity domain*.

In order to evaluate the performance of the proposed method in attenuating the sunlight flicker patterns, two underwater videos, "pool" and "caesarea"[1], are considered. Figure 6.9 shows the 100th frame of the two underwater videos. Underwater surface in the "pool" underwater video is shallower than that in the "caesarea" underwater video. Consequently, the sunlight flicker pattern in the former is much stronger than that in the latter. Both videos are processed by employing the 3-D IIR velocity filter designed in Section 6.5.1 and by employing the state-of-the-art method proposed in [109]. Figures 6.10 and 6.11 show the 100th frame of the processed "pool" and "caesarea" underwater videos, respectively. Note that the strong sunlight flicker pattern in the "pool" underwater video is completely attenuated by the proposed method whereas some residues of the attenuated sunlight flicker pattern can be observed in the underwater video processed using the method in [109]. *This indicates the better performance of the proposed method, especially attenuating the strong sunlight flicker patterns.*

In general, underwater videos are captured by cameras mounted on fixed structures or autonomous underwater vehicles. In the first case, the object of interest, i.e. the seabed or coral reefs, is static in the underwater video and the object of

[1]Original videos are available at `http://webee.technion.ac.il/~yoav/research/flicker.html`

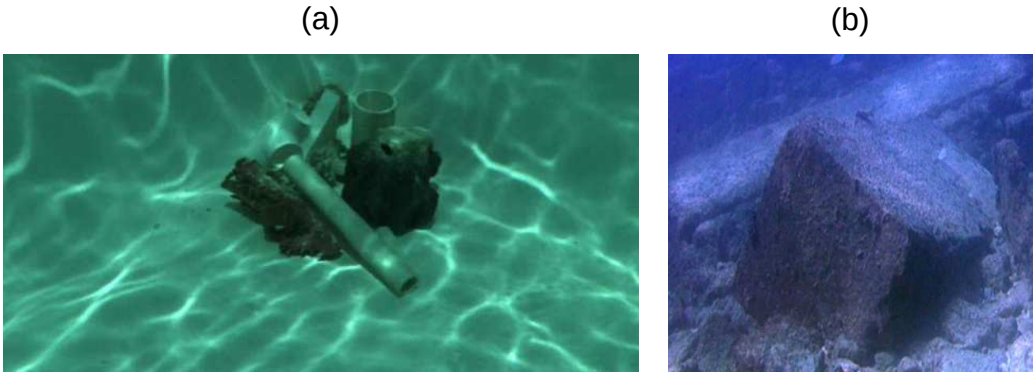


Figure 6.9: The 100th frame of the (a) "pool" underwater video; (b) "caesarea" underwater video.

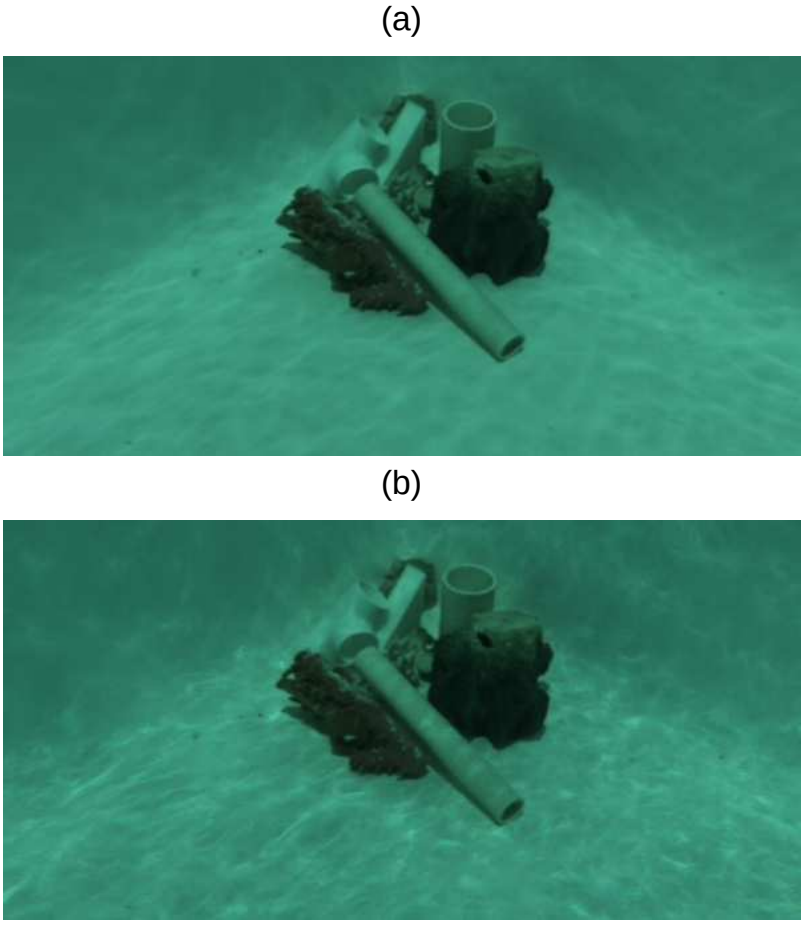


Figure 6.10: The 100th frame of the processed "pool" underwater video (a) proposed method; (b) method in [109].

interest can be enhanced directly by employing the cone filter bank. That is, the processing with the two 2-D spatial variable-shift filters is not necessary. In the sec-

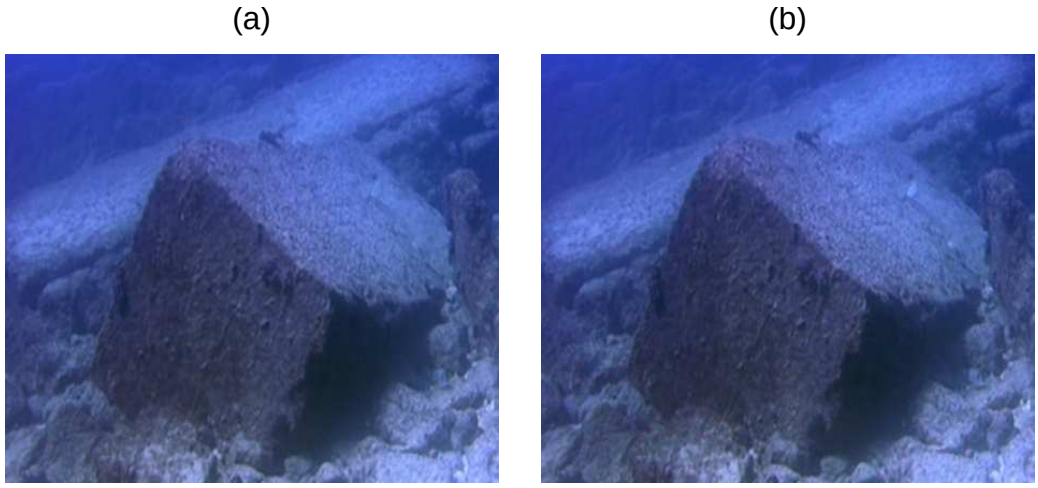


Figure 6.11: The 100th frame of the processed "caesarea" underwater video (a) proposed method; (b) method in [109].

ond case, it is assumed that the autonomous underwater vehicle follows a smooth approximately-linear trajectory, and the motion of the autonomous underwater vehicle is required to be estimated in order to calculate the 2-D spatial shifts. In both cases, the underwater video has motion-related errors. In the first case, the object of interest is not perfectly static due to slight deviations of the camera position and, in the second case, the motion estimation errors are unavoidable. Next, we show that the proposed 3-D IIR velocity filtering method is *robust* to such motion-related errors. To this end, uniformly distributed random 2-D spatial shifts are artificially introduced to both underwater videos in order to represent the motion estimation errors, and both videos are processed with the 3-D IIR velocity filter designed in Section 6.5.1. In order to numerically evaluate the performance of the proposed method, the average PSNR is estimated for different levels of 2-D spatial-shift errors. In the PSNR estimation, the outputs of the 3-D IIR velocity filter corresponding to the original motion-related-error-free underwater videos are considered as the noise-free ground truths. Figure 6.12 shows the average PSNR obtained for different levels of 2-D spatial-shift errors. As expected, the average PSNR is reduced with the increased level of 2-D spatial-shift errors. However, the average PSNR is greater than 30 dB for both underwater videos, when the 2-D spatial-shift errors are in the range $[-6, 6]$ pixels, which is a good match for fast low-complexity motion estimation techniques. Moreover, the average PSNR of 30.57 dB and 27.43 dB are achieved for the "pool" and "caeserea" underwater videos, respectively, even for the worst case considered in the

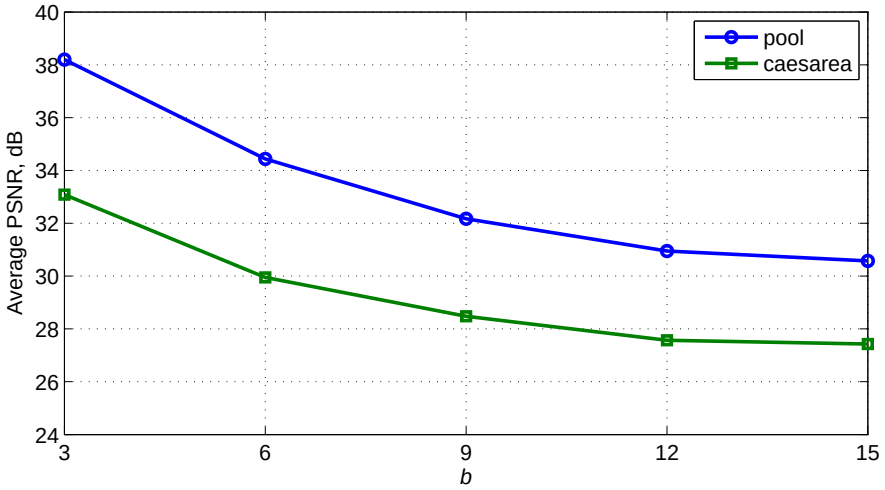


Figure 6.12: The average PSNR of the "pool" and "caesarea" underwater videos obtained with different levels of random 2-D spatial-shift errors, which are uniformly distributed in the range $[-b, b]$ pixels.

simulations, i.e. with the range $[-15, 15]$ pixels for the 2-D spatial-shift errors. These results confirm the *robustness* of the proposed method to the motion-related errors. *Considering the ultra-low complexity and robustness to the motion-related errors, the proposed method has an excellent potential to be employed in real-time applications such as real-time monitoring of coral reefs.*

6.5.4 Comparison with Previously Reported 3-D Velocity Filters

In this subsection, we briefly compare the proposed 3-D IIR velocity filter with previously reported 3-D velocity filters [27] [28] [29] [77] of which the passbands are similar to that of the proposed filter. To this end, we first consider the computational complexity of the 3-D IIR velocity filter designed in Section 6.5.1 and those of the four 3-D velocity filters considered in the study. The specifications of the filters are selected so that they provide *similar SINR improvements*. The numbers of nontrivial real multiplications and additions required to process a sample are presented in Table 6.2. Note that, in estimating the numbers of nontrivial real multiplications and additions required by the proposed 3-D IIR velocity filter, the DFT and IDFT blocks are assumed to be implemented by means of the split-radix FFT algorithm [124]. According to Table 6.2, for similar SINR improvement, *the proposed 3-D IIR velocity*

Table 6.2: Numbers of nontrivial real multiplications and additions required to process a sample by different 3-D velocity filters.

Filter	Mul.	Add.
Proposed 3-D IIR filter	25.22	59.63
3-D FIR filter [77]	52.84	177.5
3-D IIR filter [29]†	231	440
3-D IIR filter [28]†	1411	1377
3-D optimal FIR filter [27]†	3375	3374

† Obtained from [29].

filter provides significant reduction in the computational complexity compared to the other four 3-D velocity filters. This reduction is achieved mainly due to the fact that *no nontrivial arithmetic operation* is required for the implementations of the 2-D allpass filters and the 2-D allstop filters and the number of 2-D spatial IIR highpass filters ($2M_h$) is considerably small compared to the number of bands (M) of the 1-D temporal modified DFT filter bank. For example, $M_h = 4$ and $M = 128$ for the 3-D IIR velocity filter designed in Section 6.5.1. Compared to the 3-D FIR velocity filter presented in [77], the proposed 3-D IIR velocity filter requires 52.28% and 66.41% less nontrivial real multiplications and additions, respectively.

The proposed 3-D IIR velocity filter and the 3-D FIR velocity filter presented in [77] have similar temporal group delays. However, the temporal group delay of the proposed 3-D IIR velocity filter is considerably higher compared to those of the 3-D velocity filters presented in [27] [28] [29]. For example, the temporal group delay of the 3-D IIR velocity filter designed in Section 6.5.1 is 319 samples whereas those of the 3-D velocity filters presented in [27] [28] [29] may be maximally a few tens of samples. As shown in Section 6.4.3, the temporal group delay of the proposed 3-D IIR velocity filter depends on the number of bands M of the 1-D temporal modified DFT filter bank. Consequently, by employing a less number of bands, the temporal group delay of the proposed 3-D IIR velocity filter may be reduced. However, this will results a less accurate approximation for the ideal passband (See Figures 6.4 and 6.5 (b)). Furthermore, the proposed 3-D IIR velocity filter may provide a poor approximation to the ideal passband at frequencies near the origin compared to the 3-D velocity

filters proposed in [28] [29]. This results as a consequence of the lower number of 2-D spatial IIR highpass filters employed in the cone filter bank. This drawback of the proposed 3-D IIR velocity filter may be partially avoided by employing a 1-D *non-uniform* temporal filter bank, however, at the expense of additional computational complexity.

6.6 Summary

An ultra-low complexity 3-D linear-phase IIR velocity filter is proposed for attenuating moving artifacts in videos while enhancing objects moving on smooth approximately-linear trajectories. The proposed 3-D IIR velocity filter consists of an ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank between two 2-D spatial variable-shift filters. The use of the 2-D spatial shifts has the effect of skewing (or shearing) the region of support of the spectrum of an LT signal so that it is symmetric with respect to the temporal frequency axis. Therefore, the wedge-shaped passband of the ultra-low complexity 3-D wide-angle linear-phase IIR cone filter can be selected to be symmetric with respect to the temporal frequency axis and can be approximated by employing a low-complexity filter bank structure. The ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank is designed by employing a 1-D temporal modified DFT filter bank and 2-D spatial allpass, IIR highpass and allstop filters. The ultra-low complexity is achieved by exploiting the maximal decimation in the temporal dimension, by employing 2-D spatial allpass filters and 1-D spatial IIR lowpass filters to realize the 2-D spatial IIR highpass filters and due to the higher percentage of the 2-D spatial allpass and allstop filters, which require no nontrivial arithmetic operations. Furthermore, the linear phase response is achieved by employing the zero-phase filtering for the 1-D spatial IIR lowpass filters.

The experimental results obtained using a test video signal indicate that the proposed 3-D IIR velocity filter provides significant reduction of the computational complexity compared to the previously proposed 3-D velocity filters for similar SINR improvement. In particular, a typical 3-D linear-phase IIR velocity filter of order $4 \times 4 \times 510$ provides a SINR improvement of 13.86 dB, which is almost independent of the scene complexity, and the proposed 3-D linear-phase IIR velocity filter requires 52.28% and 66.41% less real multiplications and additions, respectively, compared to the 3-D FIR velocity filter presented in [77]. The proposed 3-D IIR velocity filter is employed to attenuate the sunlight flicker patterns in shallow underwater videos. The

experimental results confirm the effectiveness of the proposed method and its robustness to motion-related errors. Considering the ultra-low complexity and robustness to the motion-related errors, the proposed method has an excellent potential to be employed in real-time applications.

The proposed 3-D IIR velocity filter can be directly employed as the 5-D IIR filter $H_{uvt}(\mathbf{z})$ in the design of 5-D IIR depth-velocity filters. Furthermore, the 5-D IIR filters $H_{xu}(\mathbf{z})$ and $H_{yv}(\mathbf{z})$ can be designed by slightly modifying the design method of the proposed filter. In particular, the 2-D spatial variable-shift filters need to be replaced with 1-D spatial variable-shift filters, and the 2-D spatial IIR highpass filters need to be replaced with 1-D spatial IIR highpass filters. Furthermore, the 1-D spatial variable-shift filters should be designed to provide fractional-shifts in addition to the integer shifts.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In this dissertation, the spectra of moving objects in LFVs are analyzed, and it is shown that such moving objects can be enhanced based on their depth and velocity by employing 5-D depth-velocity filters. In the spectral analysis, it is first shown that a Lambertian point source moving with a constant velocity is represented as a 3-D hypersurface of constant value in the corresponding 5-D continuous-domain LFV. When the motion of the Lambertian point source is parallel to the camera plane, i.e. the Lambertian point source moves at a constant depth, the 3-D hypersurface is reduced to a 3-D hyperplane. For this case, the spectrum and its ROS are derived in closed form. The ROS of the spectrum is a plane through the origin in the 5-D continuous frequency domain. The spectral ROS of a continuous-domain LFV corresponding to a Lambertian object moving with a constant velocity and at a constant depth is derived in closed form. It is shown that the ROS is a skewed 3-D hyperfan in the 5-D continuous frequency domain. The degree of skewness of the 3-D hyperfan depends on both velocity and depth of the moving object whereas the angle of the 3-D hyperfan depends on the depth range occupied by the moving object. The 3-D hyperfans corresponding to the ROSs of the spectra of Lambertian objects moving with different constant velocities or at different constant depths do not overlap except at the origin in the 5-D continuous frequency domain. This allows enhancement of such objects based on their depth and velocity by employing 5-D depth-velocity filters. The essential bandwidth of the spectrum is finite along the temporal frequency dimension and, consequently, the corresponding discrete-domain LFV can be generated

with negligible aliasing by employing a sufficiently high temporal sampling rate. The ROS of the spectrum of a discrete-domain LFV corresponding to a Lambertian object moving with a constant velocity and at a constant depth is derived in closed form. Furthermore, the spectral ROS is numerical simulated. The numerically simulated spectral ROS is well agree with the theoretically predicted spectral ROS.

A novel 5-D linear-phase FIR depth-velocity filter having a planar passband in the 5-D frequency domain is proposed. The planar passband of the proposed 5-D FIR depth-velocity filter is realized by cascading three 5-D linear-phase FIR filters having 4-D hyperplanar passbands of appropriate orientations in the 5-D discrete frequency domain. The 5-D FIR filters having 4-D hyperplanar passbands are designed by employing the so-called windowing method [73](ch. 3.3) [74](ch. 5.1). The concept of the 5-D depth-velocity filtering of LFVs to enhance moving objects is experimentally proved by employing the proposed 5-D FIR depth-velocity filter to enhance a numerically generated LFV and an LFV of a real scene, generated using a commercially available Lytro LF camera. The experimental results obtained using the Lytro-LF-camera-based LFV indicate that the proposed 5-D FIR depth-velocity filter is capable of exposing heavily occluded parts of a scene and of attenuating noise significantly. The very high noise attenuation is obtained as a consequence of most of the noise energy reside in the 5-D hypervolume corresponding to the stopband of the 5-D FIR depth-velocity filter.

A novel ultra-low complexity 5-D IIR depth-velocity filter is proposed for enhancing objects moving with constant velocity and at constant depth in LFVs. The proposed 5-D IIR depth-velocity filter is realized by cascading three first-order 5-D IIR filters having 4-D hyperplanar passbands of appropriate orientations. The first-order 5-D IIR filters are designed by appropriately extending the first-order 3-D IIR planar filter design method proposed in [24]. The proposed 5-D IIR depth-velocity filter is practical-BIBO stable. Numerical simulation results indicate that the proposed 5-D IIR depth-velocity filter outperforms the 3-D IIR velocity filters [24] and the 4-D IIR depth filters [44] in enhancing moving objects in LFVs. Furthermore, by employing the Lytro-LF-camera-based LFV, it is shown that the performance of the proposed 5-D IIR depth-velocity filter is comparable to that of the proposed 5-D FIR depth-velocity filter. Most importantly, the proposed 5-D IIR depth-velocity filter requires less than 1% of the arithmetic operations required by the 5-D FIR depth-velocity filter to process a sample. Considering the ultra-low complexity, the proposed 5-D IIR depth-velocity filter has a significant potential to be employed in

real-time applications.

The 5-D depth-velocity filtering technique is extended to a more general case, where objects moving with constant velocity but at non-constant depth are enhanced. First, the spectrum of a Lambertian object moving in an LFV at non-constant depth is analyzed, and it is shown that the ROS of the spectrum can be approximated as a sequence of ROSs, each of which is a skewed 3-D hyperfan, in the 5-D continuous frequency domain. Based on the spectral analysis, a novel ultra-low complexity 5-D IIR adaptive depth-velocity filter is proposed for enhancing objects moving at non-constant depth. The proposed 5-D IIR adaptive depth-velocity filter is realized by cascading three first-order 5-D IIR adaptive filters having time-variant 4-D hyperplanar passbands of appropriate orientations. The three first-order 5-D IIR adaptive filters are designed by appropriately extending the first-order 3-D IIR adaptive planar filter design method presented in [25]. The time-variant coefficients of the three first-order 5-D IIR adaptive filters are derived in closed form. The proposed 5-D IIR adaptive depth-velocity filter is practical-BIBO stable. The performance of the proposed 5-D IIR adaptive depth-velocity filter is confirmed by employing a numerically generated LFV and a Lytro-LF-camera-based LFV of a real scene. Considering the ultra-low complexity and the availability of the closed-form expressions for the time-variant coefficients, the proposed 5-D IIR adaptive depth-velocity filter has a significant potential to be employed in real-time applications.

An ultra-low complexity 3-D linear-phase IIR velocity filter that can be incorporated to design 5-D IIR depth-velocity filters is proposed. The proposed 3-D linear-phase IIR velocity filter consists of an ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank between two 2-D spatial variable-shift filters. The ultra-low complexity 3-D wide-angle linear-phase IIR cone filter bank is designed by employing a 1-D temporal modified DFT filter bank and 2-D spatial allpass, IIR highpass and allstop filters. The ultra-low complexity is achieved by exploiting the maximal decimation in the temporal dimension and by realizing the nonseparable 2-D spatial IIR highpass filters using the 2-D spatial allpass filters and 1-D spatial IIR lowpass filters. Furthermore, the linear phase response is achieved by employing zero-phase filtering for the 1-D spatial IIR lowpass filters. A typical 3-D linear-phase IIR velocity filter of order $4 \times 4 \times 510$, applied to enhance a heavily corrupted test video signal, provides a SINR improvement of 13.86 dB, which is almost independent of the scene complexity. The 3-D linear-phase IIR velocity filter of order $4 \times 4 \times 510$ requires only 26 real multiplications and 60 real additions to process a sample, which

is significantly lower compared to previously reported 3-D FIR and IIR velocity filters. In order to illustrate the performance of the proposed 3-D linear-phase IIR velocity filter in enhancing videos, it is employed to attenuate sunlight flicker patterns in shallow underwater videos. Experimental results confirm the effectiveness of the 3-D linear-phase IIR velocity filter in attenuating the sunlight flicker patterns and its robustness to motion estimation errors.

7.2 Future Work

Potential extensions of the work presented in this dissertation are briefly discussed in this section.

5-D Depth-Velocity Filtering of Plenacoustic Signals

The acoustic counterpart of the 7-D plenoptic function is called the *plenacoustic function* [125]. The plenacoustic function has two forms: omnidirectional and directional. The omnidirectional plenacoustic function is defined as a 4-D function that describes sound pressure at location (x, y, z) in the 3-D space and at time t given an acoustic event in a room [125]. On the other hand, the definition of the directional plenacoustic function is more similar to the definition of the 7-D plenoptic function. That is, the directional plenacoustic function is a 7-D function that describes sound pressure at location (x, y, z) and direction (θ, ϕ) in the 3-D space, for frequency ω and at time t [125]. The concept of the directional plenacoustic function has been exploited in acoustic signal processing for number of applications including source localization and analysis [126] [127]. Considering the similarity of the 7-D plenoptic function and the 7-D directional plenacoustic function, the concept of 5-D depth-velocity filtering may be applied to applications in the field acoustic signal processing, for example, to localization of moving acoustic sources.

Multi-Dimensional FIR Filter Design by Rotation

As mentioned in Section 3.2.2, the passband of the 5-D FIR filter $H_{uvt}(\mathbf{z})$, which is employed to design the 5-D FIR depth-velocity filter proposed in Chapter 3, suffers from undesired ROSs for some values of the design parameters a_u and a_v. This may be avoided by employing a more general ROS for the passband (with more cutoff frequencies) in the derivation of the ideal impulse response. As pointed out in [128]

for the design of 2-D FIR fan filters, this may lead to very complicated expressions and numerical difficulty. Instead, we may employ an appropriate 1-D FIR lowpass prototype filter and coordinated rotation [128] in the design of $H_{uvt}(\mathbf{z})$. Furthermore, this technique may be generalized to design other multi-dimensional filters.

Hardware Architectures for 5-D Depth-Velocity Filters

In real-time applications of the 5-D depth-velocity filters, the processing speed of the software implementations may not be fast enough to provide the demanded throughput by a real-time application. In such a case, the 5-D depth-velocity filters are required to be implemented in fast dedicated hardware such as field-programmable gate arrays. To this end, it is worthwhile to investigate new systolic-array hardware architectures especially for the 5-D IIR depth-velocity filters, considering their ultra-low complexity. Furthermore, the hardware architectures can be optimized with respect to power consumption, throughput and so on to be compatible with the requirements of different applications, for example, to achieve the lowest power consumption (at a reduced throughput) in the case of a battery-powered robot [8] or an underwater autonomous vehicle [17].

Appendix A

Derivation of the Spectrum of a Lambertian Point Source Moving at a Constant Depth

The 5-D continuous-domain Fourier transform of $l_{5\mathrm{C}}(\mathbf{x})$ given in (2.15) can be expressed as

$$\begin{aligned} L_{5\mathrm{C}}(\boldsymbol{\Omega}) &= \int_{x,y,u,v,t=-\infty}^{\infty} \cdots \int l_{5\mathrm{C}}(\mathbf{x})\, \mathrm{e}^{-j\boldsymbol{\Omega}^T\mathbf{x}}\, \mathrm{d}\mathbf{x} \\ &= \int_{x,y,u,v,t=-\infty}^{\infty} \cdots \int l_0\, \delta(mx+u+k_x t+c_x) \\ &\qquad\qquad \times \delta(my+v+k_y t+c_y)\, \mathrm{e}^{-j\boldsymbol{\Omega}^T\mathbf{x}}\, \mathrm{d}\mathbf{x} \\ &= l_0 \int_{t=-\infty}^{\infty} L_{3\mathrm{C}}(\Omega_x,\Omega_u,t)\, L_{3\mathrm{C}}(\Omega_y,\Omega_v,t)\, \mathrm{e}^{-j\Omega_t t}\, \mathrm{d}t, \end{aligned} \tag{A.1}$$

where

$$L_{3\mathrm{C}}(\Omega_x,\Omega_u,t) = \iint_{x,u=-\infty}^{\infty} \delta(mx+u+k_x t+c_x)\, \mathrm{e}^{-j(\Omega_x x+\Omega_u u)}\, \mathrm{d}u\mathrm{d}x \tag{A.2a}$$

$$L_{3\mathrm{C}}(\Omega_y,\Omega_v,t) = \iint_{y,v=-\infty}^{\infty} \delta(my+v+k_y t+c_y)\, \mathrm{e}^{-j(\Omega_y y+\Omega_v v)}\, \mathrm{d}v\mathrm{d}y. \tag{A.2b}$$

Now we evaluate (A.2a) as

$$\begin{aligned} L_{3\mathrm{C}}(\Omega_x, \Omega_u, t) &= \iint_{x,u=-\infty}^{\infty} \delta(mx + u + k_x t + c_x)\, \mathrm{e}^{-j(\Omega_x x + \Omega_u u)}\, \mathrm{d}u\mathrm{d}x \\ &= \int_{x=-\infty}^{\infty} \mathrm{e}^{-j(\Omega_x x + \Omega_u(-mx - k_x t - c_x))}\, \mathrm{d}x \\ &= \mathrm{e}^{j\Omega_u(k_x t + c_x)} \int_{x=-\infty}^{\infty} \mathrm{e}^{-j(\Omega_x - m\Omega_u)x}\, \mathrm{d}x \\ &= 2\pi\, \delta(\Omega_x - m\Omega_u)\, \mathrm{e}^{j\Omega_u(k_x t + c_x)}. \end{aligned} \tag{A.3}$$

Here, we have used the fact that $\delta(mx+u+k_x t+c_x) \neq 0$ only when $mx+u+k_x t+c_x = 0$ [86](ch. 6.2) [87](ch. 2.1). Similarly, $L_{3\mathrm{C}}(\Omega_y, \Omega_v, t)$ can be obtained as

$$L_{3\mathrm{C}}(\Omega_y, \Omega_v, t) = 2\pi\, \delta(\Omega_y - m\Omega_v)\, \mathrm{e}^{j\Omega_v(k_y t + c_y)}. \tag{A.4}$$

By substituting (A.3) and (A.4) into (A.1), $L_{5\mathrm{C}}(\boldsymbol{\Omega})$ can be obtained as

$$\begin{aligned} L_{5\mathrm{C}}(\boldsymbol{\Omega}) &= 4\pi^2 l_0\, \delta(\Omega_x - m\Omega_u)\, \delta(\Omega_y - m\Omega_v)\, \mathrm{e}^{j(\Omega_u c_x + \Omega_v c_y)} \\ &\quad \times \int_{t=-\infty}^{\infty} \mathrm{e}^{-j(\Omega_t - k_x\Omega_u - k_y\Omega_v)t}\, \mathrm{d}t \\ &= 8\pi^3 l_0\, \delta(\Omega_x - m\Omega_u)\, \delta(\Omega_y - m\Omega_v) \\ &\quad \times \delta(\Omega_t - k_x\Omega_u - k_y\Omega_v)\, \mathrm{e}^{j(\Omega_u c_x + \Omega_v c_y)}. \end{aligned} \tag{A.5}$$

Appendix B

Derivation of the Ideal Infinite-Extent Impulse Response of $H_{xu}(\mathbf{z})$

By using the 5-D Fourier series expansion of $H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}})$ [73](ch. 1.3), the ideal infinite-extent impulse response $h^I_{xu}(\mathbf{n})$ of $H_{xu}(\mathbf{z})$ can be obtained as

$$\begin{aligned} h^I_{xu}(\mathbf{n}) &= \frac{1}{(2\pi)^5} \int_{\omega_x,\omega_y,\omega_u,\omega_v,\omega_t=-\pi}^{\pi} \cdots \int H_{xu}(\mathrm{e}^{j\boldsymbol{\omega}}) \mathrm{e}^{j\boldsymbol{\omega}^T\mathbf{n}} \,\mathrm{d}\boldsymbol{\omega} \\ &= \frac{1}{(2\pi)^2} \int_{\omega_u=-\pi}^{\pi} \int_{\omega_x=a_u\omega_u-b_x}^{a_u\omega_u+b_x} \mathrm{e}^{j(\omega_x n_x+\omega_u n_u)} \,\mathrm{d}\omega_x \mathrm{d}\omega_u \\ &\qquad \times \frac{1}{(2\pi)^3} \int_{\omega_y=-\pi}^{\pi} \mathrm{e}^{j\omega_y n_y} \mathrm{d}\omega_y \int_{\omega_v=-\pi}^{\pi} \mathrm{e}^{j\omega_v n_v} \mathrm{d}\omega_v \int_{\omega_t=-\pi}^{\pi} \mathrm{e}^{j\omega_t n_t} \mathrm{d}\omega_t \\ &= g^I_{xu}(n_x, n_u)\, \delta(n_y)\delta(n_v)\delta(n_t), \end{aligned} \tag{B.1}$$

where

$$g^I_{xu}(n_x, n_u) = \frac{1}{(2\pi)^2} \int_{\omega_u=-\pi}^{\pi} \int_{\omega_x=a_u\omega_u-b_x}^{a_u\omega_u+b_x} \mathrm{e}^{j(\omega_x n_x+\omega_u n_u)} \,\mathrm{d}\omega_x \mathrm{d}\omega_u. \tag{B.2}$$

Note that, in deriving (B.1), we employed the 1-D discrete-time Fourier transform pair $\delta(n) \Leftrightarrow 1$ [129](ch. 2.9). After some manipulation and considering the special

cases for which the denominator of $g^I_{xu}(n_x, n_u)$ is zero, the closed-form expressions for $g^I_{xu}(n_x, n_u)$ can be obtained from (B.2) as

$$g^I_{xu}(n_x, n_u) = \frac{b_x}{\pi}, \qquad n_x = 0 \ \text{ and } \ n_u = 0 \tag{B.3a}$$

$$g^I_{xu}(n_x, n_u) = \frac{b_x \sin(n_u\pi)}{n_u\pi^2}, \qquad n_x = 0 \ \text{ and } \ n_u \neq 0 \tag{B.3b}$$

$$g^I_{xu}(n_x, n_u) = \frac{\sin(b_x n_x)}{n_x\pi}, \qquad n_x \neq 0 \ \text{ and } \ a_u n_x + n_u = 0 \tag{B.3c}$$

$$g^I_{xu}(n_x, n_u) = \frac{\sin(b_x n_x)\sin[(a_u n_x + n_u)\pi]}{n_x(a_u n_x + n_u)\pi^2}, \quad n_x \neq 0 \ \text{ and } \ a_u n_x + n_u \neq 0. \tag{B.3d}$$

Appendix C

Derivation of the Ideal Infinite-Extent Impulse Response of $H_{uvt}(\mathbf{z})$

By using the 5-D Fourier series expansion of $H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}})$ [73](ch. 1.3), the ideal infinite-extent impulse response $h^I_{uvt}(\mathbf{n})$ of $H_{uvt}(\mathbf{z})$ can be obtained as

$$\begin{aligned} h^I_{uvt}(\mathbf{n}) &= \frac{1}{(2\pi)^5} \int_{\omega_x,\omega_y,\omega_u,\omega_v,\omega_t=-\pi}^{\pi} \cdots \int H_{uvt}(\mathrm{e}^{j\boldsymbol{\omega}}) \mathrm{e}^{j\boldsymbol{\omega}^T\mathbf{n}} \,\mathrm{d}\boldsymbol{\omega} \\ &= \frac{1}{(2\pi)^3} \int_{\omega_u=-\pi}^{\pi} \int_{\omega_v=-\pi}^{\pi} \int_{\omega_t=a_u\omega_u+a_v\omega_v-b_t}^{\omega_t=a_u\omega_u+a_v\omega_v+b_t} \mathrm{e}^{j(\omega_u n_u+\omega_v n_v+\omega_t n_t)} \,\mathrm{d}\omega_t \mathrm{d}\omega_v \mathrm{d}\omega_u \\ &\qquad \times \frac{1}{(2\pi)^2} \int_{\omega_x=-\pi}^{\pi} \mathrm{e}^{j\omega_x n_x} \mathrm{d}\omega_x \int_{\omega_y=-\pi}^{\pi} \mathrm{e}^{j\omega_y n_y} \mathrm{d}\omega_y \\ &= g^I_{uvt}(n_u, n_v, n_t)\, \delta(n_x)\delta(n_y), \end{aligned} \tag{C.1}$$

where

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{1}{(2\pi)^3} \int_{\omega_u=-\pi}^{\pi} \int_{\omega_v=-\pi}^{\pi} \int_{\omega_t=a_u\omega_u+a_v\omega_v-b_t}^{\omega_t=a_u\omega_u+a_v\omega_v+b_t} \mathrm{e}^{j(\omega_u n_u+\omega_v n_v+\omega_t n_t)} \,\mathrm{d}\omega_t \mathrm{d}\omega_v \mathrm{d}\omega_u. \tag{C.2}$$

Here, in deriving (C.1), the 1-D discrete-time Fourier transform pair $\delta(n) \Leftrightarrow 1$ [129](ch. 2.9) was employed. After some manipulation and considering the special cases for which the denominator of $g^I_{uvt}(n_u, n_v, n_t)$ is zero, the closed-form expressions for $g^I_{uvt}(n_u, n_v, n_t)$ can be obtained from (C.2) as

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{b_t}{\pi}, \qquad n_u = 0, n_v = 0 \text{ and } n_t = 0 \tag{C.3a}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{b_t \sin(n_u\pi)}{n_u\pi^2}, \qquad n_u \neq 0, n_v = 0 \text{ and } n_t = 0 \tag{C.3b}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{b_t \sin(n_v\pi)}{n_v\pi^2}, \qquad n_u = 0, n_v \neq 0 \text{ and } n_t = 0 \tag{C.3c}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{b_t \sin(n_u\pi)\sin(n_v\pi)}{n_u n_v\pi^3}, \qquad n_u \neq 0, n_v \neq 0 \text{ and } n_t = 0 \tag{C.3d}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{\sin(b_t n_t)}{n_t\pi}, \qquad a_u n_t + n_u = 0, a_v n_t + n_v = 0 \text{ and } n_t \neq 0 \tag{C.3e}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{\sin(b_t n_t)\sin[(a_u n_t + n_u)\pi]}{n_t(a_u n_t + n_u)\pi^2}, \qquad a_u n_t + n_u \neq 0, a_v n_t + n_v = 0 \text{ and } n_t \neq 0 \tag{C.3f}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{\sin(b_t n_t)\sin[(a_v n_t + n_v)\pi]}{n_t(a_v n_t + n_v)\pi^2}, \qquad a_u n_t + n_u = 0, a_v n_t + n_v \neq 0 \text{ and } n_t \neq 0 \tag{C.3g}$$

$$g^I_{uvt}(n_u, n_v, n_t) = \frac{\sin(b_t n_t)\sin[(a_u n_t + n_u)\pi]\sin[(a_v n_t + n_v)\pi]}{n_t(a_u n_t + n_u)(a_v n_t + n_v)\pi^3}, \qquad a_u n_t + n_u \neq 0, a_v n_t + n_v \neq 0 \text{ and } n_t \neq 0. \tag{C.3h}$$

Appendix D

Proof of the Near-Perfect Reconstruction of the 3-D Wide-Angle Linear-Phase IIR Cone Filter Bank

The near-perfect reconstruction of the 3-D wide-angle linear-phase IIR cone filter bank, henceforth referred to as the cone filter bank for brevity, can be proved following an approach similar to that employed in [114]. To this end, the input-output relationship of the cone filter bank is expressed, following [29] and [117], as

$$\widehat{I}_{out}(z_x, z_y, z_t) = z_t^{-M/2} \sum_{l=0}^{M/2-1} A_l(z_x, z_y, z_t) \widehat{I}_{in}(z_x, z_y, z_t W_M^{2l}), \tag{D.1}$$

where $\widehat{I}_{in}(z_x, z_y, z_t)$ and $\widehat{I}_{out}(z_x, z_y, z_t)$ are the 3-D $\mathcal{Z}$ transforms of the input signal $\widehat{i}_{in}(n_x, n_y, n_t)$ and the output signal $\widehat{i}_{out}(n_x, n_y, n_t)$ of the cone filter bank, respectively, and

$$A_l(z_x, z_y, z_t) = \frac{1}{M} \sum_{k=0}^{M-1} H_k(z_t W_M^{2l}) G_k(z_x, z_y) F_k(z_t),$$
$$l = 0, 1, \ldots, (M/2 - 1). \tag{D.2}$$

The output signal $\widehat{I}_{out}(z_x, z_y, z_t)$, given in (D.1), consists of the desired *alias-free* component $D(z_x, z_y, z_t)$ corresponding to $l = 0$ and given by

$$D(z_x, z_y, z_t) = \frac{z_t^{-M/2}}{M} \sum_{k=0}^{M-1} H_k(z_t) G_k(z_x, z_y) F_k(z_t) \widehat{I}_{in}(z_x, z_y, z_t), \tag{D.3}$$

and the undesired *aliased* component $A(z_x, z_y, z_t)$ having $(M/2 - 1)$ aliasing terms corresponding to $l = 1, 2, \ldots, (M/2 - 1)$ and given by

$$\begin{aligned} A(z_x, z_y, z_t) &= \frac{z_t^{-M/2}}{M} \sum_{l=1}^{M/2-1} \left(\sum_{k=0}^{M-1} H_k(z_t W_M^{2l}) G_k(z_x, z_y) F_k(z_t) \right) \widehat{I}_{in}(z_x, z_y, z_t W_M^{2l}) \\ &= z_t^{-M/2} \sum_{l=1}^{M/2-1} A_l(z_x, z_y, z_t) \widehat{I}_{in}(z_x, z_y, z_t W_M^{2l}). \end{aligned} \tag{D.4}$$

In order to achieve the *perfect reconstruction*, the magnitude and the phase responses of the cone filter bank should be *unity* and *linear*, respectively, inside the passband. Moreover, the aliased component $A(z_x, z_y, z_t)$ should be *zero* inside the passband of the cone filter bank. In what follows, we show that the cone filter bank approximately satisfies the above three conditions and, consequently, achieves the *near-perfect reconstruction*. To this end, (D.2) is rewritten, by employing (6.6), as

$$A_l(z_x, z_y, z_t) = \frac{z_t^{-N_P} W_M^{-lN_P}}{M} \sum_{k=0}^{M-1} P(z_t W_M^{2l+k}) G_k(z_x, z_y) P(z_t W_M^k). \tag{D.5}$$

Because the prototype filter $P(z_t)$, designed in Section 6.4.1, has a stopband edge less than $2\pi/M$ and a fairly high stopband attenuation, it can be shown that [130]

$$\begin{aligned} P(z_t W_M^{2l+k}) P(z_t W_M^k) \approx 0, \qquad & l = 1, \ldots, (M/2 - 1), \\ & k = 0, 1, \ldots, (M - 1). \end{aligned} \tag{D.6}$$

By substituting (D.6) into (D.5), we can show that

$$A_l(z_x, z_y, z_t) \approx 0, \qquad l = 1, 2, \ldots, (M/2 - 1). \tag{D.7}$$

Furthermore, by substituting (D.7) into (D.4), it can be shown that $||A(z_x, z_y, z_t)||_2$ can be made *arbitrarily small* in relation to $||\widehat{I}_{in}(z_x, z_y, z_t)||_2$. In this case, the cone

filter bank is *almost alias free*, and the transfer function $C(z_x, z_y, z_t)$ can be expressed as

$$\begin{aligned} C(z_x, z_y, z_t) &\approx \frac{z_t^{-M/2}}{M} \sum_{k=0}^{M-1} H_k(z_t) G_k(z_x, z_y) F_k(z_t) \\ &\approx \frac{z_t^{-(N_P+M/2)}}{M} \sum_{k=0}^{M-1} \left[P(z_t W_M^k)\right]^2 G_k(z_x, z_y). \end{aligned} \tag{D.8}$$

Considering the fact that $G_k(z_x, z_y) = 0$ for all the 2-D spatial allstop filters, (D.8) can be rewritten as

$$C(z_x, z_y, z_t) \approx \frac{z_t^{-(N_P+M/2)}}{M} \sum_{k\in\mathcal{C}} \left[P(z_t W_M^k)\right]^2 G_k(z_x, z_y), \tag{D.9}$$

where $\mathcal{C} = \{0, 1, \ldots, M_h, M-M_h, M-(M_h+1), \ldots, M-1\}$. Because $P(\mathrm{e}^{j\omega_t}) \approx \sqrt{M}$ and $G_k(\mathrm{e}^{j\omega_x}, \mathrm{e}^{j\omega_y}) \approx 1$, $k \in \mathcal{C}$, inside their respective passbands, the frequency response of the cone filter bank can be obtained from (D.9) as

$$C(\mathrm{e}^{j\omega_x}, \mathrm{e}^{j\omega_y}, \mathrm{e}^{j\omega_t}) \approx z_t^{-(N_P+M/2)}, \qquad (\omega_x, \omega_y, \omega_t) \in \mathcal{P}, \tag{D.10}$$

where $\mathcal{P}$ denotes the region of support corresponding to the passband of the cone filter bank in the 3-D frequency domain. It is clear from (D.10) that the magnitude response $|C(\mathrm{e}^{j\omega_x}, \mathrm{e}^{j\omega_y}, \mathrm{e}^{j\omega_t})|$ of the cone filter bank is approximately *unity* inside the passband whereas the phase response $\angle C(\mathrm{e}^{j\omega_x}, \mathrm{e}^{j\omega_y}, \mathrm{e}^{j\omega_t})$ of the cone filter bank is approximately *linear* inside the passband. Therefore, the cone filter bank achieves the *near-perfect reconstruction*. □

Bibliography

[1] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Computation Models of Visual Processing* (M. Landy and J. A. Movshon, eds.), pp. 3–20, Cambridge, MA: MIT Press, 1991.

[2] C. Zhang and T. Chen, "A survey on image-based rendering—representation, sampling and compression," *Signal Process.: Image Commun.*, vol. 19, pp. 1–28, Jan. 2004.

[3] H.-Y. Shum, S. B. Kang, and S.-C. Chan, "Survey of image-based representations and compression techniques," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 1020–1037, Nov. 2003.

[4] R. S. Higa, Y. Iano, R. B. Leite, R. F. L. Chavez, and R. Arthur, "Employing light field cameras in surveillance: An analysis of light field cameras in a surveillance scenario," *3D Research*, vol. 5, pp. 1–11, Mar. 2014.

[5] H. Qian, X. Wu, and Y. Xu, *Intelligent Surveillance Systems.* New York: Springer, 2011.

[6] S.-C. Chan, S. Zhang, J.-F. Wu, H.-J. Tan, J. Ni, and Y. Hung, "On the hardware/software design and implementation of a high definition multiview video surveillance system," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, pp. 248–262, June 2013.

[7] W. Limprasert, A. Wallace, and G. Michaelson, "Real-time people tracking in a camera network," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, pp. 263–271, June 2013.

[8] F. Dong, S.-H. Ieng, X. Savatier, R. Etienne-Cummings, and R. Benosman, "Plenoptic cameras in real-time robotics," *Int. J. Robot. Research*, vol. 32, pp. 206–217, Feb. 2013.

[9] D. G. Dansereau, *Plenoptic Signal Processing for Robust Vision in Field Robotics*. PhD thesis, Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, NSW, Australia, 2014.

[10] L. Belcour, C. Soler, K. Subr, N. Holzschuch, and F. Durand, "5D covariance tracing for efficient defocus and motion blur," *ACM Trans. Graph.*, vol. 32, pp. 31:1–31:18, June 2013.

[11] K. Vaidyanathan, J. Munkberg, P. Clarberg, and M. Salvi, "Layered light field reconstruction for defocus blur," *ACM Trans. Graph.*, vol. 34, pp. 23:1–23:12, Feb. 2015.

[12] R. Ronfard and G. Taubin, eds., *Image and Geometry Processing for 3-D Cinematography*. Heidelberg: Springer, 2010.

[13] C. Zhu, Y. Zhao, L. Yu, and M. Tanimoto, eds., *3D-TV System with Depth-Image-Based Rendering*. New York: Springer, 2013.

[14] A. Madanayake, C. Wijenayake, D. G. Dansereau, T. K. Gunaratne, L. T. Bruton, and S. B. Williams, "Multidimensional (MD) circuits and systems for emerging applications including cognitive radio, radio astronomy, robot vision and imaging," *IEEE Circuits Syst. Mag.*, vol. 13, no. 1, pp. 10–43, 2013.

[15] D. G. Dansereau and S. B. Williams, "Seabed modeling and distractor extraction for mobile auvs using light field filtering," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 1634–1639, 2011.

[16] B. Ouyang, F. Dalgleish, A. Vuorenkoski, W. Britton, B. Ramos, and B. Metzger, "Visualization and image enhancement for multistatic underwater laser line scan system using image-based rendering," *IEEE J. Ocean. Eng.*, vol. 38, pp. 566–580, July 2013.

[17] O. Pizarro, S. B. Williams, M. V. Jakuba, M. Johnson-Roberson, I. Mahon, M. Bryson, D. Steinberg, A. Friedman, D. Dansereau, N. Nourani-Vatani, D. Bongiorno, M. Bewley, A. Bender, N. Ashan, and B. Douillard, "Benthic monitoring with robotic platformsthe experience of australia," in *Proc. IEEE Int. Underwater Technol. Symp.*, pp. 1–10, 2013.

[18] L. T. Bruton, N. R. Bartley, and Z. Q. Liu, "On the classification of moving objects in image sequences using 3D adaptive recursive tracking filters and neural networks," in *Proc. 29th Asilomar Conf. Signals, Syst., Comput.*, vol. 2, pp. 1006–1010, 1995.

[19] J.-Y. Chang, C.-C. Cheng, S.-Y. Chien, and L.-G. Chen, "Relative depth layer extraction for monoscopic video by use of multidimensional filter," in *Proc. IEEE Int. Conf. Multimedia and Expo*, pp. 221–224, 2006.

[20] S. Schauland, J. Velten, and A. Kummert, "Multidimensional linear shift invariant velocity filters for vision-based automotive applications," in *Proc. Int. Workshop Multidim. (nD) Syst.*, pp. 83–87, 2007.

[21] S. Schauland, J. Velten, and A. Kummert, "Detection of moving objects in image sequences using 3D velocity filters," *Int. J. Appl. Math. Comput. Sci.*, vol. 18, pp. 21–31, Mar. 2008.

[22] S. Schauland, J. Velten, and A. Kummert, "Motion-based object detection for automotive applications using multidimensional wave digital filters," in *Proc. IEEE Veh. Technol. Conf. (Spring)*, pp. 2700–2704, 2008.

[23] H. L. Kennedy, "Multidimensional digital filters for point-target detection in cluttered infrared scenes," *J. Electron. Imag.*, vol. 23, pp. 063019–1–063019–11, Nov./Dec. 2014.

[24] L. T. Bruton and N. R. Bartley, "Three-dimensional image processing using the concept of network resonance," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 664–672, July 1985.

[25] L. T. Bruton and N. Bartley, "The enhancement and tracking of moving objects in digital images using adaptive three-dimensional recursive filters," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 604–612, June 1986.

[26] S.-C. Pei, W.-Y. Kuo, and W.-T. Huang, "Tracking moving objects in image sequences using 1-D trajectory filter," *IEEE Signal Process. Lett.*, vol. 13, pp. 13–16, Jan. 2006.

[27] K. Kondo and N. Hamada, "Design of optimal filter for detecting linear trajectory signals utilizing object shape and velocity vector," *Electron. and Commun. in Japan, Part I*, vol. 83, pp. 42–51, Feb. 2000.

[28] B. Anderson and L. T. Bruton, "Non-uniform bandwidth frequency-planar (NUB-FP) filter banks," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. I–809–I–812, 2002.

[29] B. Kuenzle and L. T. Bruton, "3-D IIR filtering using decimated DFT-polyphase filter bank structures," *IEEE Trans. Circuits Syst. I*, vol. 53, pp. 394–408, Feb. 2006.

[30] T. Schwerdtfeger, J. Velten, and A. Kummert, "A multidimensional wave digital filter bank for video-based motion analysis," *Multidim. Syst. Signal Processing*, vol. 25, pp. 295–311, Apr. 2014.

[31] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 31–42, 1996.

[32] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 43–54, 1996.

[33] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Tech. Rep. CTSR 2005-02, Stanford Univ., Stanford, CA, 2005.

[34] R. Ng, "Fourier slice photography," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 735–744, 2005.

[35] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 69–1–69–12, 2007.

[36] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *Proc. IEEE Int. Conf. Comput. Photogr.*, pp. 1–8, 2009.

[37] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Linear volumetric focus for light field cameras," *ACM Trans. Graph.*, vol. 34, pp. 15:1–15:20, Feb. 2015.

[38] V. Vaish, B. Wilburn, N. Joshi, and M. Levoy, "Using plane + parallax for calibrating dense camera arrays," in *Proc. IEEE Conf. Comput. Vision and Pattern Recogn.*, vol. 1, pp. I–2–I–9, 2004.

[39] V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, and S. B. Kang, "Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures," in *Proc. IEEE Conf. Comput. Vision and Pattern Recogn.*, vol. 2, pp. 2331–2338, 2006.

[40] N. Joshi, S. Avidan, W. Matusik, and D. J. Kriegman, "Synthetic aperture tracking: Tracking through occlusions," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 1–8, 2007.

[41] A. Madanayake, R. Wimalagunarathne, D. G. Dansereau, R. J. Cintra, and L. T. Bruton, "VLSI architecture for 4-D depth filtering," *Signal, Image and Video Process.*, pp. 1–10, 2013.

[42] T. Yang, Y. Zhang, X. Tong, X. Zhang, and R. Yu, "A new hybrid synthetic aperture imaging model for tracking and seeing people through occlusion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, pp. 1461–1475, Sep. 2013.

[43] A. Isaksen, L. McMillan, and S. J. Gortler, "Dynamically reparameterized light fields," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 297–306, 2000.

[44] D. Dansereau and L. Bruton, "A 4D frequency-planar IIR filter and its application to light field processing," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 4, pp. IV–476–IV–479, 2003.

[45] V. Vaish, G. Garg, E.-V. Talvala, E. Antunez, B. Wilburn, M. Horowitz, and M. Levoy, "Synthetic aperture focusing using a shear-warp factorization of the viewing transform," in *Proc. IEEE Conf. Comput. Vision and Pattern Recogn.*, pp. 129–129, 2005.

[46] D. Dansereau and L. T. Bruton, "A 4-D dual-fan filter bank for depth filtering in light fields," *IEEE Trans. Signal Process.*, vol. 55, pp. 542–549, Feb. 2007.

[47] D. G. Dansereau, D. L. Bongiorno, O. Pizarro, and S. B. Williams, "Light field image denoising using a linear 4D frequency-hyperfan all-in-focus filter," in *Proc. SPIE Comput. Imag. XI*, vol. 8657, pp. 86570P–1–86570P–14, 2013.

[48] A. Madanayake, R. Wimalagunaratne, D. G. Dansereau, and L. T. Bruton, "Design and FPGA-implementation of 1st-order 4D IIR frequency-hyperplanar digital filters," in *Proc. IEEE 54th Int. Midwest Symp. Circuits Syst.*, pp. 1–4, 2011.

[49] R. Wimalagunarathne, A. Madanayake, D. G. Dansereau, and L. T. Bruton, "A systolic-array architecture for first-order 4-D IIR frequency-planar digital filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 3069–3072, 2012.

[50] R. Wimalagunarathne, C. Wijenayake, A. Madanayake, D. G. Dansereau, and L. T. Bruton, "Integral form 4-D light field filters using Xilinx FPGAs and 45 nm CMOS technology," *Multidim. Syst. Signal Process.*, vol. 26, no. 1, pp. 47–65, 2015.

[51] C. Hahne and A. Aggoun, "Embedded FIR filter design for real-time refocusing using a standard plenoptic video camera," in *Proc. SPIE Electron. Imag.*, vol. 9023, pp. 902305–1–902305–12, 2014.

[52] B. S. Wilburn, M. Smulski, H.-H. K. Lee, and M. A. Horowitz, "Light field video camera," in *Proc. SPIE Media Processors*, vol. 4674, pp. 29–36, 2002.

[53] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance imaging using large camera arrays," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 765–776, 2005.

[54] S.-C. Chan, K.-T. Ng, Z.-F. Gan, K.-L. Chan, and H.-Y. Shum, "The plenoptic video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 1650–1659, Dec. 2005.

[55] A. Akin, O. Cogal, K. Seyid, H. Afshari, A. Schmid, and Y. Leblebici, "Hemispherical multiple camera system for high resolution omni-directional light field imaging," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 3, pp. 137–144, June 2013.

[56] K. Venkataraman, D. Lelescu, J. Duparré, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar, "PiCam: An ultra-thin high performance monolithic camera array," *ACM Trans. Graph.*, vol. 32, pp. 166:1–166:13, Nov. 2013.

[57] S. Tambe, A. Veeraraghavan, and A. Agrawal, "Towards motion aware light field video for dynamic scenes," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 1009–1016, 2013.

[58] D. C. Schedl, C. Birklbauer, and O. Bimber, "Coded exposure HDR light-field video recording," *Comput. Graph. Forum*, vol. 33, pp. 33–42, May 2014.

[59] J. Neumann and C. Fermüller, "Plenoptic video geometry," *The Visual Comput.*, vol. 19, pp. 395–404, Oct. 2003.

[60] J. Neumann, C. Fermuller, and Y. Aloimonos, "Polydioptric camera design and 3D motion estimation," in *Proc. IEEE Conf. Comput. Vision and Pattern Recogn.*, vol. 2, pp. II–294–II–301, 2003.

[61] D. G. Dansereau, I. Mahon, O. Pizarro, and S. B. Williams, "Plenoptic flow: Closed-form visual odometry for light field cameras," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, pp. 4455–4462, 2011.

[62] S.-C. Chan and H.-Y. Shum, "A spectral analysis for light field rendering," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, pp. 25–28, 2000.

[63] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic sampling," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 307–318, 2000.

[64] A. Levin, S. W. Hasinoff, P. Green, F. Durand, and W. T. Freeman, "4D frequency analysis of computational cameras for depth of field extension," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 97:1–97:14, 2009.

[65] C. Zhang and T. Chen, "Spectral analysis for sampling image-based rendering data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 1038–1050, Nov. 2003.

[66] M. N. Do, D. Marchand-Maillet, and M. Vetterli, "On the bandwidth of the plenoptic function," *IEEE Trans. Image Process.*, vol. 21, pp. 708–717, Feb. 2012.

[67] C. Gilliam, P. L. Dragotti, and M. Brookes, "A closed-form expression for the bandwidth of the plenoptic function under finite field of view constraints," in *Proc. IEEE Int. Conf. Image Process.*, pp. 3965–3968, 2010.

[68] C. Gilliam, P.-L. Dragotti, and M. Brookes, "On the spectrum of the plenoptic function," *IEEE Trans. Image Process.*, vol. 23, pp. 502–516, Feb. 2014.

[69] P. Zhou, L. Yu, and G. Zhong, "The non-lambertian reflection in plenoptic sampling," in *Proc. IEEE Int. Conf. Image Process.*, pp. 2154–2157, 2013.

[70] P. Zhou, L. Yu, and C. Pak, "The spectrum broadening in the plenoptic function," in *Proc. Int. Conf. Internet Multimedia Comput. Service*, pp. 130–135, 2014.

[71] H. Zhang, "A Fourier theory of the light field with the resolution of the sampling camera," in *Proc. Int. Conf. Internet Multimedia Comput. Service*, pp. 1–5, 2015.

[72] C. U. S. Edussooriya, D. G. Dansereau, L. T. Bruton, and P. Agathoklis, "Five-dimensional depth-velocity filtering for enhancing moving objects in light field videos," *IEEE Trans. Signal Process.*, vol. 63, pp. 2151–2163, Apr. 2015.

[73] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[74] J. W. Woods, *Multidimensional Signal, Image, and Video Processing and Coding*. Academic Press, 2006.

[75] C. U. S. Edussooriya, L. T. Bruton, and P. Agathoklis, "A 5-D IIR depth-velocity filter for enhancing objects moving on linear-trajectories in light field videos," in *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 2381–2384, 2015.

[76] C. U. S. Edussooriya, L. T. Bruton, and P. Agathoklis, "Enhancing moving objects in light field videos using 5-D IIR adaptive depth-velocity filters," in *Proc. IEEE Pacific Rim Conf. Commun., Comput., Signal Process.*, pp. 1–5, 2015.

[77] C. U. S. Edussooriya, L. T. Bruton, and P. Agathoklis, "A low-complexity 3D spatio-temporal FIR filter for enhancing linear trajectory signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 1165–1169, 2014.

[78] C. U. S. Edussooriya, L. T. Bruton, and P. Agathoklis, "Velocity filtering for attenuating moving artifacts in videos using an ultra-low complexity 3-D linear-phase IIR filter," *Multidim. Syst. Signal Process.*, pp. 1–20, 2015.

[79] R. Szeliski, *Computer Vision: Algorithms and Applications*. London: Springer, 2011.

[80] H.-Y. Shum, S.-C. Chan, and S. B. Kang, *Image-Based Rendering*. New York: Springer, 2007.

[81] E. Camahort, A. Lerios, and D. Fussell, "Uniformly sampled light fields," Tech. Rep. TR98-09, Univ. Texas at Austin, Austin, TX, 1998.

[82] I. Ihm, S. Park, and R. K. Lee, "Rendering of spherical light fields," in *Proc. Pacific Conf. Comput. Graph. Appl.*, pp. 59–68, 1997.

[83] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge: Cambridge University Press, 7th ed., 1999.

[84] S. F. Ray, *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video, and Electronic Imaging*. Oxford: Focal Press, 2nd ed., 1994.

[85] L. Hogben, ed., *Handbook of Linear Algebra*. Boca Raton, FL: Chapman & Hall/CRC, 2007.

[86] A. Antoniou, *Digital Signal Processing: Signals, Systems and Filters*. NY: McGraw-Hill, 2006.

[87] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications (Eastern Economy Edition)*. New Delhi: Prentice-Hall, 3rd ed., 1996.

[88] Y. Wang, J. Ostermann, and Y.-Q. Zhang, *Video Processing and Communications*. Upper Saddle River, NJ: Prentice-Hall, 2002.

[89] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *Int. J. Comput. Vision*, vol. 1, no. 1, pp. 7–55, 1987.

[90] J. S. Lim, *Two-Dimensional Signal and Image Processing*. Upper Saddle River, NJ: Prentice-Hall, 1990.

[91] Lytro, Inc., "Technical specifications." [Online], `https://www.lytro.com/camera/specs/`, 2014.

[92] T. Georgiev, Z. Yu, A. Lumsdaine, and S. Goma, "Lytro camera technology: theory, algorithms, performance analysis," in *Proc. SPIE Electron. Imag.*, pp. 86671J–1–86671J–10, 2013.

[93] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in *Proc. IEEE Conf. Comput. Vision and Pattern Recogn.*, pp. 1027–1034, 2013.

[94] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Trans. Graph.*, vol. 18, pp. 1–34, Jan. 1999.

[95] F. Durand, N. Holzschuch, C. Soler, E. Chan, and F. X. Sillion, "A frequency analysis of light transport," in *Proc. Annu. Conf. Comput. Graph. (SIGGRAPH)*, pp. 1115–1126, 2005.

[96] C.-K. Liang, Y.-C. Shih, and H. H. Chen, "Light field analysis for modeling image formation," *IEEE Trans. Image Process.*, vol. 20, pp. 446–460, Feb. 2011.

[97] L. T. Bruton, "Three-dimensional cone filter banks," *IEEE Trans. Circuits Syst. I*, vol. 50, pp. 208–216, Feb. 2003.

[98] L. T. Bruton and N. R. Bartley, "Highly selective three-dimensional recursive beam filters using intersecting resonant planes," *IEEE Trans. Circuits Syst.*, vol. CAS-30, pp. 190–193, Mar. 1983.

[99] D. M. Goodman, "Some difficulties with double bilinear transformation in 2-D recursive filter design," *Proc. IEEE*, vol. 66, pp. 796–797, July 1978.

[100] D. Goodman, "Some stability properties of two-dimensional linear shift-invariant digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 201–208, Apr. 1977.

[101] P. Agathoklis and L. T. Bruton, "Practical-BIBO stability of n-dimensional discrete systems," *IEE Proc. G (Electron. Circuits and Syst.)*, vol. 130, pp. 236–242, Dec. 1983.

[102] S. Schauland, J. Velten, A. Kummert, and K. Galkowski, "Insufficiencies of practical BIBO stable nD systems," *Multidim. Syst. Signal Process.*, vol. 25, pp. 3–15, Jan. 2014.

[103] J.-L. Starck, F. Murtagh, E. J. Candes, and D. L. Donoho, "Gray and color image contrast enhancement by the curvelet transform," *IEEE Trans. Image Process.*, vol. 12, pp. 706–717, June 2003.

[104] J.-H. Han, S. Yang, and B.-U. Lee, "A novel 3-D color histogram equalization method with uniform 1-D gray scale histogram," *IEEE Trans. Image Process.*, vol. 20, pp. 506–512, Feb. 2011.

[105] T. Celik and T. Tjahjadi, "Automatic image equalization and contrast enhancement using gaussian mixture modeling," *IEEE Trans. Image Process.*, vol. 21, pp. 145–156, Jan. 2012.

[106] S.-W. Jung, "Image contrast enhancement using color and depth histograms," *IEEE Signal Process. Lett.*, vol. 21, pp. 382–385, Apr. 2014.

[107] G. Jiang, C. Y. Wong, S. C. F. Lin, M. A. Rahman, T. R. Ren, N. Kwok, H. Shi, Y.-H. Yu, and T. Wu, "Image contrast enhancement with brightness preservation using an optimal gamma correction and weighted sum approach," *J. Mod. Opt.*, vol. 62, no. 7, pp. 536–547, 2015.

[108] Y. Y. Schechner and N. Karpel, "Attenuating natural flicker patterns," in *Proc. MTTS/IEEE Oceans*, vol. 3, pp. 1262–1268, 2004.

[109] N. Gracias, S. Negahdaripour, L. Neumann, R. Prados, and R. Garcia, "A motion compensated filtering approach to remove sunlight flicker in shallow water images," in *Proc. Oceans*, pp. 1–7, 2008.

[110] A. Shihavuddin, N. Gracias, and R. García, "Online sunflicker removal using dynamic texture prediction," in *Proc. Int. Joint Conf. Comput. Vision, Imag. and Comput. Graph. Theory, Appl. (VISAPP)*, pp. 161–167, 2012.

[111] B. Kuenzle and L. T. Bruton, "A novel low-complexity spatio-temporal ultra wide-angle polyphase cone filter bank applied to sub-pixel motion discrimination," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 3, pp. 2397–2400, 2005.

[112] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, pp. 4–24, Apr. 1988.

[113] C. U. S. Edussooriya, L. T. Bruton, M. A. Naeini, and P. Agathoklis, "Using 1-D variable fractional-delay filters to reduce the computational complexity of 3-D broadband multibeam beamformers," *IEEE Trans. Circuits Syst. II*, vol. 61, pp. 279–283, Apr. 2014.

[114] C. U. S. Edussooriya, L. T. Bruton, P. Agathoklis, and T. K. Gunaratne, "Low-complexity maximally-decimated multirate 3D spatio-temporal FIR cone and frustum filters," *IEEE Trans. Circuits Syst. I*, vol. 60, pp. 1845–1856, July 2013.

[115] C. U. S. Edussooriya, L. T. Bruton, and P. Agathoklis, "Under-decimated 3D FIR space-time cone filters using DFT polyphase filter banks for attenuation of radio frequency interference," in *Proc. IEEE Int. Midwest Symp. Circuits Syst.*, pp. 1–4, 2011.

[116] N. J. Fliege, "Computational efficiency of modified DFT polyphase filter banks," in *Proc. 27th Asilomar Conf. Signals, Syst., Comput.*, vol. 2, pp. 1296–1300, 1993.

[117] T. Karp and N. J. Fliege, "Modified DFT filter banks with perfect reconstruction," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 1404–1414, Nov. 1999.

[118] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[119] H. S. Malvar, "Extended lapped transforms: properties, applications, and fast algorithms," *IEEE Trans. Signal Process.*, vol. 40, pp. 2703–2714, Nov. 1992.

[120] M. G. Bellanger and J. L. Daguet, "TDM-FDM transmultiplexer: Digital polyphase and FFT," *IEEE Trans. Commun.*, vol. COM-22, pp. 1199–1204, Sep. 1974.

[121] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[122] H. V. Sorensen, D. L. Jones, M. T. Heideman, and C. S. Burrus, "Real-valued fast fourier transform algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 849–863, June 1987.

[123] R. E. Blahut, *Fast Algorithms for Signal Processing*. NY: Cambridge University Press, 2010.

[124] P. Duhamel and H. Hollmann, "'Split radix' FFT algorithm," *Electron. Lett.*, vol. 20, pp. 14–16, Jan. 1984.

[125] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," *IEEE Trans. Signal Process.*, vol. 54, pp. 3790–3804, Oct. 2006.

[126] D. Markovic, F. Antonacci, A. Sarti, and S. Tubaro, "Soundfield imaging in the ray space," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, pp. 2493–2505, Dec. 2013.

[127] D. Markovic, F. Antonacci, A. Sarti, and S. Tubaro, "Multiview soundfield imaging in the projective ray space," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, pp. 1054–1067, June 2015.

[128] S.-C. Pei and S.-B. Jaw, "Two-dimensional general fan-type FIR digital filter design," *Signal Process.*, vol. 37, pp. 265–274, May 1994.

[129] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 2nd ed., 1999.

[130] Y.-P. Lin and P. P. Vaidyanathan, "Application of DFT filter banks and cosine modulated filter banks in filtering," in *Proc. IEEE Asia-Pacific Conf. Circuits Syst.*, pp. 254–259, 1994.