

COMPUTATIONAL IMAGING FOR MINIATURE CAMERAS

by

Basel Salahieh

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Basel Salahieh, titled Computational Imaging For Miniature Cameras and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: September 3, 2015
Jeffrey J. Rodriguez

_____ Date: September 3, 2015
Rongguang Liang

_____ Date: September 3, 2015
Ali Bilgin

_____ Date: September 3, 2015
Thomas D. Milster

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: September 3, 2015
Co-Dissertation Director: Jeffrey J. Rodriguez

_____ Date: September 3, 2015
Co-Dissertation Director: Rongguang Liang

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Basel Salahieh

ACKNOWLEDGEMENTS

My sincere gratitude goes first to my co-advisors Dr. Rongguang Liang and Dr. Jeffrey J. Rodriguez who have been very supportive and motivating throughout my dissertation work and for being there for me when I needed it the most. Also, I would like to kindly express my appreciation to Dr. Amit Ashok and Dr. Mark Neifeld for setting me on the computational imaging path through the first phase of my Ph.D. program. In addition, I would like to thank the members of my dissertation committee for their useful comments and discussions. The experience that I have gained during the course of my research is certainly invaluable in preparing me for a career in this field.

I would like to thank each and every member of OCPL, SaIL, and applied optics groups that I have had a chance to interact with for all their help and guidance. Also, I would like to thank Google's ATAP group for their kind support of the image deconvolution project.

My heartfelt thanks go to my wife, my parents, and my overseas family who have provided me with their encouragement and support throughout this journey. Without their love and support, it would have been impossible to be where I am today. Many thanks to all my friends in the Tucson community for all the wonderful times spent together.

Last but foremost, I thank Allah for all the blessings he has given me through the wonderful opportunities, the inspiring people, and the strength to pursue my ambitions.

DEDICATION

This dissertation is dedicated to my beloved country, Syria!

TABLE OF CONTENTS

LIST OF FIGURES	9
LIST OF ABBREVIATIONS.....	13
Abstract.....	14
1. Introduction to Computational Imaging.....	16
1.1. Basics of Imaging Systems	16
1.2. The Design Factors of Imaging Systems	17
1.3. Limitations of Miniature Imaging Systems	18
1.4. Computational Imaging as a Solution.....	20
1.4.1. Motivation to Develop a Realistic Superresolution Technique	21
1.4.2. Motivation to Develop a Full-Focus Imaging Technique.....	22
1.4.3. Motivation to Develop a Depth Acquisition Technique for HDR Objects	23
1.5. Organization of the Dissertation	24
2. Direct Superresolution for Realistic Image Reconstruction	26
2.1. Introduction.....	26
2.2. The Resolvability of Spatial Frequencies in the Imaging Systems.....	28
2.3. Mathematics of Multi-Shift Images for Superresolution	29
2.4. Multi-Shift Imaging Systems.....	30
2.5. Parameters Estimate from the LR Images	32
2.5.1. Shift Estimate.....	32
2.5.2. Blurring Estimate	34
2.5.3. Noise Estimate	35
2.6. Direct Superresolution Algorithm.....	35
2.7. Adaptive Frequency-Based Filtering Scheme	39
2.8. Simulation Results	42
2.8.1. Evaluating Various SR Techniques under Ideal Imaging Conditions .	43
2.8.2. Studying the Adaptive Frequency-Based Filtering Scheme at Noisy Environments	45
2.8.3. Assessing the Impact of Blurring When Imaging Through Different Observation Models	47
2.8.4. Validating the Replicating Boundary Assumption Made in the Direct SR Technique.....	51
2.8.5. Evaluating the Impact of the Shift Misestimates on the Direct SR Restorations.....	52

2.8.6. Evaluating the Performance at a Combination of Non-idealities	54
2.9. Conclusion and Future Work	56
2.9.1. The Motion Blur and Dynamic Scenes	57
2.9.2. Shift Type.....	58
2.9.3. The Chromatic and Off-Axis Aberrations	58
2.9.4. The Scanning Filtering Path and the Selection Criteria.....	58
2.9.5. Categorized Training and Reference SR Techniques	59
3. Computational Depth-Based Deconvolution for Full-Focus Imaging.....	61
3.1. Introduction.....	61
3.2. Depth-Based Deconvolution Technique	63
3.3. Addressing the Ringing Artifacts.....	65
3.3.1. Tackling the Boundary Artifacts.....	65
3.3.2. Tackling the Edges Artifacts.....	66
3.3.3. Tackling the Depth-Transition Artifacts.....	67
3.4. Local Focus Decision Metrics	69
3.5. Simulation Results	70
3.5.1. Impact of Pre-Processing the Boundaries of the Blurred Images	72
3.5.2. Impact of Adaptive Regularization for Planar Objects set at Different Depth Planes	74
3.5.3. Impact of Block-Wise Deconvolution	77
3.5.4. Impact of Depth-Based Masking Approach	78
3.5.5. Comparison of Full-Focus Deconvolution Results after Stitching.....	79
3.6. Conclusion and Future Work.....	82
4. Multi-Polarization Fringe Projection Imaging for High Dynamic Range Objects	84
4.1. Introduction.....	84
4.2. Multi-Polarization Fringe Projection Imaging Algorithm	85
4.3. Evaluation of Multi-Polarization Fringe Projection Algorithm with Various Objects	88
4.3.1. Simple Object with Three Different Surfaces.....	89
4.3.2. Microscopic Spatial Filter Stage as HDR Object	93
4.3.3. Circuit Board of Various Intensities	94
4.3.4. Object at Different Exposures.....	95
4.4. Conclusion	97
5. Summary.....	98

5.1. Summarizing the Direct Superresolution Technique.....	98
5.2. Summarizing the Full-Focus Depth-Based Deconvolution	99
5.3. Summarizing the Depth Acquisition for HDR Objects	99
5.4. Closing Thoughts	100
References	101

LIST OF FIGURES

Fig. 1.1. The design parameters of an optical imaging system.	18
Fig. 1.2. Imaging components of an optical system.	20
Fig. 2.1. The resolvability of spatial frequencies through the MTF frequency response of an optical system.	29
Fig. 2.2. Illustration of the shift impact on spatial sampling and angular bandwidth: sensor shift (left) and camera shift (right) as compared with the non-shifted case (center).	31
Fig. 2.3. An iterative subpixel shift estimate scheme using simulated annealing search.	33
Fig. 2.4. RMSE calculated at iterative shifts searched by the simulated annealing algorithm. The final reconstructed image with the estimated shifts using direct SR technique is shown on right.	34
Fig. 2.5. Simple imaging scheme illustrating the linear combinations of HR pixels including the boundaries forming the LR images (top) and the resulting vector-matrix representation that direct SR solves (bottom) in a noisy environment.	36
Fig. 2.6. Structure of the H matrix (right) showing the sparse diagonal-like distribution (locations of nonzero elements) and the color-coded repetitive blocks for a 16×16 object (left) using 4×4 blocks and 4 LR images. The colored blocks below the horizontal axis of the H matrix are highlighting a group of columns that can be copied (with circular shifts) from any other block’s columns belonging to the same colored group.	38
Fig. 2.7. The steps (top) and a block diagram (bottom) summarizing the direct SR technique.	39
Fig. 2.8. Block diagram illustrating the adaptive frequency-based filtering scheme in the training stage (top) and the testing stage (bottom).	42
Fig. 2.9. Training images used in adaptive frequency-based filtering scheme.	42
Fig. 2.10. Visual comparison of various SR techniques at 2× and 4× subsampling factors using sensor shift model, known non-integer-pixel shifts, $\sigma_{\text{blur}} = 0.5$ pixel, and noiseless environment for U.S. Air Force (USAF) target and cameraman image. Shown at the first row: sample of LR images and the HR images. Rows 2-5 displays from left to right: bicubic interpolation, SR of mixed priors, MCBD, and direct SR. The associated RMSE% is shown at their left.	44
Fig. 2.11. Reconstruction error vs. noise of various SR techniques and the AFFS results for the sharpness circle.	46
Fig. 2.12. Visual comparison of the reconstructed sharpness circle (zoomed) at NS = 0.1% and 0.5% for various SR techniques and averaged over 20 noise realizations. A sample LR image and the ground-truth HR image are shown on the left of the first row. The three trained multi-gray masks (wDirect, wBicubic, and wMCBD) are color encoded as red for direct, green for bicubic, and blue for MCBD in the composite RGB masks displayed on the right of the first row.	47
Fig. 2.13. Reconstruction errors vs. σ_{blur} when imaging the boat image through two observation models: solid curves for sensor shift (SS) and dashed curves for camera shift (CS) using various SR techniques.	48

Fig. 2.14. Visual comparison of the reconstructed boat at $\sigma_{\text{blur}} = 0.75$ and 1 pixel for various SR techniques using sensor shift (second and third rows, respectively) and $\sigma_{\text{blur}} = 1$ using camera shift (forth row). A sample LR image is shown on the left of the first row. The three trained averaged masks (wDirect, wBicubic, and wMCBD) are color encoded as red for direct, green for bicubic, and blue for MCBD in the composite RGB masks displayed on the right of the first row..... 50

Fig. 2.15. Reconstruction errors vs. BTol when reconstructing the peppers image with various SR techniques..... 51

Fig. 2.16. Visual comparison of the reconstructed peppers (zoomed) at BTol=10% (middle row) and 20% (bottom row) for various SR techniques. A sample LR image and the original HR image are shown on the left of the first row. The three trained multi-gray masks (wDirect, wBicubic, and wMCBD) are color encoded as red for direct, green for bicubic, and blue for MCBD in the composite RGB masks displayed on the right of the first row. 52

Fig. 2.17. Reconstruction errors vs. STol when restoring the endoscopy image using various SR techniques. 53

Fig. 2.18. Visual comparison of the reconstructed endoscopy image (zoomed) at STol=0.1% (middle row) and 1% (bottom row) for various SR techniques. A sample LR image and the original HR image are shown on the left of the first row. The three trained multi-gray masks (wDirect, wBicubic, and wMCBD) are color encoded as red for direct, green for bicubic, and blue for MCBD in composite RGB masks displayed on the right of the first row. 54

Fig. 2.19. Visual comparison of the reconstructed USAF (zoomed) at two combinations of non-idealities; Set1 (middle row) and Set2 (bottom row) for various SR techniques. A sample LR image and the original HR image are shown on the left of the first row. The three trained multi-gray masks (wDirect, wBicubic, and wMCBD) are color encoded as red for direct, green for bicubic, and blue for MCBD and displayed on the right of the first row. 55

Fig. 3.1. Block diagram of the computational depth-based deconvolution technique. 64

Fig. 3.2. Sketch illustrating the developed depth-based masking approach for planar USAF target set at the three depth planes utilized in the simulation study. 69

Fig. 3.3. Block-wise depth map estimation by finding the best focus over the axially deconvolved blocks. Samples of the axial blocks are encoded in different colors..... 70

Fig. 3.4. USAF (first row) and cameraman (second row) images captured at various object distances. The associated axial PSF profile of 64×64 pixels is shown at the bottom-right corner of each USAF image. 71

Fig. 3.5. Stitching planar regions according to a depth map (left) to form MD objects for USAF (center) and cameraman (right). 71

Fig. 3.6. Blurred images after preprocessing the boundaries with various techniques: zero padding (left), edge taper (center), and block tiling (right) for USAF (first row) and cameraman (second row) objects captured at 100 mm distance. 73

Fig. 3.7. Comparison of deblurred images under various boundary pre-processing techniques using adaptive regularization for planar USAF (first row) and cameraman (second row) objects at 100 mm distance. The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.....	73
Fig. 3.8. Visual comparison of different deconvolution techniques for USAF target at different depth planes when deconvolved with the corresponding axial PSF profiles. The last row shows the final reference maps used as local regularizers in Tsai's technique. The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.	75
Fig. 3.9. Visual comparison of different deconvolution techniques for cameraman image at different depth planes when deconvolved with the corresponding axial PSF profiles. The last row shows the final reference maps used as local regularizers in Tsai's technique. The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.	76
Fig. 3.10. Visual comparison of block-wise deconvolution results for MD USAF and cameraman images using two different block sizes of 100×100 pixels (first two rows) and 200×200 pixels (last two rows) when deconvolved with different axial PSF profiles (sorted column-wise). The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.	78
Fig. 3.11. Visual comparison of deconvolution results for MD USAF and cameraman images without any depth processing (first two rows) and with depth-based masking approach (last two rows) when deconvolved with different axial PSF profiles (sorted column-wise). The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.	79
Fig. 3.12. Visual comparison of final full-focus deconvolution results after stitching based on a depth map for USAF and cameraman images showing reference MD objects (first row), deblurring results when considering SD objects followed by stitching to mimic MD object (second row), deblurring results of MD object without any depth processing (third row), and deblurring results of MD object with depth-based masking approach (last row). The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.....	81
Fig. 4.1. Multi-polarization fringe projection (MPFP) imaging system.	86
Fig. 4.2. Steps of multi-polarization fringe projection technique (MPFP) for HDR objects, vectors denote image-level operations while scalars denote pixel-level operations.....	87
Fig. 4.3. Single-polarization fringe projection imaging of simple object. (a) Simple three-surface object captured by unpolarized camera. (b) Raw polarized data of first distorted fringes. (c) Fringe contrast of various polarization channels at two cross-sections of first distorted fringes (black-white-black tapes on left and metal surface on right). (d) Shape rendering of five fringe images at separate polarizations.....	91
Fig. 4.4. Multi-polarization fringe projection imaging of simple object. (a) Multi-polarization decision map. (b) Merging results of first fringe images. (c) Phase retrieval. (d) Shape rendering of five enhanced fringe images.	92
Fig. 4.5. Multi-polarization fringe projection imaging of microscopic spatial filter stage. (a) Raw polarized data of first distorted fringes. (b) Stage captured by	

regular camera. (c) Decision map. (d) Merging results of first fringe images. (e) Shape rendering of five enhanced fringe images. 94

Fig. 4.6. Multi-polarization fringe projection imaging of circuit board object. (a) Circuit board captured by unpolarized camera. (b) Decision map. (c) Merging results shown for first fringe. (d) Shape rendering of five enhanced fringe images. 95

Fig. 4.7. Multi-polarization fringe projection imaging of an object under different exposures. (a) Scissors captured by unpolarized camera. (b) Decision map. (c) Shape rendering. The images shown in (b) and (c) are sorted left to right according to the utilized exposure time. 96

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
ADC	Analog-to-digital conversion
AFFS	Adaptive frequency-based filtering scheme
BRDF	Bidirectional reflectance distribution function
CS	Camera shift
DFT	Discrete Fourier transform
DOF	Depth of field
FFT	Fast Fourier transform
FOV	Field of view
FPA	Focal plane array
HDR	High dynamic range
HR	High-resolution
IDFT	Inverse discrete Fourier transform
LR	Low-resolution
MAP	Maximum a posteriori
MCBD	Multi-channel blind deconvolution
MD	Multi-depth
MPFP	Multi-polarization fringe projection
MTF	Modulation transfer function
NA	Numerical aperture
NS	Noise strength
POCS	Projections onto convex sets
PSF	Point spread function
RGB	Red, green, and blue
RMSE	Root mean square error
SD	Single-depth
SNR	Signal-to-noise ratio
SR	Superresolution
SS	Sensor shift
SSIM	Structural similarity index measure
USAF	1951 U. S. air force target

Abstract

Miniature cameras play a key role in numerous imaging applications ranging from endoscopy and metrology inspection devices to smartphones and head-mount acquisition systems. However, due to the physical constraints, the imaging conditions, and the low quality of small optics, their imaging capabilities are limited in terms of the delivered resolution, the acquired depth of field, and the captured dynamic range.

Computational imaging jointly addresses the imaging system and the reconstructing algorithms to bypass the traditional limits of optical systems and deliver better restorations for various applications. The scene is encoded into a set of efficient measurements which could then be computationally decoded to output a richer estimate of the scene as compared with the raw images captured by conventional imagers. In this dissertation, three task-based computational imaging techniques are developed to make low-quality miniature cameras capable of delivering realistic high-resolution reconstructions, providing full-focus imaging, and acquiring depth information for high dynamic range objects.

For the superresolution task, a non-regularized direct superresolution algorithm is developed to achieve realistic restorations without being penalized by improper assumptions (e.g., optimizers, priors, and regularizers) made in the inverse problem. An adaptive frequency-based filtering scheme is introduced to upper bound the reconstruction errors while still producing more fine details as compared with previous methods under realistic imaging conditions. For the full-focus imaging task, a computational depth-based deconvolution technique is

proposed to bring a scene captured by an ordinary fixed-focus camera to a full-focus based on a depth-variant point spread function prior. The ringing artifacts are suppressed on three levels: block tiling to eliminate boundary artifacts, adaptive reference maps to reduce ringing initiated by sharp edges, and block-wise deconvolution or depth-based masking to suppress artifacts initiated by neighboring depth-transition surfaces. Finally for the depth acquisition task, a multi-polarization fringe projection imaging technique is introduced to eliminate saturated points and enhance the fringe contrast by selecting the proper polarized channel measurements. The developed technique can be easily extended to include measurements captured under different exposure times to obtain more accurate shape rendering for very high dynamic range objects.

1. Introduction to Computational Imaging

1.1. Basics of Imaging Systems

Imaging is the process of collecting valuable information (e.g., spatial irradiance, angular distribution, spectral response, polarization information) about the scene by focusing the desired information through the optics onto the sensing medium (usually a two-dimensional (2D) focal plane array (FPA)). The conventional optical systems inherently apply a simple form of computation [1] to the captured object's field to create the image. For instance, the imaging optics inherently implements a set of inversion, magnification, filtering, aberrating, and cropping transformations on the imaged field while the imaging sensor array inherently performs 3D to 2D mapping, sampling, quantization, and noise addition operations on the captured data.

The optical system also sets the field of view (FOV) which is the extent of the scene to be captured, the depth of field (DOF) which is the depth range where the scene remains in acceptable focus, the numerical aperture (NA) which is the span of ray angles that can be captured by the optical system, and the magnification of the captured scene. Together the FOV, the DOF, the magnification, and the sensor array define how many image pixels are utilized to sample a certain region in the object space.

Also, the objective lens through its aperture acts as a low-pass filter with optical resolution set by the diffraction limit (assuming no aberrations) given by the Rayleigh criterion as $R = 0.61 \lambda / \text{NA}$ where λ is the wavelength of light. The image sensor, according to the Nyquist criterion, has to sample at less than half

the smallest feature size R (i.e. sampling rate is larger than twice the highest spatial frequency available in the captured data) so no aliasing is induced. Thus, a smaller pixel size in the image sensor is desired but this may also impose limitations on the dynamic range and the signal-to-noise ratio (SNR). In short, the image's spatial resolution is determined by the number of resolvable points or lines per unit length which in turn is limited by the optics (e.g., lens blur, aberration effects, aperture diffraction) and the image sensors (e.g., sampling density, quantization, noise).

1.2. The Design Factors of Imaging Systems

The specifications of an imaging system, shown in Fig. 1.1, may fall in four categories: objects, optics, digitization, and application. The objects to be imaged can be characterized by their geometry, spectrum, bidirectional reflectance distribution function (BRDF), and polarization. The optics determines FOV, DOF, NA, and magnification of the system. The digitization stage governs the analog-to-digital conversion (ADC), the high dynamic range (HDR) capturing, SNR, and sampling resolution. Furthermore, parameters like camera size, acquisition time, monetary cost, and consumed power may be of special importance for certain applications (e.g., endoscopy, smartphones, and depth acquisition).

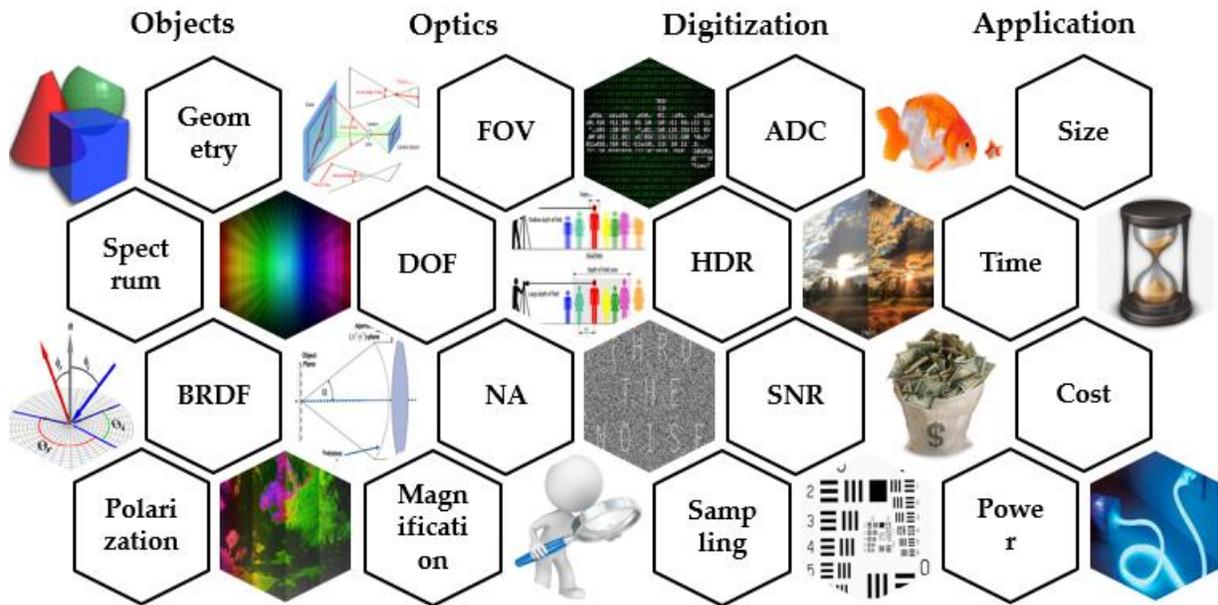


Fig. 1.1. The design parameters of an optical imaging system.

From a design perspective, many of these parameters may conflict with each other in conventional imaging systems. For instance, larger NA mandates smaller FOV and narrower DOF [2, 3] given a fixed focal length and finite detector array. Also, better resolution requires high-quality optics and dense sensors which in turn increases the physical size and design cost. Moreover, smaller FPA detectors (finer resolution) results in a longer acquisition time to collect a sufficient number of photons (acceptable SNR) under the same illumination conditions. Therefore, there is a need for computational imaging techniques to break such traditional limitations.

1.3. Limitations of Miniature Imaging Systems

Miniature imaging devices such as endoscopes are usually designed to cover a relatively wide field of view ($\sim 120^\circ$) but are also physically restricted in the lens aperture, the focal length, and the captured resolution. In addition, the low-illumination conditions and temperature variations in the miniature scenes limit

the acquired dynamic range. These overall result in a small NA, more aberrations, low SNR, and low-resolution (LR) acquisition. Therefore, accurate image analysis is difficult to achieve as the fine features are not being captured.

Furthermore, a fixed-focus camera can only bring the scene to focus beyond half its hyper-focal distance [4] (if focused at infinity) which blurs the close objects and narrows the DOF. As a result, tasks such as detection, recognition, and classification will be difficult to function on the blurred objects as they require a certain degree of clarity (i.e. retrieving certain spatial frequencies) [5].

In addition, traditional fringe projection imagers are unable to render a complete 3D shape of the captured scenes if specular reflections and high dynamic range objects are included. For example, saturated regions of specular reflections (e.g., seen on shiny surgical tools) in a surgical endoscopic scene may completely block any fringe patterns, leading to loss in the retrieved depth information.

Finally in traditional imaging, it is generally accepted that higher performance comes at the cost of complexity. For instance, to increase the resolution of a camera, one needs to increase the number of elements in its lens to combat the aberrations that limit resolution. However, this may not be an acceptable solution for miniature applications. In contrast, computational imaging allows a designer to shift complexity from hardware to computations [6]. For instance, a high-resolution (HR) image can be achieved by post-processing an image captured with simple optics.

1.4. Computational Imaging as a Solution

Computational imaging introduces novel schemes to perform imaging by means of a powerful convergence [7] between the image processing algorithms and the imaging components (e.g., digital sensors, optical elements, and illuminating sources). In other words, the camera and the algorithms are designed to take computation into account. One motivation for developing computational cameras is to create new imaging functionalities that would be difficult, if not impossible, to achieve using the traditional camera model [8]. The new functionality may also give the ability to manipulate the optical settings of an image (e.g., focus, depth of field, viewpoint, resolution, lighting) after the image has been captured.

Computational imagers apply computations throughout the imaging components, presented in Fig. 1.2, by either altering the illumination conditions (e.g., projecting fringe patterns, changing illumination angles), or modulating the propagated rays (e.g., coded aperture imaging, wave-front coding), or translating at least one of the optical elements (e.g., shifting sensor array, moving camera), or inserting new elements (e.g., lenslet array, pixelated polarizer array), or developing computational algorithms (e.g., deriving mathematical tools, implementing advanced image processing techniques), or even by better characterizing the imaging path (e.g., finding the multi-dimensional point spread function) and utilizing this knowledge in the reconstructing algorithms.

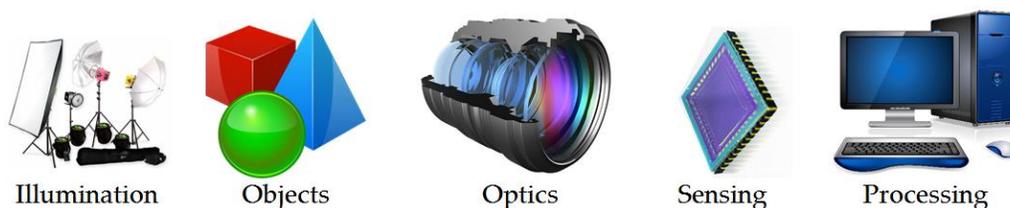


Fig. 1.2. Imaging components of an optical system.

Instead of making a naive one-to-one mapping between the objects and the measurement, the scene could be encoded into a set of measurements which could then be decoded to estimate the scene. The general idea is to modify or better characterize the imaging system so that measurements of the scene can be made more intelligently to estimate the scene more efficiently [1]. This introduces an extra computation step in order to recover the original image from the recorded data.

Through computational imaging, systems can be designed and optimized for specific applications as opposed to the traditional approach of putting high-quality lenses in front of dense detector arrays. A survey on various imaging trends and technologies that may enable the development of computational imaging systems can be found in [6, 9, 10].

In this dissertation, three computational imaging techniques have been developed to deliver a richer representation of the scene for various applications – specifically, by achieving higher resolution, extending the depth of field, and acquiring depth information of a HDR scene. The optical system as well as the reconstruction strategies are discussed in detail. But first, the motivation behind each developed technique is described.

1.4.1. Motivation to Develop a Realistic Superresolution Technique

Superresolution (SR) techniques are employed to retrieve HR details beyond the diffraction limit mandated by the imaging optics and the sensing medium from a set of LR images. Each LR image has to contribute additional information which may be attained by non-integer-pixel shifts with respect to other LR

measurements. By SR techniques, multiple LR measurements are integrated to replace the need for a HR sensor array; hence, the hardware cost is traded for the computational cost [11].

In general, current SR techniques employ optimizers, priors, and regularizers to deliver stable, appealing restorations even though deviating from the real, ground-truth scene. For instance, objective functions may be set to enhance sharpness and reduce ringing artifacts, which makes the reconstructions visually better but may not necessarily resolve fine details (small features) which are critical for accurate image analysis.

A non-regularized SR technique is developed to directly solve a fully-characterized multi-shift imaging reconstruction problem and achieve realistic restorations without being penalized by improper assumptions made in the inverse problem. An adaptive frequency-based filtering scheme is further introduced to enable the technique to function and still produce more fine details as compared with other SR techniques when inaccurate shift estimation, boundary variations, noise, and blurring scenarios are considered.

1.4.2. Motivation to Develop a Full-Focus Imaging Technique

Hardware-based approaches extend the depth of field by inserting or modifying or even translating optical elements through the imaging process to produce a depth-invariant point spread function (PSF). However, these solutions may result in reduction of SNR, longer acquisition time, larger physical size, and additional monetary cost. Software-based approaches can output an all-in-focus scene from a single image captured by ordinary compact cameras. However, most such

techniques imprecisely approximate the PSF as a 2D depth-invariant profile, which leads to suboptimal restorations with ringing artifacts especially when objects in the captured scene belong to a broad range of depth planes.

A computational depth-based deconvolution technique is developed to bring a scene captured by an ordinary fixed-focus camera into full focus based on a depth-variant PSF prior. The captured image is brought into focus at different depth planes by deconvolving it with axial slices of the inverse PSF volume. The focused details from the deconvolved images are then stitched together to deliver the full-focus image. The ringing artifacts are suppressed on three levels: block tiling to eliminate boundary artifacts, adaptive reference maps to reduce ringing initiated by sharp edges, and block-wise deconvolution or depth-based masking to suppress artifacts initiated by multi-depth objects.

1.4.3. Motivation to Develop a Depth Acquisition Technique for HDR Objects

Conventional fringe projection imagers comprise a projector-camera pair in which successive phase-shifted fringe patterns are projected onto objects, become distorted, and are then captured by the camera. These captured distorted fringes carry valuable information about the object's depth, which can be retrieved through phase shifting algorithms. However, these conventional imagers struggle to estimate the shape of high dynamic range objects where detected fringes are of limited visibility. Moreover, saturated regions of specular reflections can completely block any fringe patterns, leading to a loss in the depth information.

In this work, the conventional camera is replaced by a pixelated polarizer camera capable of acquiring richer information about the scene. Besides, a multi-

polarization fringe projection imaging technique is developed on the computational side to eliminate saturated points and enhance the fringe contrast by selecting the proper polarized channel measurements. The developed technique can be easily extended to include measurements captured under different exposure times to obtain more accurate shape rendering for very high dynamic range objects.

1.5. Organization of the Dissertation

The theme for the work described in this dissertation is to develop computational techniques that empower low-quality miniature cameras to deliver multi-task imaging capabilities – in particular, developing realistic superresolution reconstructions, full-focus imaging, and depth acquisition for HDR objects. In order to develop these novel computational techniques, a deep understanding of both the imaging system as well as the reconstruction techniques is the key focus in this work. The following chapters are organized as follows.

Chapter 2 starts by illustrating the resolvability of spatial frequencies captured by the optical system and explaining why multi-shift imaging systems are good candidates for the superresolution techniques. Then, different observation models are characterized and efficient parallel computation of the PSF matrix is prescribed. Next, the adaptive frequency-based filtering scheme is introduced and filtering masks are trained against different imaging conditions. Afterward, a thorough simulation study is carried out under ideal and realistic imaging scenarios and comparison with various SR techniques is presented. Finally, the conclusion and future research directions are listed.

Chapter 3 describes the computational depth-based deconvolution technique. First, the image boundaries are processed by a block tiling approach to eliminate artifacts. Then, the iterative deconvolution objective function is analyzed and an adaptive regularization scheme is introduced. Afterward, a block-wise deconvolution and depth-based masking are proposed. Simulation results are shown for planar objects set at different depths and for multi-depth objects. This is followed by the conclusion and future research directions.

Chapter 4 illustrates the multi-polarization fringe projection technique and presents the improvement in depth estimation when merging distorted fringe images captured through different polarization channels. The algorithm is validated experimentally for different HDR objects and demonstrated for different exposure scenarios.

Chapter 5 summarizes the developed computational imaging techniques and highlights the computational capabilities enabled in the miniature cameras.

2. Direct Superresolution for Realistic Image Reconstruction

2.1. Introduction

Many commercialized imaging systems do not capture the high-resolution (HR) features of scenes due to the additional monetary costs of the high-quality optics and the dense sensors, the longer acquisition time needed to collect sufficient photons at smaller imaging units, and the physical restrictions imposed by miniature imaging applications such as endoscopic systems and industrial fiberscopes which limits the number of detectors on the focal plane array (FPA). Superresolution (SR) techniques [11-13] are a powerful way to bypass the diffraction limits mandated by the imaging optics and the sensing medium so the HR details within the scene can be retrieved. Many SR techniques integrate multiple low-resolution (LR) images, each contributing additional information, through an inverse problem to output a relatively richer scene of greater resolving power with less aliasing.

The SR inverse problem, in general, is an ill-posed problem as the solution is not guaranteed to be existent, stable, or unique [14]. Therefore, SR techniques approximate the HR image by optimizing weights for non-uniform interpolation [15], or implementing Tikhonov's regularization component to overcome noise [16], or employing stochastic reconstruction methods such as maximum likelihood (ML) [17] or maximum a posteriori (MAP) estimation [18], or combining deterministic and stochastic regularizers such as sparse and non-sparse priors [19, 20] to constrain the restorations (i.e., the reconstructed images), or iterating the projections onto convex sets (POCS) of priors to assure convergence

while satisfying the assumed priors [21, 22], or adopting back-projection schemes [23] to iteratively minimize the differences between the measured LR images and the synthesized ones, or even alternating the optimization using deconvolution techniques [24] to reverse the acquisition process. These optimization schemes may ensure a converging solution, but at the cost of deviating from the real ground-truth restoration since they are inherently biased by the type of priors, the utilized regularizers, and the optimized objective functions. For instance, regularizing to enhance smoothness and reduce aliasing can make the reconstructions visually better but may not necessarily resolve the actual HR details critical for accurate image analyses (e.g., medical diagnostics).

A fast SR technique [25] is proposed to deliver an unconstrained accurate physical HR solution without being biased by regularizing terms or optimization schemes. A forward observation model is first characterized using information such as an estimate of the blurring kernels (resulting from the imaging optics and the sensor's finite size) and the relative shifts between the acquired LR images to find a full-rank matrix representation of the multi-shift imaging point spread function (PSF), and by characterizing the noise statistics. The HR reconstruction is then found by directly solving a set of sufficient linear equations to reconstruct a rich HR image. This technique produces a unique exact restoration under ideal shift estimates and in-focus noiseless measurements, making the inverse a well-posed problem. The realistic scenarios of inaccurate shift estimates and blurred noisy measurements in some cases may result in more unknowns than the number of equations or small changes in some variables which may hurt the stability of

the solution. For such scenarios, an adaptive frequency-based filtering scheme is introduced to upper bound the reconstruction errors while still producing more fine details as compared with other regularized SR techniques.

2.2. The Resolvability of Spatial Frequencies in the Imaging Systems

The spatial frequency content captured by an optical systems is usually characterized by the modulation transfer function (MTF), shown in Fig. 2.1. Beyond the MTF cutoff f_c , spatial frequencies are permanently lost as they are extinguished by the optics (assuming no active illumination such structured projections or variably-illuminated scenes implemented for Fourier Ptychography [26-29]). The image sensor array through its sampling rate $f_s = \frac{1}{\text{detector pitch}}$ determines the maximum frequency $f_{max} = f_s/2$ to be captured without aliasing (according to Nyquist criterion). The rest frequencies between f_{max} and f_c that pass through the optics will be aliased (folded in the low frequency range) but they can be unrolled by the SR techniques. SR removes the ambiguity in the aliased frequency components by integrating multiple aliased LR images of the scene, each taken with some change in the observation parameters (e.g., camera motion, sensor shift, aperture synthesizing). A mathematical proof of how HR details can be resolved from a set of shifted images (in the camera motion case) is discussed next.

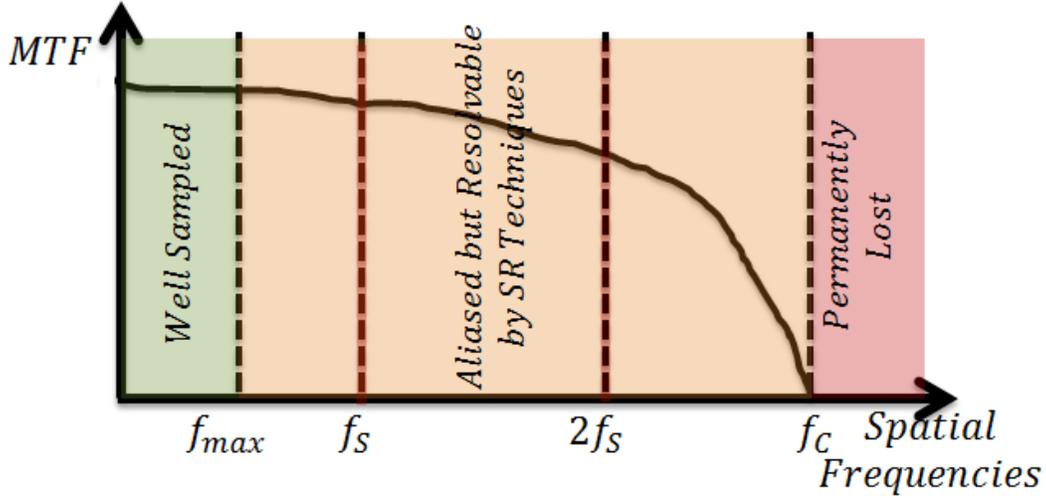


Fig. 2.1. The resolvability of spatial frequencies through the MTF frequency response of an optical system.

2.3. Mathematics of Multi-Shift Images for Superresolution

Consider a $2\times$ HR image that has a discrete Fourier transform (DFT) (1D notation

is used here for simplicity) given by;
$$I_{HR}(f) = S(f)H(f)\text{sinc}\left(\frac{f}{2f_s}\right) \quad (1)$$

where the *sinc* term, H and S are the DFT of sensor blur, optical blur and the scene, respectively.

By the aliasing property [30], the corresponding spectrum of the LR image is given by;

$$I_{LR}(f) = S(f)H(f)\text{sinc}\left(\frac{f}{f_s}\right) + S(f_s - f)H(f_s - f)\text{sinc}\left(1 - \frac{f}{f_s}\right) \quad (2)$$

Capturing a second LR image translated by δ in the spatial domain results in phase terms in the Fourier domain as in the following;

$$I'_{LR}(f) = \exp(j2\pi f\delta) S(f)H(f)\text{sinc}\left(\frac{f}{f_s}\right) + \exp(j2\pi(f_s - f)\delta) S(f_s - f)H(f_s - f)\text{sinc}\left(1 - \frac{f}{f_s}\right) \quad (3)$$

Let us define a term $G(f) = \text{sinc}(\frac{f}{f_s})/\text{sinc}(\frac{f}{2f_s})$ to reflect the relative signal loss between the LR and HR images due to the pixelization term. Then equations (1), (2), and (3) can be rewritten in a matrix format [31] as follows;

$$\begin{bmatrix} I_{LR}(f) \\ I'_{LR}(f) \end{bmatrix} = \begin{bmatrix} G(f) & G(f_s - f) \\ \exp(j2\pi f\delta) G(f) & \exp(j2\pi(f_s - f)\delta) G(f_s - f) \end{bmatrix} \begin{bmatrix} I_{HR}(f) \\ I_{HR}(f_s - f) \end{bmatrix} \quad (4)$$

This leads to a linear system relating the DFT of the pair of LR images to that of the HR image. By solving this system, the aliased frequencies will be unrolled leading to a richer HR scene. Note that the displacement must be a non-integer-pixel value $\delta = \alpha/f_s$ where α is non-integer, otherwise the phase terms will be eliminated. i.e. $\exp(j2\pi f\delta) = \exp(j2\pi\alpha f/f_s) = 1$ since f/f_s is integer too. This will make the second measurement in (4) redundant and the SR algorithms to fail. Note also that due to the *sinc* ripples, the frequencies at multiple integer of f_s are not retrievable, see Fig. 2.1.

Finally the frequencies beyond f_c , which are extinguished by the optics $H(f)$, are independent of the application of the shift operator, hence, are permanently lost. From an information perspective, optical blur represents a true loss of information whereas aliasing is only a scrambling of information. The SR algorithms can only unscramble the existing information [31].

2.4. Multi-Shift Imaging Systems

In multi-shift imaging systems, the LR images are linearly related to the unknown HR object through blurring, sampling, shifting, and addition of noise. To assure diversity in the collected LR measurements, consecutive non-integer (subpixel or large) shifts at the acquisition stage can be introduced by shifting the sensor array

itself or the camera as a whole. Shifting the sensor may be accomplished by small piezoelectric actuators while maintaining the optics and camera body fixed. Having actuators shifting only the sensor may be advantageous, for example, in endoscopic imaging since it enables automating accurate shifts within the limited physical space while maintaining the stability of the optics and reducing the translated load. Camera shift is a second possible case. For general imaging purposes, shifting the camera may be simpler as successive measurements are captured in motion (keeping in mind that multi-head systems are a special case of camera shift model). A distinction between sensor shift and camera shift, presented in Fig. 2.2, is the captured content itself; shifting the sensor resembles spatial resampling within the same angular bandwidth captured by the optics whereas shifting the camera represents spatial sampling of shifted angular bandwidth as new rays propagate through the system.

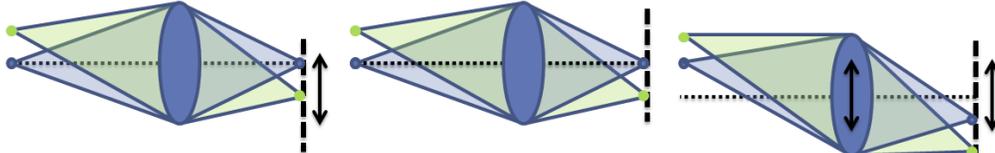


Fig. 2.2. Illustration of the shift impact on spatial sampling and angular bandwidth: sensor shift (left) and camera shift (right) as compared with the non-shifted case (center).

Consider a vector notation such that the 2D image is rasterized block-wise lexicographically into a 1D vector. For the sensor shift (SS) case, the imaged HR object \vec{f} is first optically blurred by kernel B , then undergoes K consecutive non-integer-pixel shifts S_k where $k = 1, \dots, K$, then is subsampled by the FPA operator D which includes the sensor blur as well, and then perhaps is combined with an additive thermal noise \vec{n}_k , resulting in a captured image. The mathematical

acquisition model for the SS case can be described by $\vec{g}_k^{SS} = D(S_k(B * \vec{f})) + \vec{n}_k$, where \vec{g}_k^{SS} is the k^{th} captured image using sensor shift, \vec{n}_k is the k^{th} additive noise, and $*$ is the convolution operator. For the camera shift (CS) case, the shift and blurring order is alternated so the observation model becomes $\vec{g}_k^{CS} = D(B * S_k(\vec{f})) + \vec{n}_k$ where \vec{g}_k^{CS} is the k^{th} captured image using camera shift. The global shift vectors and the blurring and subsampling kernels are employed to generate the PSF matrix \mathbf{H} , so the multi-shift imaging system can be simplified to a linear system incorporating all K measurements as $\vec{g} = \mathbf{H}\vec{f} + \vec{n}$.

2.5. Parameters Estimate from the LR Images

2.5.1. Shift Estimate

The relative shifts between captured images can be estimated using spatial-based or frequency-based registration techniques [32-34] with subpixel accuracy of about 1/10-1/100 pixel. Yet this accuracy is not adequate for solving the linear system as a small variation in any of the parameters can result in a large deviation from the correct solution; therefore, a more accurate approach, which may involve setting an iterative simulated annealing search [35], may be used to find the best shift combinations in XY directions for all the LR images.

An iterative subpixel shift estimate scheme, sketched in Fig. 2.3, starts by estimating the coarse relative shifts between the captured LR Images and feeding them as initial values to the simulated annealing search. The search through the iterations aims to find the best shifts that minimize the root mean square errors (RMSE) between the reconstructed images delivered by the developed direct SR

technique and at least one reference SR technique. Since the LR images may be originally captured with large non-integer shifts (larger than pixel pitch), they are first registered with pixel-accuracy and the overlapped LR regions are extracted and input into various SR techniques to calculate the reconstruction errors. Note that the non-overlapped measurements are ignored as they are insufficient to make the PSF matrix \mathbf{H} a full-rank matrix.

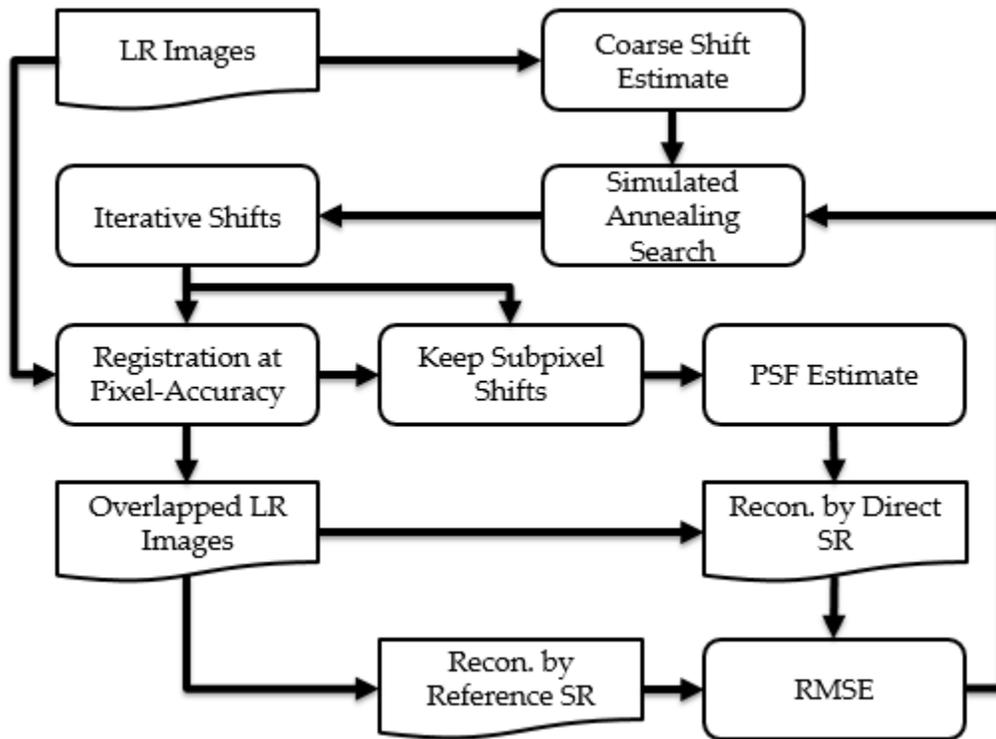


Fig. 2.3. An iterative subpixel shift estimate scheme using simulated annealing search.

To demonstrate the subpixel estimation accuracy that can be delivered in the proposed scheme, 4 LR images of USAF target were captured at $2\times$ SR factor with (X,Y) shifting vectors set as $(0,0; \text{Shift}_x,0; 0,\text{Shift}_y; \text{Shift}_x,\text{Shift}_y)$ where $\text{Shift}_x = 14.4349$ and $\text{Shift}_y = 7.92$ in LR pixel unit. The simulated annealing algorithm was configured to find two unknown variables ($\text{Est_Shift}_x, \text{Est_Shift}_y$) over 2000 iterations at max with searching window ± 0.25 around the coarse

estimated shifts (14.5, 8). Three reference SR techniques were used to increase the reliability by minimizing the RMSE averaged over all; Bicubic SR [15], MCBD SR [24], and dictionary-based SR [36]. The iterative RMSE vs. the iterative shifts are shown in Fig. 2.4. The best combination of shift vectors estimated were $\text{Est_Shift}_x = 14.4370643$ and $\text{Est_Shift}_y = 7.91964983$. This shows subpixel accuracy of $\frac{1}{462} \times \frac{1}{2855.7}$ pixel which is one-to-two orders better than what is reported so far in the literature on other registration techniques [32-34]. Yet this simulated annealing search is too slow (~6 hours) as compared with others (~1 min).

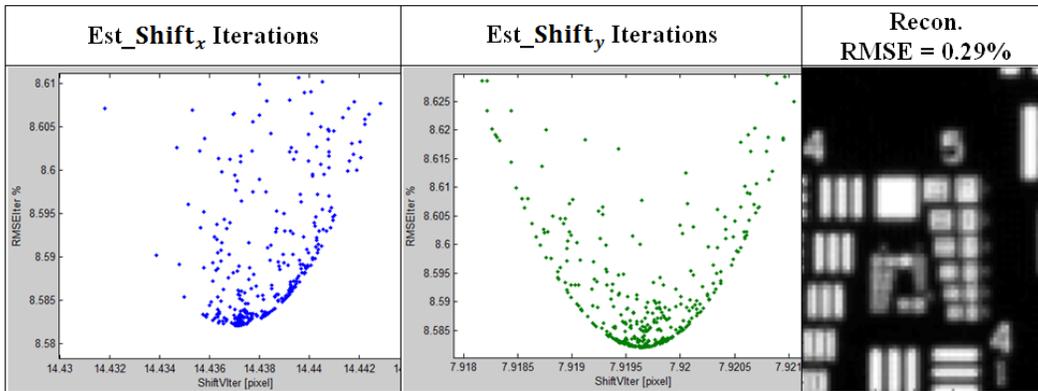


Fig. 2.4. RMSE calculated at iterative shifts searched by the simulated annealing algorithm. The final reconstructed image with the estimated shifts using direct SR technique is shown on right.

2.5.2. Blurring Estimate

The blurring kernel is assumed to be a spatially invariant operator (ignoring aberrations) and representing the optical blur due to the finite aperture of the imaging optics. The blurring profile (shape and strength) may be estimated through blind deconvolution techniques [24, 37].

2.5.3. Noise Estimate

The additive noise is assumed to follow the normal distribution (additive white Gaussian noise) where the noise statistics (mean and variance) can be estimated using the pixel-wise adaptive Wiener method [38].

2.6. Direct Superresolution Algorithm

Assuming N pixels at the HR level and the dimensional subsampling factor d , the minimum number of LR images (each having N/d^2 pixels) required to make the linear imaging system a well-posed SR problem is $K = d^2$. Fig. 2.5 illustrates the direct SR algorithm by a simple case of a 4×4 object f_j (j indexing the HR pixels) divided into 2×2 blocks, subsampling factor $d = 2$, four LR captured images g_i^k with noise n_i^k (k indexing the LR image and i indexing its pixels), and equal non-integer-pixel shifts in a noisy environment to obtain a full-rank system \mathbf{H} . Each column in \mathbf{H} is found by letting the associated object's pixel be one while setting the rest of the pixels to zero and recording the response through the observation model on the FPA. The weighted linear combinations of HR pixels, as shown in Fig. 2.5, form the measured LR images; solving this set of unique linear equations can grant the exact restoration of the HR object under ideal imaging conditions.

$$\begin{aligned}
f'_1 & \rightarrow g_1^1 = h_{1,1}f_1 + h_{1,2}f_2 + h_{1,3}f_3 + h_{1,4}f_4 + n_1^1 \\
f'_1 & \rightarrow g_1^2 = h_{2,2}f_2 + h_{2,4}f_4 + h_{2,5}f_5 + h_{2,7}f_7 + n_1^2 \\
f'_3 & \rightarrow g_1^3 = h_{3,3}f_3 + h_{3,4}f_4 + h_{3,9}f_9 + h_{3,10}f_{10} + n_1^3 \\
f'_3 & \rightarrow g_1^4 = h_{4,4}f_4 + h_{4,7}f_7 + h_{4,10}f_{10} + h_{4,13}f_{13} + n_1^4 \\
& \vdots \\
f'_9 & \rightarrow g_4^1 = h_{13,13}f_{13} + h_{13,14}f_{14} + h_{13,15}f_{15} + h_{13,16}f_{16} + n_4^1 \\
f'_{11} & \rightarrow g_4^2 = h_{14,14}f_{14} + h_{14,14}^{ur}f'_{14} + h_{14,16}f_{16} + h_{14,16}^{br}f'_{16} + n_4^2 = h_{14,14}^*f_{14} + h_{14,16}^*f_{16} + n_4^2 \\
f'_{11} & \rightarrow g_4^3 = h_{15,15}f_{15} + h_{15,16}f_{16} + h_{15,15}^{bl}f'_{15} + h_{15,16}^{br}f'_{16} + n_4^3 = h_{15,15}^*f_{15} + h_{15,16}^*f_{16} + n_4^3 \\
f'_{11} & \rightarrow g_4^4 = h_{16,16}f_{16} + h_{16,16}^{ur}f'_{16} + h_{16,16}^{bl}f'_{16} + h_{16,16}^{br}f'_{16} + n_4^4 = h_{16,16}^*f_{16} + n_4^4
\end{aligned}$$

$$\begin{bmatrix} g_1^1 \\ g_1^2 \\ g_1^3 \\ g_1^4 \\ \vdots \\ g_4^1 \\ g_4^2 \\ g_4^3 \\ g_4^4 \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & h_{1,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_{2,2} & 0 & h_{2,4} & h_{2,5} & 0 & h_{2,7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h_{3,3} & h_{3,4} & 0 & 0 & 0 & 0 & h_{3,9} & h_{3,10} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_{4,4} & 0 & 0 & h_{4,7} & 0 & 0 & h_{4,10} & 0 & 0 & h_{4,13} & 0 & 0 & 0 \\ \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{13,13} & h_{13,14} & h_{13,15} & h_{13,16} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{14,14}^* & 0 & h_{14,16}^* \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{15,16}^* & h_{15,16}^* \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{16,16}^* \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ \vdots \\ f_{13} \\ f_{14} \\ f_{15} \\ f_{16} \end{bmatrix} + \begin{bmatrix} n_1^1 \\ n_1^2 \\ n_1^3 \\ n_1^4 \\ \vdots \\ n_4^1 \\ n_4^2 \\ n_4^3 \\ n_4^4 \end{bmatrix}$$

Fig. 2.5. Simple imaging scheme illustrating the linear combinations of HR pixels including the boundaries forming the LR images (top) and the resulting vector-matrix representation that direct SR solves (bottom) in a noisy environment.

However, there is an inherent dependency between the adjacent LR pixels (including the boundary conditions involved in the subsampling process, which are shown as f'_j in Fig. 2.5) which forces us to solve the whole system globally (and not individually at the block level). This may result in a giant \mathbf{H} matrix for large images; nevertheless, imposing block-wise lexicographical ordering in both the object and the measurement domains results in a diagonal-like sparse matrix, shown in Fig. 2.6, that can be directly and efficiently solved using Gaussian elimination or linear least-squares techniques [39, 40] (without the need to find its inverse). For the boundary pixels, the assumption made is that the HR objects are smoothly varying at the boundaries, which allows us to approximate their values by the closest HR pixels $f'_j \approx f_j$ so no additional unknowns are introduced and the related boundary weights $h_{(i-1)K+k,j}^{ul,ur,bl,br}$ can be compacted into $h_{(i-1)K+k,j}^*$ (with

superscript *) in the \mathbf{H} matrix; see boundary weights of block 4 in Fig. 2.5 for extra clarification.

An additional advantage of adopting the block approach is the repetitive pattern found in the \mathbf{H} matrix for similar blocks (inner blocks, boundary blocks, corner blocks), where repetitive blocks are color coded in Fig. 2.6. For instance, when generating \mathbf{H} , the response of the object's points is first computed for a single inner block (the columns of \mathbf{H} associated with that block) and then the computed response is copied to all other columns of \mathbf{H} associated with the other inner blocks while maintaining the proper shifts representing the spatial location of these blocks within the HR image. This means that only the impulse response of a single block within each of the nine color-encoded groups, displayed in Fig. 2.6-left, needs to be computed and then copied (with the relative circular shifts) to other blocks belonging to the same group. The size of the blocks is also set adaptively according to the subsampling rate and the size of the HR object so the computational loads made to find and copy the blocks' responses are balanced. This makes generating the sparse \mathbf{H} matrix computationally efficient (highly parallelized), highly scalable (for large images and subsampling factors), and with an efficient memory allocation. For instance, superresolving to get an HR image of size 512×512 using $2 \times$ and $4 \times$ subsampling factors takes 3.25 min ($T_{PSF} = 1.54$ min, $T_{Recon} = 1.43$ min) and 20.8 min ($T_{PSF} = 4.43$ min, $T_{Recon} = 15.83$ min) respectively on a computer with 12 cores, each having a 2 GHz clock rate and 16 GB of RAM.

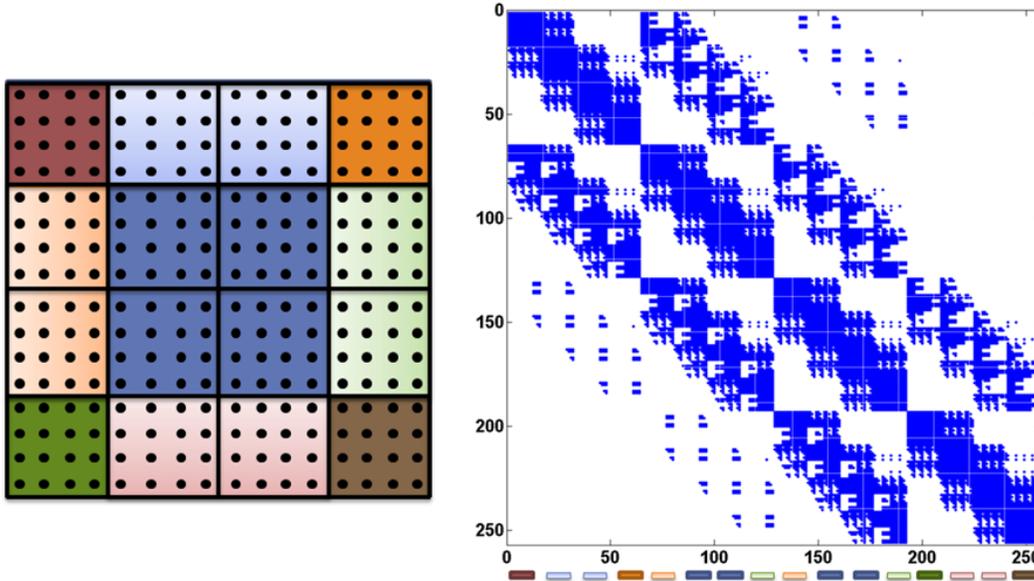


Fig. 2.6. Structure of the H matrix (right) showing the sparse diagonal-like distribution (locations of nonzero elements) and the color-coded repetitive blocks for a 16×16 object (left) using 4×4 blocks and 4 LR images. The colored blocks below the horizontal axis of the H matrix are highlighting a group of columns that can be copied (with circular shifts) from any other block's columns belonging to the same colored group.

Finally, it should be noted that the mechanism of simulating the H matrix may require an accurate knowledge of the shifting vectors (it can work with random non-integer large shifts, too), fast acquisition of the LR images (so that global shifts are sufficient as the captured scene can be assumed static), and smooth variations at the boundaries (since replications are considered when finding the responses at boundaries). Also, to deliver a stable solution, the direct SR technique favors acquisition of the LR images with good focus and low noise. The steps and a block diagram summarizing all the operations in the direct SR technique is presented in Fig. 2.7.

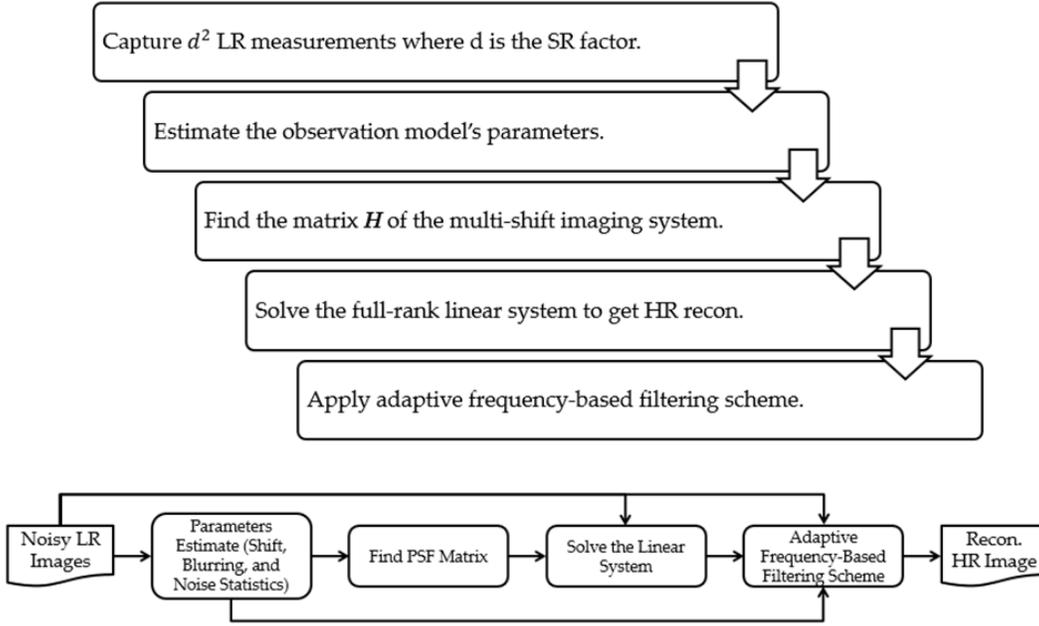


Fig. 2.7. The steps (top) and a block diagram (bottom) summarizing the direct SR technique.

2.7. Adaptive Frequency-Based Filtering Scheme

A closer look at the impact of non-idealities on the linear imaging equations, given in Fig. 2.5, reveals challenges that may prevent any robust solution: 1) The misestimate of shifting vectors or blurring kernel will result in global deviations in the \mathbf{H} weights. 2) Out-of-focus or noisy measurements may prevent the uniqueness or existence of the reconstruction. 3) The nonconformity of the replication assumption at the boundaries may affect the starred weights (h^*), altering some equations, hence, affecting the whole system.

To enable the direct SR technique to function in such realistic scenarios while maintaining its merit by not incorporating priors and optimizers directly, an adaptive frequency-based filtering scheme (AFFS) has been developed to adaptively combine the direct SR reconstruction with reference SR reconstructions in a frequency-by-frequency manner based on trained masks to

output an improved HR reconstruction having bounded errors. At least one reference SR technique (indexed $r = 1, \dots, R$) is chosen to upper bound the reconstruction errors due to non-idealities.

At the training stage, shown at the top of Fig. 2.8, a set of HR images is used to generate noisy blurred LR images through the observation model at predefined shifting vectors, blurring kernel, and subsampling factor. This can be repeated to generate multiple realizations of LR images. The LR images are then inputted into multiple SR techniques to generate corresponding HR reconstructed images, and the discrete Fourier transform (DFT) of the reconstructed images is calculated and ordered so the DC frequency component is at the center of the DFT array. Next, masking information is generated to associate each frequency with a corresponding preferred direct SR technique or a reference SR technique. An outward spiral scanning path can be chosen to step through the frequency locations in the mask, starting with zero frequency as the current frequency component. The calculation of the mask starts by taking only the currently selected frequency components of the reference reconstructed image DFT and the direct reconstructed image DFT while setting the other frequency components to zero, computing the inverse discrete Fourier transform (IDFT), and then calculating the root mean square error (RMSE) with respect to the original HR image in the spatial domain. The current frequency location in the mask is then associated with the SR technique whose frequency component results in the lowest RMSE, and a new search to find the best next frequency component (along with the already selected frequency components) is conducted till all frequency

locations have been associated with a preferred SR technique. The outward spiral pattern can be beneficial as lower frequencies usually carry more energy as compared with higher frequencies for natural scenes. The masks resulting from each realization in the training stage represent binary decision maps for the given HR image, subsampling factor, translating information, blurring kernel, noise statistics, and noise realization. By averaging masks over many different realizations, the training stage outputs multiple filtering masks (w_{Ref_r} for each reference SR technique and w_{Direct} for the direct SR technique) scaled to have real values (optimized weights) from zero to one.

Similarly at the testing stage, shown at the bottom of Fig. 2.8, the captured LR images are used to calculate reconstructed images using both the reference SR techniques and the direct SR technique. The DFT of each reconstructed image is computed, yielding $\mathcal{F}_{\text{Ref}_r}$ and $\mathcal{F}_{\text{Direct}}$. In the testing stage, the filtering masks w_{Ref_r} and w_{Direct} (averaged over all cases) are used to weight the frequency components, i.e. perform filtering, to obtain the DFT of the final reconstructed image: $\mathcal{F}_{\text{AFFS}} = \sum_r w_{\text{Ref}_r} \mathcal{F}_{\text{Ref}_r} + w_{\text{Direct}} \mathcal{F}_{\text{Direct}}$. The IDFT of $\mathcal{F}_{\text{AFFS}}$ is then computed to deliver the final reconstructed image.

Example training images to be used in this chapter are presented in Fig. 2.9. For validation purposes, any test image may be selected from this set, and the training procedure will then use masks averaged over the other images (not including the test image) with noise realizations for a given noise strength (NS), blurring kernel of standard deviation σ , and subsampling factor d .

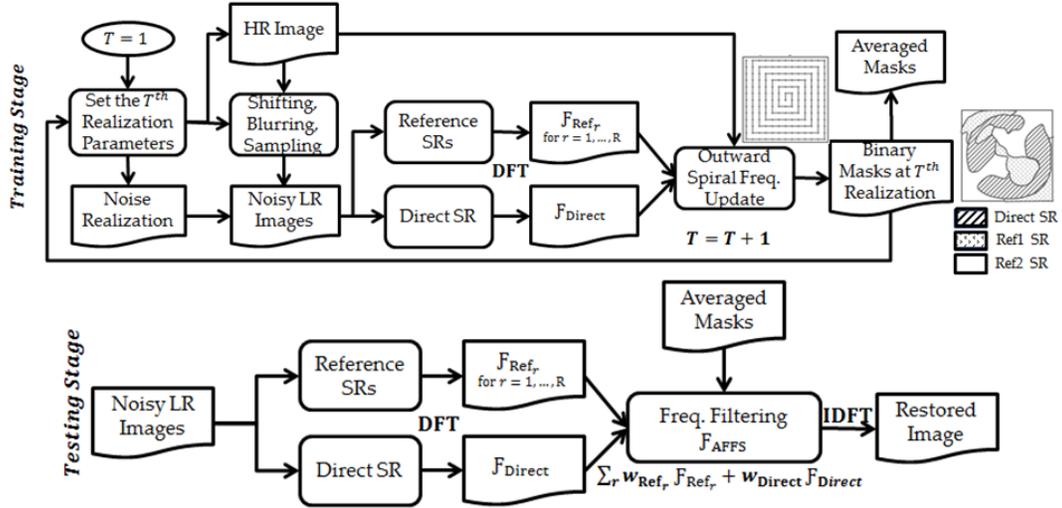


Fig. 2.8. Block diagram illustrating the adaptive frequency-based filtering scheme in the training stage (top) and the testing stage (bottom).

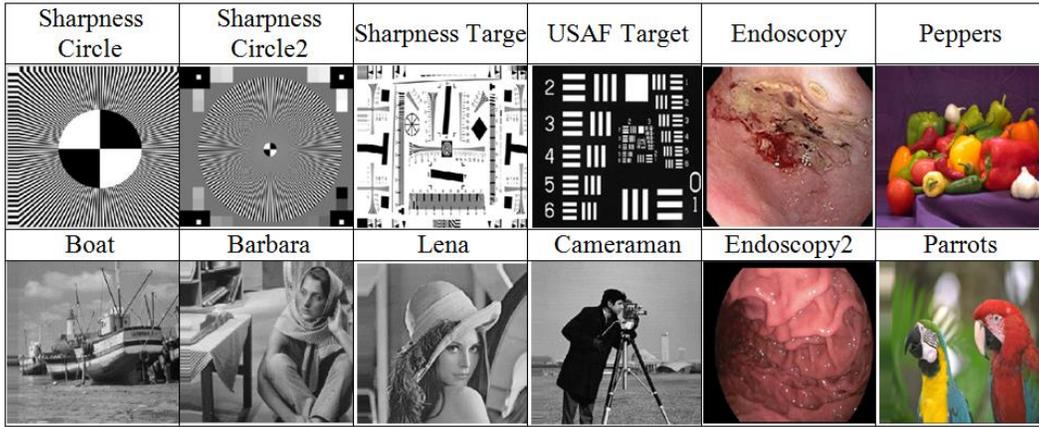


Fig. 2.9. Training images used in adaptive frequency-based filtering scheme.

2.8. Simulation Results

A simulation study has been conducted to evaluate the performance of the direct SR technique described herein and the AFFS visually and numerically in idealistic and realistic scenarios. Unless noted differently in the evaluated scenarios, the observation model to be used assumes a sensor shift with $2\times$ subsampling factor, equal known non-integer-pixel shifts along X-Y directions, in-focus capturing (B has Gaussian distribution with standard deviation $\sigma_{\text{blur}} = 0.5$ pixel), replicated boundary conditions, and noiseless environment. As case

studies, the imaging parameters in $\vec{g}_k = D(S_k(B * \vec{f})) + \vec{n}_k$ are varied one at a time and their impact on the restored images are studied; I either vary the boundary conditions in \vec{f} , or modify the blurring variance in B , or introduce shift misestimates in S_k , or alter the sequence order between B and S_k , or change the subsampling in D , or even apply different noise strengths in \vec{n}_k . At the end, the performance is evaluated at a combination of these variations.

2.8.1. Evaluating Various SR Techniques under Ideal Imaging Conditions

Four different SR techniques are implemented to visualize the impact of biasing priors and optimization schemes on the reconstructions of $2\times$ and $4\times$ subsampling factors, shown in Fig. 2.10. The multi-image SR techniques utilized here are the non-uniform bicubic interpolation [15], SR with mixed priors [19] (sparse ℓ_1 -norm of horizontal and vertical first-order differences and non-sparse simultaneous autoregressive prior with $\lambda = 0.5$), the multi-channel blind deconvolution (MCBD) [24], and the direct SR technique. To quantify the reconstruction errors (at pixel level) between different SR techniques for various objects, the relative numerical error is calculated by $\text{RMSE} = \text{sqrt}\left(\frac{1}{N}\sum_{j=1}^N (f_j - \hat{f}_j)^2\right) / L$, where f_j and \hat{f}_j are the j^{th} pixels of the HR image and the recon. image, respectively, N is the number of pixels, and L is the maximum allowed intensity (e.g., 255 for an 8-bit image); see Fig. 2.10. More simulation results are available in [41].

		Sample LR Image		HR Image (at 4X)	
		Bicubic SR	Mixed Priors SR	MCBD SR	Direct SR
2X		14.6907%	23.0428%	18.3693%	0.5589%
		10.3584%	17.7861%	11.9322%	0.0186%
4X		23.0047%	26.59%	29.6854%	0.6294%
		13.3805%	19.1719%	17.1718%	0.01497%

Fig. 2.10. Visual comparison of various SR techniques at 2 \times and 4 \times subsampling factors using sensor shift model, known non-integer-pixel shifts, $\sigma_{\text{blur}}=0.5$ pixel, and noiseless environment for U.S. Air Force (USAF) target and cameraman image. Shown at the first row: sample of LR images and the HR images. Rows 2-5 displays from left to right: bicubic interpolation, SR of mixed priors, MCBD, and direct SR. The associated RMSE% is shown at their left.

The visual quality of the restorations (especially at the resolved HR features of the USAF target, the camera body and its aligning stick) and the numerical errors of the implemented SR techniques in idealistic scenarios confirm that the direct SR technique described herein resolves the true features with fewer artifacts than the other SR techniques.

2.8.2. Studying the Adaptive Frequency-Based Filtering Scheme at Noisy Environments

Two reference SR techniques, bicubic and MCBBD, are implemented besides the direct SR technique in the training stage to increase the degrees of freedom in selecting best frequencies and weighting masks that minimize the reconstruction errors. The resulting masks are averaged over 11 training images (since the test image is not included) and 20 noise realizations of normal distribution $\mathcal{N}(0, \sigma_{\text{noise}}^2)$ at the given noise strength NS related to $\sigma_{\text{noise}} = \text{NS} \times L$, where L is the maximum allowed intensity. The sharpness circle is used as a test image, and Fig. 2.11 shows the calculated reconstruction RMSE vs. NS, averaged over 20 noise realizations using the direct SR technique, bicubic, MCBBD, and the ADFS.

The RMSE curves, shown in Fig. 2.11, demonstrate that while the bicubic and MCBBD SR techniques experience insignificant changes, the reconstructed images using direct SR suffer as NS increases. Yet, the ADFS succeeded in picking the best frequencies to suppress any artifacts while maintaining the fine structures leading to the lowest RMSE values; see also Fig. 2.12. Note that in the noiseless scenario at $\text{NS} = 0$, due to the spiral path utilized in the ADFS, the filtering masks are not picking all frequencies from the direct SR technique, leading to slightly larger errors (black curve) than the direct SR results (red curve).

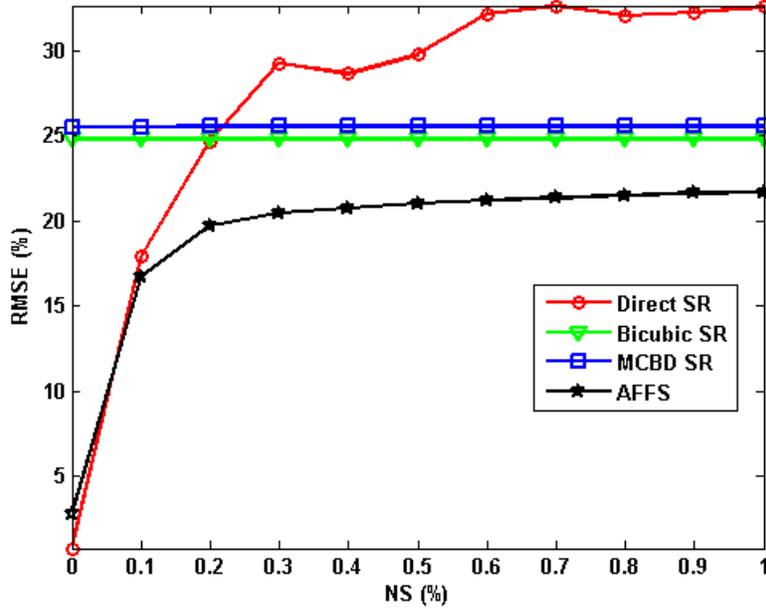


Fig. 2.11. Reconstruction error vs. noise of various SR techniques and the AFFS results for the sharpness circle.

A visual assessment of the reconstruction results, presented in Fig. 2.12, shows that despite the noisy restorations, there are still additional fine details resolved by the direct SR technique as compared with the bicubic and MCBBD SR techniques. The pixelated artifacts reside in the high-frequency regions while washed-out areas exist in the low-frequency regions of the direct SR reconstructed image; thus, the trained filtering masks succeeded in assigning the low and high-frequency locations to the reference SR techniques instead leading to the donut shape in the direct mask w_{Direct} encoded in the red channel of the displayed masks. More noise means more frequencies to be selected from, or greater weight given to, the reference SR techniques. Note that the noise has no direct impact on generating the point spread function's matrix \mathbf{H} , yet the noise distorts the captured images \vec{g}_k and drives the solution, i.e. the reconstructed image, far from the optimal answer.

	Sample LR Image	HR Image (at 2 \times)	Masks @ NS = 0.1%	Masks @ NS = 0.5%
Sharpness Circle				
	Direct SR 13.607%	Bicubic SR 26.7715%	MCBD SR 25.5453%	AFFS 13.7139%
NS = 0.1%				
	25.8734%	26.572%	20.3299%	19.101%
NS = 0.5%				

Fig. 2.12. Visual comparison of the reconstructed sharpness circle (zoomed) at NS = 0.1% and 0.5% for various SR techniques and averaged over 20 noise realizations. A sample LR image and the ground-truth HR image are shown on the left of the first row. The three trained multi-gray masks (w_{Direct} , w_{Bicubic} , and w_{MCBD}) are color encoded as red for direct, green for bicubic, and blue for MCBD in the composite RGB masks displayed on the right of the first row.

2.8.3. Assessing the Impact of Blurring When Imaging Through Different Observation Models

Simulating out-of-focus imaging can be done by widening the blurring kernel (i.e. increasing σ_{blur}), which in turn affects \mathbf{H} globally besides blurring the LR measurements. As a test case, the boat image (bottom left image in Fig. 2.9) was imaged through two different observation models, sensor shift (SS) and camera shift (CS), at various σ_{blur} values (in LR pixel units).

The general trend in all RMSE curves, plotted in Fig. 2.13, is the increment in RMSE as more blurring is introduced. Comparing the RMSE curves of the two observation models, insignificant impact can be seen in the results of the bicubic and MCBD SR techniques as they are not relying directly on \mathbf{H} in their calculation of the reconstructed image. However, when using the direct SR or the AFFS, the performance of the CS model outperforms the SS model since the blurring in SS precedes and hence impacts the shifting, causing an additional source of error.

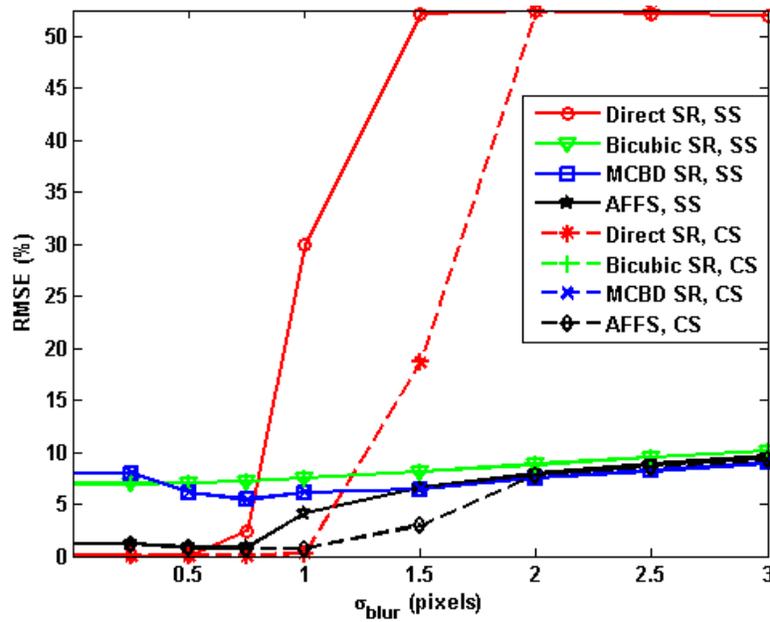


Fig. 2.13. Reconstruction errors vs. σ_{blur} when imaging the boat image through two observation models: solid curves for sensor shift (SS) and dashed curves for camera shift (CS) using various SR techniques.

The blurring through SS results in dappled artifacts in the reconstructed image of the direct SR technique, yet many fine details survived and were later chosen by the filtering masks (trained for different σ_{blur} values and observation models and averaged over the other training images), leading to much richer content (e.g., see the outrigger poles of the boats) than both the bicubic and MCBD SR

techniques (see the third row in Fig. 2.14). The same blurring amount in the CS model seems to be ineffective; thus, the masks favored the direct SR output, leading to sharp reconstruction.

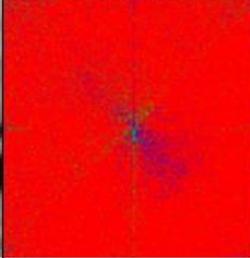
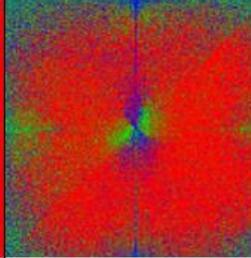
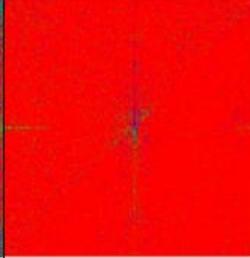
	Sample LR Image @ SS	Masks @ SS, $\sigma_{\text{blur}} = 0.75$	Masks @ SS, $\sigma_{\text{blur}} = 1$	Masks @ CS, $\sigma_{\text{blur}} = 1$
Boat				
	Direct SR	Bicubic SR	MCBD SR	AFFS
	2.428%	7.226%	5.3902%	0.8039%
SS, $\sigma_{\text{blur}} = 0.75$				
	29.9366%	7.4964%	6.1377%	4.0942%
SS, $\sigma_{\text{blur}} = 1$				
	0.2487%	7.4962%	6.1381%	0.7242%
CS, $\sigma_{\text{blur}} = 1$				

Fig. 2.14. Visual comparison of the reconstructed boat at $\sigma_{\text{blur}} = 0.75$ and 1 pixel for various SR techniques using sensor shift (second and third rows, respectively) and $\sigma_{\text{blur}} = 1$ using camera shift (forth row). A sample LR image is shown on the left of the first row. The three trained averaged masks (w_{Direct} , w_{Bicubic} , and w_{MCBD}) are color encoded as red for direct, green for bicubic, and blue for MCBD in the composite RGB masks displayed on the right of the first row.

2.8.4. Validating the Replicating Boundary Assumption Made in the Direct SR Technique

The subsampling operator D inherently makes the boundary values part of the generation of \mathbf{H} at the starred PSF values h^* ; see Fig. 2.5. A replication boundary condition is assumed in order to avoid having more unknowns than the number of equations, so the full-rank property of \mathbf{H} is preserved. To verify the impact of this assumption, the boundary values are derived from normal distributions $\mathcal{N}(\text{nearest image values}, \sigma_{\text{bound}}^2)$ while varying the boundary tolerance BTol related to the standard deviation $\sigma_{\text{bound}} = \text{BTol} \times L$, where L is the maximum allowed intensity. Again, the filtering masks to be used here are trained for each BTol and averaged over 20 boundary realizations.

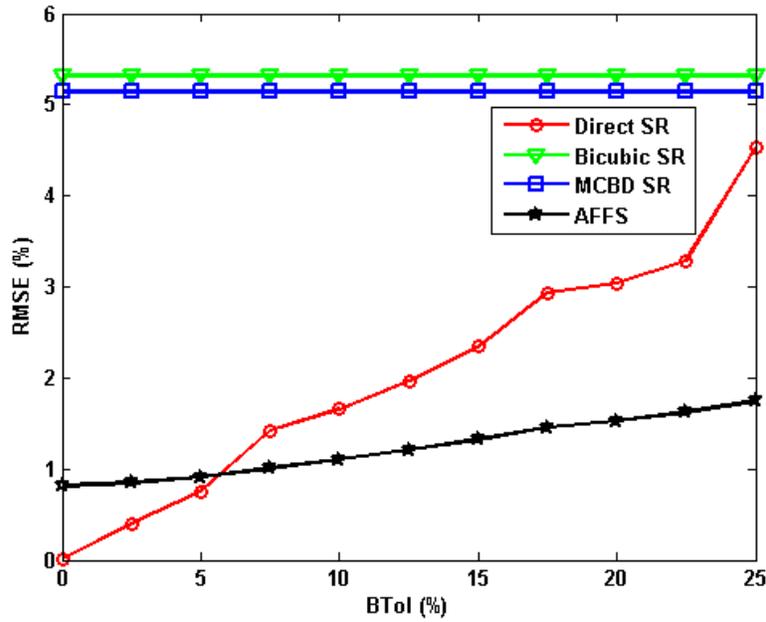


Fig. 2.15. Reconstruction errors vs. BTol when reconstructing the peppers image with various SR techniques.

The RMSE curves and visual reconstructions in Fig. 2.15 and Fig. 2.16 show that the direct SR technique is still capable of resolving fine details even with large BTol, which validates the boundary assumption made earlier. The additional

filtering step, AFFS, indeed assists in eliminating many of the boundary artifacts. Note that both the bicubic and MCBD SR techniques are totally insensitive to the boundary conditions, and their HR reconstructions are either too smooth (in the bicubic case) or have artificial ripples (in MCBD case; see the edges of the objects in the peppers image).

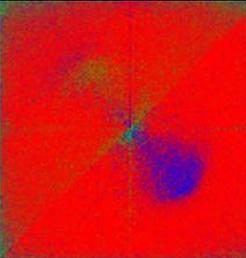
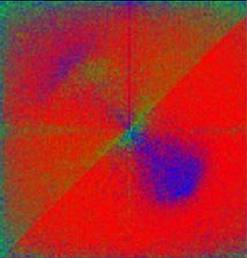
	Sample LR Image	HR Image (at 2×)	Masks @ BTol= 10%	Masks @ BTol=20%
Peppers				
	Direct SR 1.74%	Bicubic SR 6.0205%	MCBD SR 6.3612%	AFFS 1.2239%
BTol = 10%				
	3.3776%	6.0205%	6.361%	1.7945%
BTol = 20%				

Fig. 2.16. Visual comparison of the reconstructed peppers (zoomed) at BTol=10% (middle row) and 20% (bottom row) for various SR techniques. A sample LR image and the original HR image are shown on the left of the first row. The three trained multi-gray masks (w_{Direct} , w_{Bicubic} , and w_{MCBD}) are color encoded as red for direct, green for bicubic, and blue for MCBD in the composite RGB masks displayed on the right of the first row.

2.8.5. Evaluating the Impact of the Shift Misestimates on the Direct SR Restorations

The direct SR technique requires accurate subpixel shift estimation for all captured LR images as any tiny error in the estimated shift vectors affects the

generation of \mathbf{H} globally. For the $2\times$ subsampling case, there are 6 random variables (in X-Y directions) that need to be estimated with respect to one of the captured LR images selected as a reference. To test how accurate the shift estimate needs to be, the estimated shifts are set to follow normal distributions $\mathcal{N}(\text{true shifts}, \sigma_{\text{shift}}^2)$ while varying the shift tolerance STol related to the standard deviation $\sigma_{\text{shift}} = \text{STol} \times \text{DetectorPitch}$. Similarly, the filtering masks are trained for each STol and averaged over 20 shift misestimate realizations.

The RMSE curves and visual reconstructions in Fig. 2.17 and Fig. 2.18 show that direct SR technique is significantly sensitive to the STol even as low as 0.1%. Yet, there are still some fine details that survived (see blood and black edges in the reconstructed endoscopy images) and were successfully picked up by the additional AFFS (see red channels in the filtering masks) without propagating any of the shift artifacts.

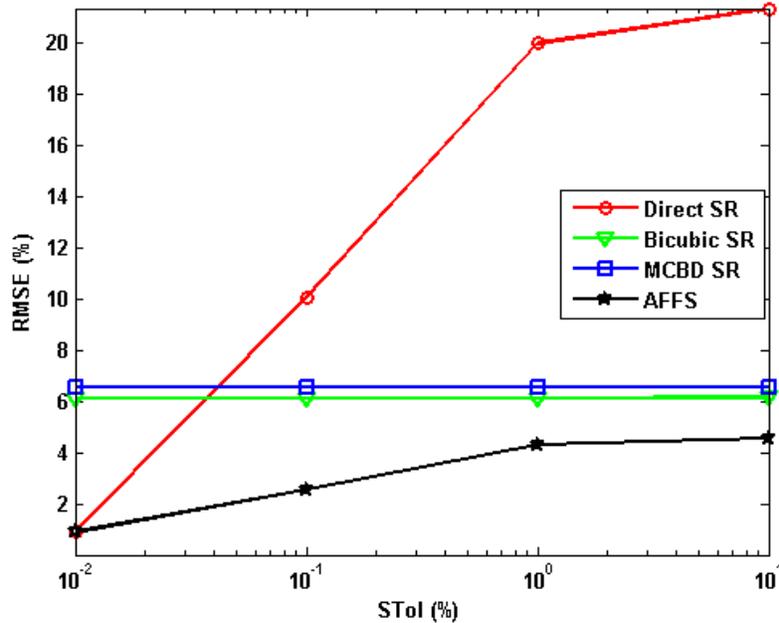


Fig. 2.17. Reconstruction errors vs. STol when restoring the endoscopy image using various SR techniques.

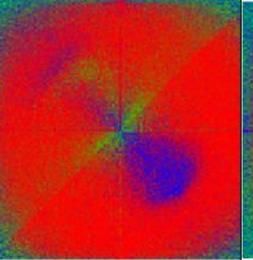
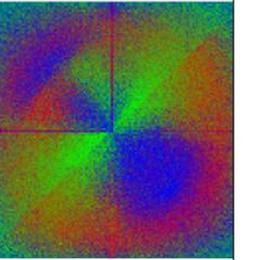
	Sample LR Image	HR Image (at 2×)	Masks @ STol = 0.1%	Masks @ STol = 1%
Endoscopy				
	Direct SR	Bicubic SR	MCBD SR	AFFS
	6.2906%	5.1328%	7.1281%	1.4463%
STol = 0.1%				
	17.454%	5.1342%	7.1281%	4.1558%
STol = 1%				

Fig. 2.18. Visual comparison of the reconstructed endoscopy image (zoomed) at STol=0.1% (middle row) and 1% (bottom row) for various SR techniques. A sample LR image and the original HR image are shown on the left of the first row. The three trained multi-gray masks (w_{Direct} , w_{Bicubic} , and w_{MCBD}) are color encoded as red for direct, green for bicubic, and blue for MCBD in composite RGB masks displayed on the right of the first row.

2.8.6. Evaluating the Performance at a Combination of Non-idealities

As a final evaluation, the SR techniques are tested over two combinations of non-idealities in the imaging conditions; Set1 (NS = 0.1%, BTol = 10 %, STol = 0.1 %) and Set2 (NS = 0.5%, BTol = 20 %, STol = 1 %). The reconstructions are carried using SS observation model, $2\times$ SRFactor, $\sigma_{\text{blur}}=0.5$ pixel, and averaged over 100 realizations. The filtering masks are trained for each set and averaged over 100 statistical realizations.

The visual reconstructions, shown in Fig. 2.19, demonstrate the fine features (e.g., elements 2 and 3 of group 2) recovered by the direct SR while both Bicubic and MCBBD SR techniques failed in that. AFFS again picked such details with minimal artifacts leading to lowest RMSE errors. Having significant non-idealities (e.g., Set2) will lead to a degraded performance in direct SR results and AFFS reconstructions will be similar to those produced by the reference SR techniques.

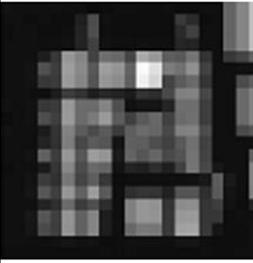
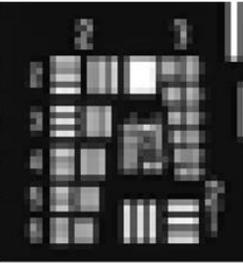
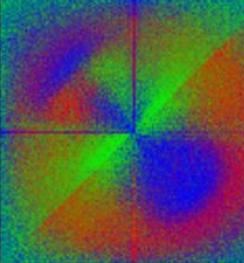
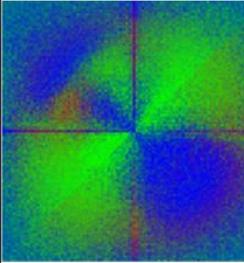
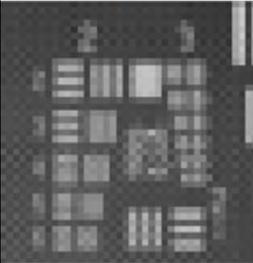
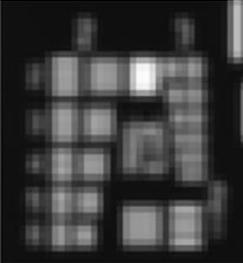
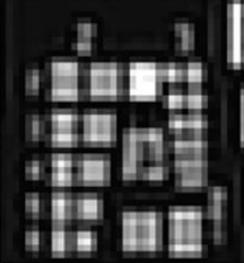
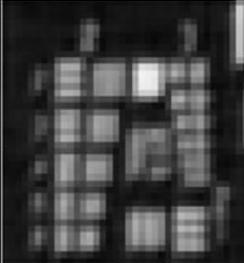
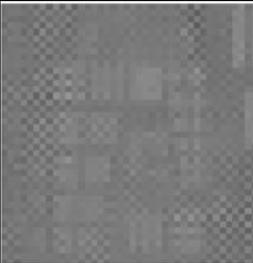
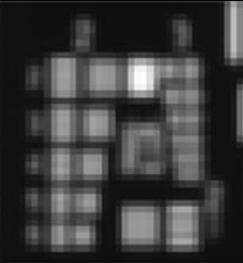
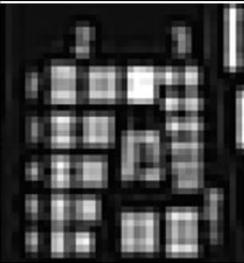
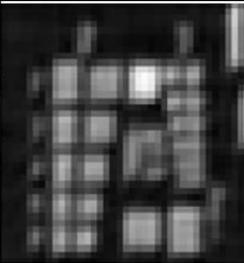
	Sample LR Image @ Set1	HR Image (at 2 \times)	Masks @ Set1	Masks @ Set2
USAF				
	Direct SR 18.5559%	Bicubic SR 15.1819%	MCBD SR 17.9145%	AFFS 11.286%
Set1				
	33.8852%	15.1818%	17.9583%	13.8382%
Set2				

Fig. 2.19. Visual comparison of the reconstructed USAF (zoomed) at two combinations of non-idealities; Set1 (middle row) and Set2 (bottom row) for various SR techniques. A sample LR image and the original HR image are shown on the left of the first row. The three trained multi-gray masks (w_{Direct} , w_{Bicubic} , and w_{MCBD}) are color encoded as red for direct, green for bicubic, and blue for MCBBD and displayed on the right of the first row.

2.9. Conclusion and Future Work

Direct superresolution (SR) is a non-regularized technique that uniquely solves a set of linear equations representing the multi-shift image reconstruction problem with sufficient measurements to deliver realistic reconstructions without any inherent bias imposed by priors, regularizers, and optimization schemes. This approach guarantees an optimal ground-truth solution in a well-posed ideal imaging conditions (in-focus noiseless measurements with known relative shifts, smoothly varying boundaries, and static scene).

An adaptive frequency-based filtering scheme (AFFS) is introduced to gain robustness against out-of-focus, shift misestimate, boundary variations, and noisy scenarios while maintaining the merit of direct SR to achieve optimal restorations without being penalized by the improper assumptions made in the inverse problem. AFFS selects at each frequency a weighted combination of discrete Fourier transform (DFT) values of reconstructed reference SR and direct SR images based on a trained masking information to output, after inverse DFT operation, a filtered image recovering the fine details besides suppressing the reconstruction artifacts.

The simulation studies show that the developed direct SR technique is less sensitive to the boundary variations as compared to the blurring, shift misestimates, and noisy environments. Yet in all simulated non-idealistic reconstruction conditions, direct SR is still capable of resolving extra details even though they were embedded in various artifacts. AFFS then successfully retrieves

these details from various artifacts leading to better reconstructions than those resulted from the SR techniques considered in the AFFS scheme.

As a future work, I will continue to address physical imaging issues in the direct SR technique and the AFFS approach to make them applicable for the physical experiment where a combination of all imperfections exists. The development will address the following additional challenges:

2.9.1. The Motion Blur and Dynamic Scenes

Each super resolved image requires sufficient low resolution (LR) images taken at multiple non-integer-pixel shifts for the exact scene. Thus for dynamically changing scenes, the multi-shift acquisition process should be fast enough so the scene appears static over the time extent that the shifted LR images are captured within. Also, the number of successive measurements may be minimized and the linear least-squares solution can be computed using Moore-Penrose pseudoinverse or linear minimum mean square error techniques. Furthermore, the whole multi-shift acquisition process has to be repeated at least 30 times if real-time acquisition is required. This raises issues about the stability and motion blur caused by the moving actuators during the acquisition. Accounting for jittering and the motion blur can either be done through digital image stabilization techniques between the multi-shift images as preprocessing stage or be done by estimating the relative motion blur vectors and adding them to the translating shifts S_k in the observation model to be part of the computed multi-shift point spread function.

2.9.2. Shift Type

To improve the resolution content in two dimensions (2D), two translating actuators are traditionally implemented in X-Y directions to impose the required non-integer-pixel shifts. This adds to the complexity of the imaging system and demands accurately estimating two variables per shifted image essential for any SR technique to function. To relax these constraints, a single rotating stage may replace the need for the X-Y motion and simplify the shift estimate process to only one variable per image, which is the rotation angle. The origin of rotation can also be part of the investigation to guarantee non-integer radially increasing shifts for all image pixels.

2.9.3. The Chromatic and Off-Axis Aberrations

Ray tracing or experimental calibration procedure can be followed to find the impulse response associated with RGB wavelengths over the field of view in a non-shifted imaging setup. Once fully characterized, the spatio-spectral variant point spread function is modified according to various estimated shifts and then applied locally on various image pixels and color channels within the direct SR technique to accurately super resolve the scene.

2.9.4. The Scanning Filtering Path and the Selection Criteria

The current adaptive training approach adopted in AFFS generates the masking information at each frequency in an outwardly spiral trajectory starting at the zero frequency. At each frequency, it identifies which of the direct SR reconstruction and reference SR reconstructions has the minimum reconstruction error based on training images at the estimated forward observation model parameters.

The simulation results using the outward spiral path show good filtering performance, however, in ideal scenarios the spiral path is favoring the reference reconstructed images at few frequency locations leading to suboptimal performance. This suggests non-optimality of the spiral path, hence other scanning trajectories need to be examined to achieve a better filtering results. One possibility might be a multiplexed random scan where multiple masking locations, selected at random from the DFT of direct SR reconstruction and reference SR reconstructions, will favor either SR technique according to the selection criteria.

The selection criteria itself can be further optimized so that minimizing the root mean square error (RMSE) may be substituted by a task-based metric that consider only the effectiveness of the SR techniques regardless the deblurring or denoising aspects (usually enhanced by different priors and regularizers). For instance in a noisy environment, the extra smoothness imposed by the bicubic SR technique may mistakenly let it score better RMSE than the direct SR although the later one may include more fine features, hence it should be the one to be chosen. The type of the task (e.g., recognition or classification) highlighted in such metric is an open question and part of a future investigation.

2.9.5. Categorized Training and Reference SR Techniques

The training images can be selected from the same category that the objects of interest belong to whether the SR application is face recognition, endoscopic imaging, or others. The training set should be captured experimentally under the same imaging conditions so the filtering masking information becomes more

representative. Also as seen in the simulations, the reconstruction errors of AFFS is upper bounded by the implemented SR techniques, hence, utilizing a variety of state-of-art SR techniques can help to further improve the filtering performance. Finally, a more thorough characteristic study about the AFFS masks will be taking into consideration the impact of the number and the category of the training images, the various implemented SR techniques, the type and the tolerance of parameters' misestimate, and the type of the scanning filtering path.

3. Computational Depth-Based Deconvolution for Full-Focus Imaging

3.1. Introduction

Conventional optical systems have a limited imaging capability beyond their depth of field (DOF) set by the numerical aperture of the optics and the pixel pitch of the focal plane array. Researchers have traditionally invested in the imaging optics and the acquisition mechanism to extend the DOF by making an imaging system that holds a depth-invariant point spread function (PSF). This has been accomplished either by reducing the imaging aperture [42], stacking multiple images taken at different focus distances [43], sweeping the focus during the image acquisition [44], or utilizing diffusion or wavefront coding photography [45]. However, the insertion, modification, or even translation of optical elements may result in reduction of signal-to-noise ratio (SNR), degradation in the captured resolution, longer acquisition time, larger physical size, and additional monetary cost.

To avoid many of the hardware-based limitations, software-based deblurring approaches have been developed to output an all-in-focus scene from a single image captured by ordinary cameras, trading off the hardware complexity for the computational cost. This has been achieved by multi-scale deconvolution [46], two-phase processing for edge and texture restoration [47], blur map estimation with sharp edge prior and guided image filter [48], edge detection and focus map estimation with blind deconvolution [49], and adaptive local regularization to impose edge and smoothness priors [50, 51]. Most such techniques imprecisely approximate the PSF as a two-dimensional (2D) depth-invariant entity which

leads to suboptimal restorations with ringing artifacts especially when objects included in the captured scene belong to a broad range of depth planes.

Ordinary fixed-focus cameras, such as low-end hand-held cameras, smartphone cameras, and surveillance cameras, have a depth-variant PSF that can be visualized as double cones when focusing at finite distances or as a single cone when focusing at infinity with apex at the on-axis in-focus point located at the hyperfocal distance. Objects become progressively more defocused as their distance to the lens gets smaller than half of the hyperfocal distance [4]. The prior knowledge of the volumetric PSF can be utilized effectively to deblur the captured image at different depth planes.

Image deconvolution is a computational technique that undoes the optical blurring to retrieve the true scene. Deconvolution, in its simplest implementation, can be carried out by a division between the captured image and the blurring kernel in the Fourier domain [52]. However, several factors can inevitably cause a boost in noise and ringing artifacts in the deconvolution process: the missing boundary information that participated in the blurring process, the discontinuity between the image boundaries, the distortion of high-frequency components of the sharp edges, the division by weak kernel values in the Fourier domain (due to PSF profiles of large support when deblurring close objects), and the unmatched depth planes between the image regions and the deconvolving PSF kernels. Besides the ringing artifacts, traditional deconvolution techniques also narrow the DOF as they usually blindly approximate the PSF as a 2D depth-invariant entity.

A framework for non-blind depth-based deconvolution [53, 54] is developed so that it computationally brings a captured blurred scene to focus at different depth planes by deconvolving it with axial slices of a pre-characterized PSF volume. A full-focus image is then formed by merging the focused details across all axially deconvolved images based on focus decision metrics or depth map prior. The ringing artifacts associated with the deconvolution are addressed on three levels. The boundary artifacts are eliminated by adopting a block-tiling approach [55]. The sharp edges' artifacts are adaptively controlled by reference maps working as local regularizers through an iterative deconvolution process [50]. Finally, artifacts initiated by neighboring depth-transition surfaces are suppressed by block-wise deconvolution approach or depth-based masking approach if the depth map is available.

3.2. Depth-Based Deconvolution Technique

A three-dimensional (3D) object f is blurred by an imaging PSF kernel h to form an image g that is captured along with noise n . Here, h acts as a 3D-to-2D mapping operator as shown in the following imaging process:

$$g(x_i, y_i) = \int dz_o \iint dx_o dy_o h(x_i, y_i, x_o, y_o, z_o) f(x_o, y_o, z_o) + n(x_i, y_i)$$

where o and i subscripts denote object and image coordinates, respectively.

Realizing the depth-dependent nature of the PSF, a computational depth-based deconvolution technique is developed, presented in Fig. 3.1. The technique inputs the single blurred image g , the multi-dimensional PSF h , and a depth map if available, and outputs a fully focused image I . First, the blurred image is padded

by interpolated smoothed blocks to eliminate the boundary artifacts. Next, the PSF volume is sliced axially into various depth planes z_d (where $d = 1, 2, \dots, N$) and fed along with the padded blurred image into an iterative adaptive non-blind deconvolution stage. The algorithm iterates to get a better estimate of a reference map employed to locally regularize a smoothing prior, which in turns balances the suppressible ringing and the resolvable details in the deconvolved image I_{z_d} associated with the applied axial PSF slice. The axially deconvolved images undergo a focus detection process to locate the focused points across them in case the depth map is not provided. Note that details in each axial image are only brought to focus if they are located at the exact depth plane defined by the applied deblurring PSF slice. Finally, these focused points are stitched together to form a full-focus image I .

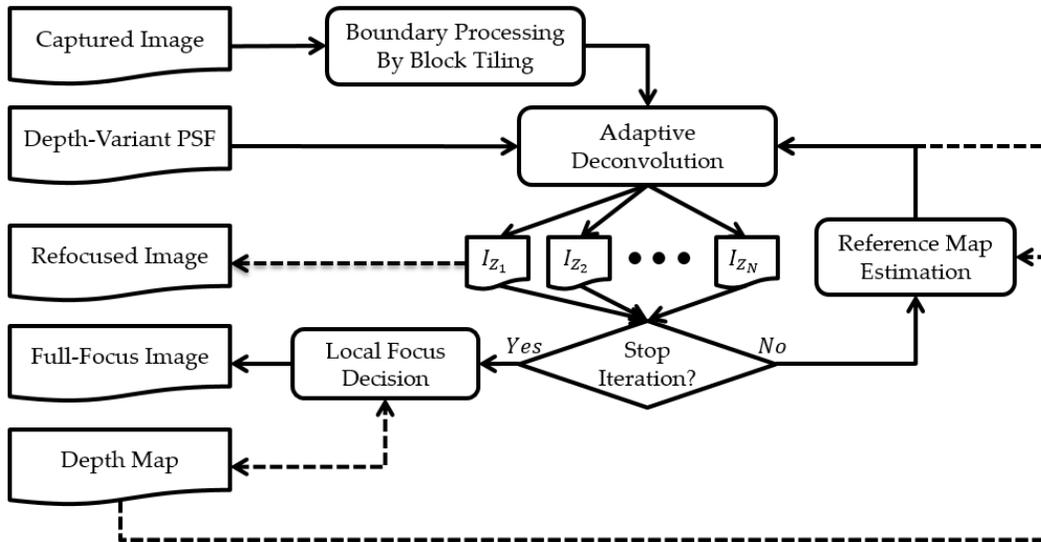


Fig. 3.1. Block diagram of the computational depth-based deconvolution technique.

Note that the proposed depth-based deconvolution technique differs from the 2.5 and 3D deconvolution techniques [56-58] which operate on a stack of images and demand a larger computational cost. For instance, deconvolving by 2D axial

PSF profiles as proposed in the developed technique is a highly parallelized process which outputs refocused images without any further processing.

3.3. Addressing the Ringing Artifacts

The deconvolution's ringing artifacts are suppressed on three levels; block tiling to eliminate boundary artifacts, local regularization to reduce ringing initiated by sharp edges, and block-wise deconvolution or depth-based masking to mitigate artifacts raised by neighboring depth-transition surfaces.

3.3.1. Tackling the Boundary Artifacts

Deconvolution is simply a division in the Fourier domain between the image and the blurring kernel. However, the discrete Fourier transform (DFT) assumes data periodicity which is not present within the captured images. Moreover, parts of the scene beyond the field of view (FOV) participate in the blurring process but are missing in the captured images. These cause the boundary artifacts to propagate through the deconvolved images, and the artifacts become more severe when the PSF profile is of a large support (usually for close objects that are far from the hyperfocal plane). Constrained image extrapolation, smoothing gradients, and block tiling [51, 55] may be applied to achieve periodicity without introducing discontinuities or sharp edges at the borders that may additionally contribute to the ringing artifacts.

Inspired by the tiling scheme presented in [55], the content of tiled blocks is calculated by a simple linear interpolation between the given boundary values, followed by adaptively blurring incrementally toward the center of the tiled blocks to impose smoothness. These pre-processing steps are sufficient to reduce

boundary artifacts without the need for heavy computations to optimize the objective function given in [55] especially when large images are being processed.

3.3.2. Tackling the Edges Artifacts

Ringings initiated by the image content is typically controlled by various regularization techniques such as Tikhonov regularization [59] which may also over-smooth the image, iteratively reweighted least squares [60] which requires hundreds of iterations, Gaussian scale mixtures [61] with conjugate gradient optimization which is a slow solver to converge, total variation [62] with Laplacian prior which may not exactly fit the image gradients, and adaptive sparse priors [63] implementing smoothness via first and second-order derivative filters. The image gradients of real-world scenes usually follow distributions of heavy tails that may be better approximated by hyper-Laplacian priors [64, 65] although being much slower than the Gaussian and Laplacian priors.

Assuming a Gaussian noise model and hyper-Laplacian prior, the maximum a posteriori (MAP) estimator after computing the negative log becomes a minimization problem:

$$\min_f \sum_{j=1}^{\#Pixels} \left((h * f - g)_j^2 + \eta \left(|(\delta_x * f)_j|^\alpha + |(\delta_y * f)_j|^\alpha \right) \right)$$

where $*$ denotes a convolution, j indexes the image's pixels, η is a scalar regularization, α is the hyper-Laplacian order, and δ_x and δ_y represent the horizontal and vertical gradient operators.

Giving η a higher value means favoring more ringing smoothing over edge sharpening. To gain a better control, local regularization based on a reference map [50, 51] (η becomes a 2D matrix) can adaptively assign larger weights to the smooth regions where ringing suppression is of interest and smaller weights for the sharp edges where maximizing the details is desired. The reference maps are found through an iterative deconvolution process to refine the edge strength calculation and assign lower weights to the texture regions and higher weights to the smooth regions. Solving the minimization problem can be done efficiently in the frequency domain through an alternating minimization procedure [66, 67], and analytical derivations can be found for certain α values (e.g., 1/2, 2/3) [65], reducing the computational effort.

3.3.3. Tackling the Depth-Transition Artifacts

In the case of multi-depth (MD) objects, surfaces belonging to different depth planes will be blurred uniquely with their corresponding blurring kernels. Deconvolving these blurred surfaces by unmatched axial PSF profiles will result in ringing denoted here as depth-transition artifacts. Two different approaches are proposed to deal with this challenge: The first approach relies on local application of the deconvolution operator while the second one relies on object's depth map to eliminate depth transitions in the axially deconvolved images.

3.3.3.1. *Block-Wise Deconvolution*

In this approach, the deconvolution is being applied on distinct blocks (or overlapping blocks for finer results) rather than on the image as a whole. This assumes that the blocks' content is most likely to belong to same depth plane especially when these blocks are small enough. As a result, ringing artifacts

initiated by different depth surfaces may be reduced. This approach also assumes that objects are smoothly varying in depth, hence, complex scenes of depth-varying fine-texture such as trees with leaves and branches will be problematic.

3.3.3.2. Depth-Based Masking Approach

Tackling depth-transition artifacts for complex scenes requires prior knowledge about the depth map, which can be acquired from a depth sensor or a stereo system. The depth map provides the masking information needed to prevent simultaneous processing of surfaces belonging to different depth planes, hence eliminating the source of such artifacts. Fig. 3.2 summarizes the depth-based masking steps proposed to suppress the depth-transition artifacts. First, the masking information is extracted from the depth map and applied to the raw blurred image to output N masked images. The masked-out (black) areas are then filled by bilinear interpolation to maintain continuity with the mask's boundaries. Scattered data interpolation [68] is used so masked-out regions of any shape for complex scenes can be inpainted. Afterward, the interpolated areas are blurred with the exact PSF profile to be used at the deconvolution step. This step will assure smoothness in the interpolated regions and guarantee matching with the deconvolving kernel so no depth-transition artifacts are induced. Finally, the N masked interpolated blurred images are adaptively deconvolved by the corresponding PSF profiles using reference maps which have also been masked accordingly (so interpolated areas are considered smooth regions during the regularization) to produce N axially deconvolved images to be merged later to form the full-focus image.

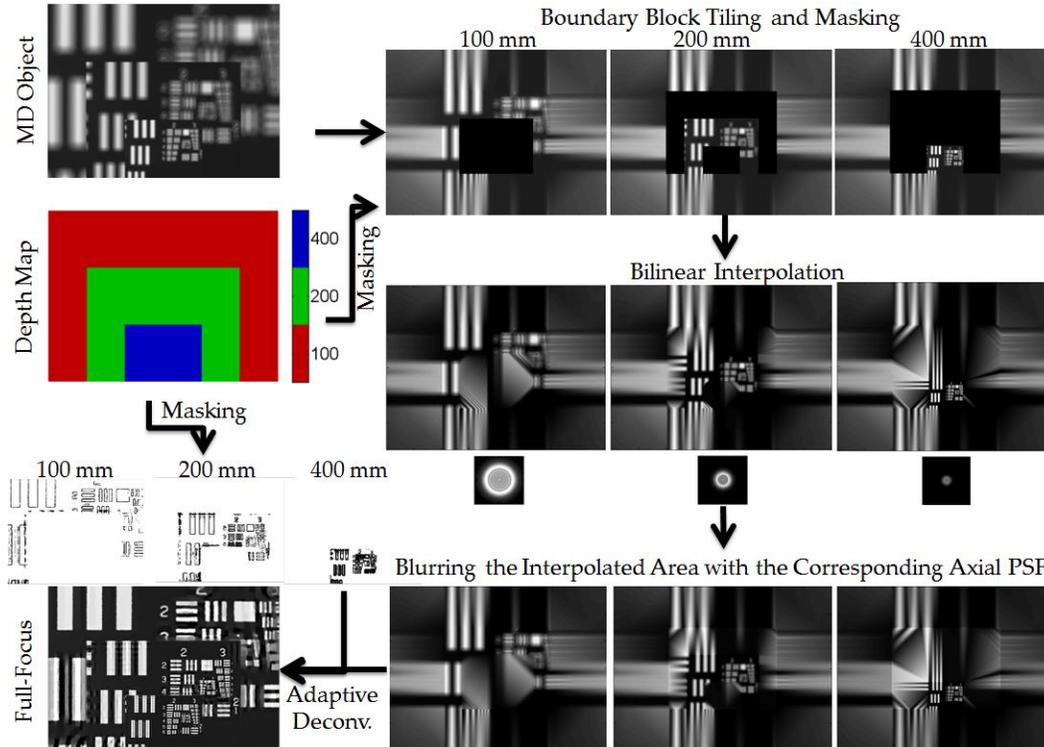


Fig. 3.2. Sketch illustrating the developed depth-based masking approach for planar USAF target set at the three depth planes utilized in the simulation study.

3.4. Local Focus Decision Metrics

In the event that depth map is not given as prior information, it can be estimated by identifying the depth planes of the focused details along the axially deconvolved images, see Fig. 3.3. This may be carried out block-wise reliably using predefined focusing metrics that exploit the fact that a block at its best focus has the maximum contrast [69], the highest energy after high pass filtering, the maximum absolute coefficient values in the wavelet transform domain [70], and the least ringing content. A survey of various focus metrics including gradient-, Laplacian-, wavelet-, statistics-, and discrete cosine transform- based operators can be found in [71]. An overlapping block-wise approach may also be adopted to have a consistent focusing decision across the neighboring pixels while maintaining a pixel-precision decision. Once the depth map is known, the focused

details are extracted and stitched together to produce a 2D full-focus image, giving the ordinary camera an extended DOF capability.

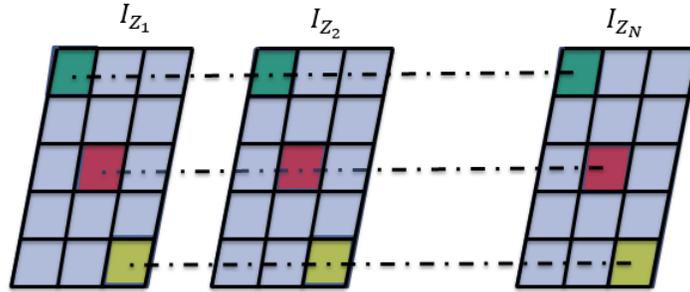


Fig. 3.3. Block-wise depth map estimation by finding the best focus over the axially deconvolved blocks. Samples of the axial blocks are encoded in different colors.

3.5. Simulation Results

A ray-tracing program, OpticStudio by Zemax, is utilized to simulate the imaging of planar objects placed at different depth planes through an off-the-shelf fixed-focus objective lens (4.5 mm focal length, 2.4 f-number, and 1.5 m hyper-focal distance) to a focal plane array of resolution 600×800 pixels with pitch size $1.4 \times 1.4 \mu\text{m}$. The blurring kernels have been similarly estimated by calculating the Huygens PSF [72] of on-axis points placed at the desired depths. The 1951 U.S. Air Force (USAF) resolution test target and cameraman image are chosen to assess the ringing artifacts and the resolved details achieved by the proposed deconvolution technique. The blurred images along with their associated axial PSF profiles, displayed in Fig. 3.4, show the variation in FOV, magnification, and amount of blur as the depth varies.

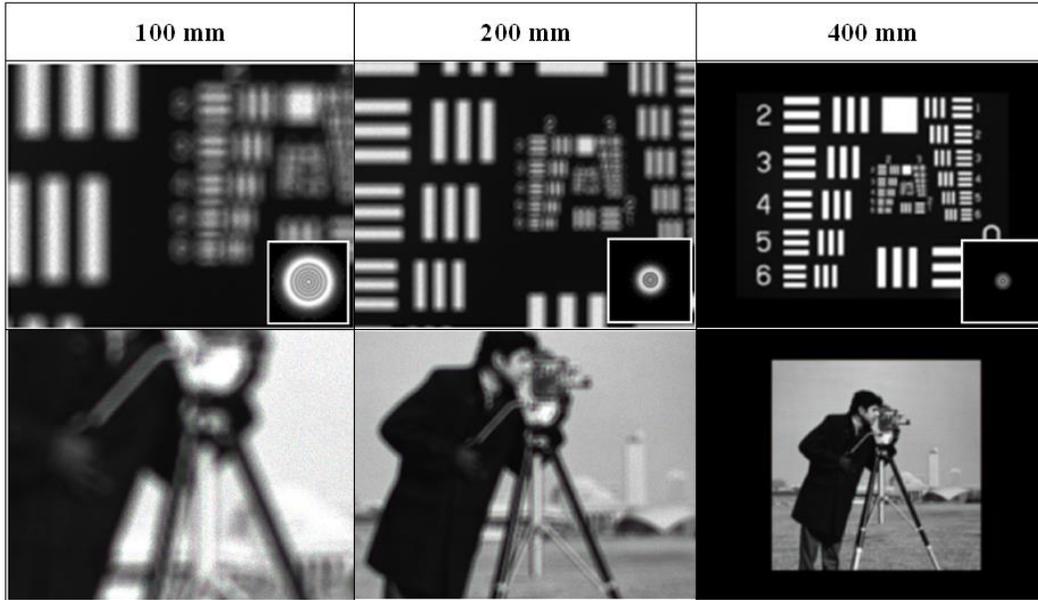


Fig. 3.4. USAF (first row) and cameraman (second row) images captured at various object distances. The associated axial PSF profile of 64×64 pixels is shown at the bottom-right corner of each USAF image.

To imitate MD objects, different regions of planar objects traced at various depth planes are stitched together according to a customized depth map, presented in Fig. 3.5.



Fig. 3.5. Stitching planar regions according to a depth map (left) to form MD objects for USAF (center) and cameraman (right).

The adaptive deconvolution technique [50], reported here, operates using the regularization parameters: 3 iterations, 3 reference levels, hyper-Laplacian $\alpha = 2/3$, threshold $T = 0.1 \times \max$ value of the reference map, and $\eta \in [10^{-5}, 10^{-3}]$. Note that in the case of color images, the deconvolution is applied to the luminance channel only, which is afterward recombined with the chrominance channels. The deconvolution numerical errors are quantified using two metrics:

the structural similarity index measure (SSIM) [73] which mimics the human visual system to extract information based on structure, and the root mean square error (RMSE) which quantifies the errors on a pixel level compared to a reference image. For illustration purposes, the various ringing artifacts are addressed one at a time through the following simulation scenarios.

3.5.1. Impact of Pre-Processing the Boundaries of the Blurred Images

The impact of pre-processing the image boundaries is evaluated in three scenarios: zero padding by 32 pixels (i.e., half the blurring kernel) along each boundary, edge tapering by blurring the boundaries with the exact PSF profile to be used for deconvolution, and block tiling to impose continuity and smoothness across the boundaries. The blurred images for close USAF and cameraman objects after pre-processing the boundaries are presented in Fig. 3.6 and their deblurred results using adaptive regularization technique are displayed in Fig. 3.7.

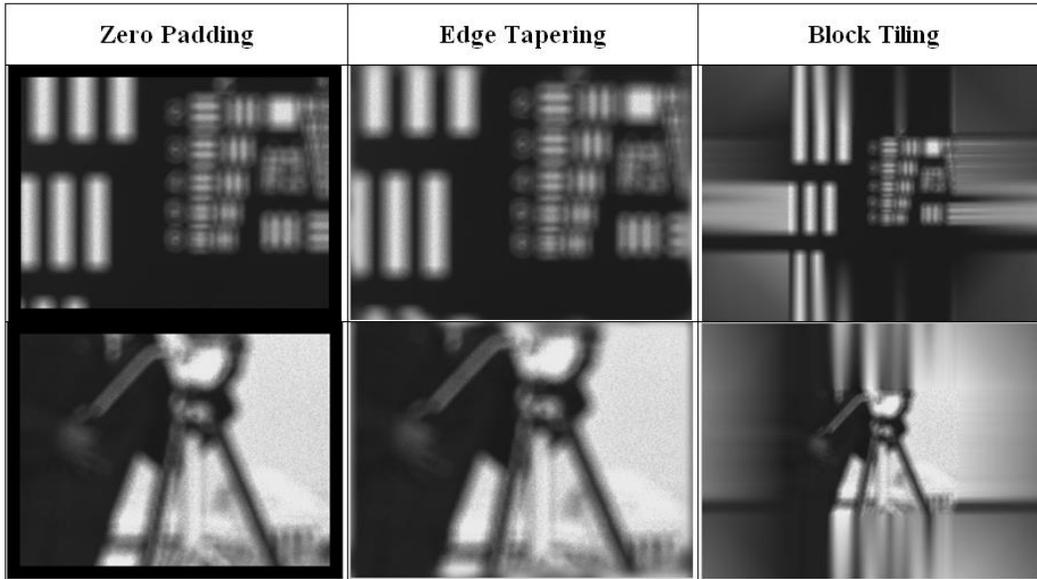


Fig. 3.6. Blurred images after preprocessing the boundaries with various techniques: zero padding (left), edge taper (center), and block tiling (right) for USAF (first row) and cameraman (second row) objects captured at 100 mm distance.

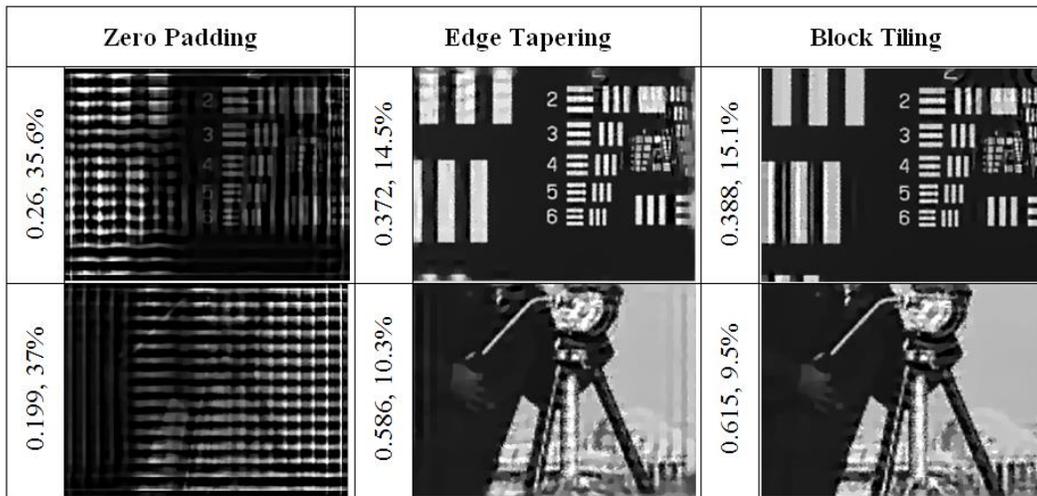


Fig. 3.7. Comparison of deblurred images under various boundary preprocessing techniques using adaptive regularization for planar USAF (first row) and cameraman (second row) objects at 100 mm distance. The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.

Fig. 3.7 shows how a simple zero padding (and similarly in the case of no boundary pre-processing) can lead to ringing all over the deconvolved images especially when the PSF has large support. Tapering the boundaries can help reduce the ringing, but rippling can still be observed near boundaries (see both USAF and cameraman deblurring results in Fig. 3.7). These results demonstrate

that the block tiling approach is superior with the highest SSIM values and barely noticeable boundary artifacts.

3.5.2. Impact of Adaptive Regularization for Planar Objects set at Different Depth Planes

To study the impact of adaptive regularization at different depth planes, the planar objects shown in Fig. 3.4, captured at various depth planes were deconvolved by their corresponding axial PSF profiles. Four deconvolution techniques are evaluated: the Lucy-Richardson Bayesian-based iterative method [74, 75] with 100 iterations, Fortunato's deconvolution using sparse adaptive priors [63] with regularization parameters (0.001, 20, 0.033, 0.05), Krishnan's method with hyper-Laplacian prior [65] of $\alpha = 2/3$ and scalar regularization $\lambda = 10^4$, and Tsai's adaptive deconvolution [50] with the regularization parameters defined at the beginning of section 3.5. The parameters are chosen so the deconvolution techniques can resolve the same level of details (as quantified by USAF) so the ringing appearance can be fairly judged between them. The visual reconstructions and numerical errors for USAF and cameraman images are shown in Fig. 3.8 and Fig. 3.9, respectively. The regularizing reference maps of Tsai's algorithm are shown at the last row of each figure, where the white regions represent the areas to be smoothed and the black regions represent the edges to be preserved.

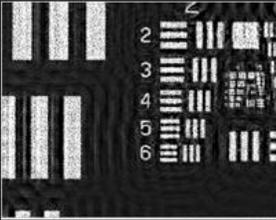
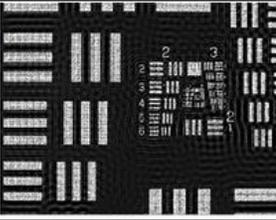
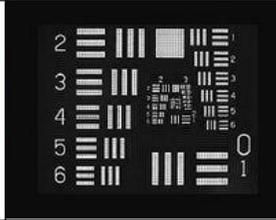
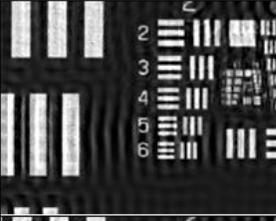
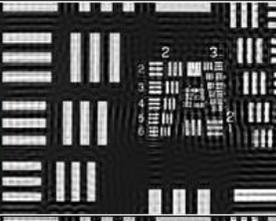
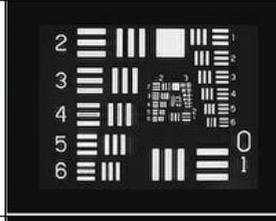
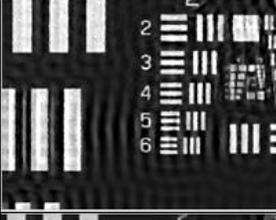
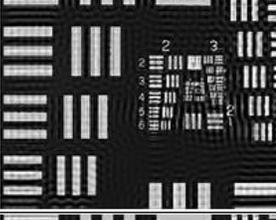
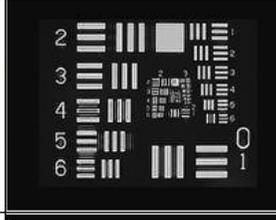
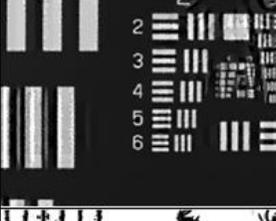
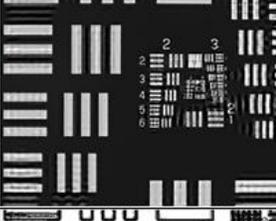
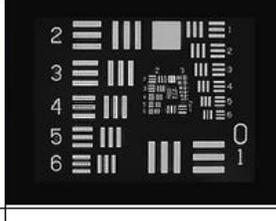
	100 mm	200 mm	400 mm
Lucy	0.211, 16.6% 	0.24, 19.1% 	0.291, 14.2% 
Fortunato	0.356, 14.6% 	0.337, 15.4% 	0.319, 12.4% 
Krishnan	0.351, 15% 	0.328, 16.8% 	0.371, 14% 
Tsai	0.388, 15.1% 	0.391, 15.8% 	0.439, 13.4% 
Ref. Map			

Fig. 3.8. Visual comparison of different deconvolution techniques for USAF target at different depth planes when deconvolved with the corresponding axial PSF profiles. The last row shows the final reference maps used as local regularizers in Tsai's technique. The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.

		100 mm		200 mm		400 mm
Lucy	0.16, 19.4%		0.169, 22.8%		0.132, 16%	
Fortunato	0.411, 15.4%		0.614, 7.4%		0.31, 9%	
Krishnan	0.386, 17.7%		0.572, 8.9%		0.246, 10.3%	
Tsai	0.615, 9.5%		0.723, 7.5%		0.333, 9.6%	
Ref. Map						

Fig. 3.9. Visual comparison of different deconvolution techniques for cameraman image at different depth planes when deconvolved with the corresponding axial PSF profiles. The last row shows the final reference maps used as local regularizers in Tsai’s technique. The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.

A visual and numerical comparison of the reconstruction results in Fig. 3.8 and Fig. 3.9 demonstrate the effectiveness of Tsai’s adaptive deconvolution over other techniques in suppressing the edge artifacts and the noisy appearance while maximizing the image details and sharpness. This becomes more noticeable as the objects get far from the DOF range (e.g., when the power of the blurring kernel is

spread over a larger support in the case of shorter imaging distances). Note that the adaptive deconvolution results in a higher structural similarity (SSIM) as compared with others. The RMSE metric seems to be inaccurate in reflecting the visual quality of Tsai's adaptive deconvolution as compared with Fortunato's deconvolution since RMSE is more related to errors on a pixel level (which cannot capture the ringing presence) rather than the overall appearance of the deconvolved images.

3.5.3. Impact of Block-Wise Deconvolution

Blocks, extracted from the captured blurred MD object then pre-processed by the block tiling approach, are adaptively deconvolved (using Tsai algorithm) by various axial PSF profiles and then stitched together to form the axially deconvolved images, shown in Fig. 3.10. Two different block sizes 100×100 and 200×200 pixels are investigated to evaluate the deconvolution results. Smaller blocks will better suit the assumption made in the block-wise deconvolution to reduce the depth-transition artifacts (e.g., see the left vertical bars of USAF when deconvolving with PSF @ 100 and 200 mm). However, the deconvolution process itself requires sufficient extent of spatial regions in order to converge to the right solution (e.g., see the three vertical bars at the upper left corner of USAF when deconvolving with PSF @ 100 mm).

	PSF @ 100 mm	PSF @ 200 mm	PSF @ 400 mm
Blocks of 100×100	0.187, 44.3%	0.174, 30.9%	0.177, 29.1%
	0.375, 33%	0.441, 21%	0.361, 20.6%
Blocks of 200×200	0.219, 39.8%	0.181, 31.2%	0.191, 28.9%
	0.352, 32.2%	0.432, 21.2%	0.366, 20.6%

Fig. 3.10. Visual comparison of block-wise deconvolution results for MD USAF and cameraman images using two different block sizes of 100×100 pixels (first two rows) and 200×200 pixels (last two rows) when deconvolved with different axial PSF profiles (sorted column-wise). The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.

3.5.4. Impact of Depth-Based Masking Approach

Given the depth map, shown on the left in Fig. 3.5, the regions belong to different depth planes can be masked out according to the procedure described in section 3.3.3.2. To evaluate the effectiveness of depth masking, bilinear interpolation, and blurring by the associated axial PSF profile, the resulting axially deconvolved images (last two rows in Fig. 3.11) are compared against deblurring results when no depth processing is implemented (first two rows in Fig. 3.11). It is clear from the visual restorations, in addition to the low RMSE and the high SSIM values,

that many of the depth-transition ringing artifacts are successfully filtered out by the depth-based masking approach.

	PSF @ 100 mm	PSF @ 200 mm	PSF @ 400 mm
Without Any Depth Processing	0.218, 42.9% 	0.185, 32.1% 	0.273, 24.9%
	0.314, 35.4% 	0.409, 23% 	0.619, 16.6%
Depth-Based Masking	0.671, 12.9% 	0.792, 8% 	0.956, 6.2%
	0.785, 11.2% 	0.863, 8% 	0.944, 8.9%

Fig. 3.11. Visual comparison of deconvolution results for MD USAF and cameraman images without any depth processing (first two rows) and with depth-based masking approach (last two rows) when deconvolved with different axial PSF profiles (sorted column-wise). The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.

3.5.5. Comparison of Full-Focus Deconvolution Results after Stitching

The final step to form a full-focus image is merging the axially deconvolved images according to a given or estimated depth map, presented on the left in Fig. 3.5. To completely isolate the depth-transition artifacts, raw planar single-depth (SD) images, shown in Fig. 3.4, are block tiled, adaptively deconvolved by Tsai's technique with their corresponding PSF profiles, and finally stitched

according to the depth map (see second row in Fig. 3.12) to mimic the MD deconvolution scenario. On the other extreme, the depth-transition artifacts can dominate the MD deconvolution results if no special processing is involved (see third row in Fig. 3.12). The depth-based masking approach can help a lot in reducing such artifacts (see last row in Fig. 3.12), although it is not fully restoring some of the image details (e.g., compare the vertical bars on the right of restored USAF images between the depth-based masking case and the SD deblurred case).

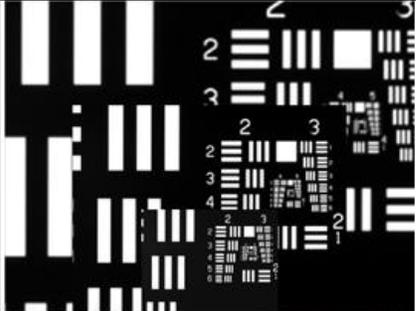
Reference MD Objects	1, 0%		
Deblurring SD Objects then Stitching	0.471, 16.3%		
Deblurring MD Object without Processing	0.332, 28.3%		
Deblurring MD Object with Depth-Based Masking	0.434, 16.4%		

Fig. 3.12. Visual comparison of final full-focus deconvolution results after stitching based on a depth map for USAF and cameraman images showing reference MD objects (first row), deblurring results when considering SD objects followed by stitching to mimic MD object (second row), deblurring results of MD object without any depth processing (third row), and deblurring results of MD object with depth-based masking approach (last row). The numerical errors in terms of SSIM and RMSE are shown to the left of each image, respectively.

3.6. Conclusion and Future Work

This chapter introduces a computational depth-based deconvolution technique that equips ordinary cameras with an extended depth-of-field capability by deconvolving the captured scene with pre-calibrated depth-variant point spread function (PSF) profiles. This can bring the blurred scene to focus at different depth planes. Afterward, the focused features from the axially deconvolved images are stitched together based on prior or estimated depth map to form a full-focus image. The developed technique can accomplish the tasks of depth map estimation, image refocusing, and full-focus imaging simultaneously.

The ringing artifacts have been analyzed and addressed on three levels. The boundary artifacts are eliminated by adopting a block-tiling approach that imposes continuity and smoothness along the boundaries. The edge artifacts are adaptively controlled by iterative reference maps which penalize the ringing in the smooth regions and favor the details in the textured regions. Finally, artifacts initiated by neighboring depth-transition surfaces are suppressed by a block-wise deconvolution or depth-based masking approach to isolate processing surfaces belonging to different depth planes simultaneously, hence eliminating the source of such artifacts over a wide depth range. The effectiveness of the proposed ringing processing is demonstrated for planar objects and multi-depth objects.

Future work will experimentally validate the performance of the developed deconvolution technique on continuous multi-depth objects captured by a real camera. Also, other filtering schemes to deal with depth-transition artifacts will be explored while relaxing the need for a depth map as an extra prior. Moreover, to

eliminate the need for frequent calibration as the depth-variant PSF volume varies with time, temperature, or due to mechanical shock, a hybrid deconvolution scheme will be investigated to combine the adaptive depth-based deconvolution with iterative blind deconvolution techniques so the PSF volume is kept updated with minimal calibration effort.

4. Multi-Polarization Fringe Projection Imaging for High Dynamic Range Objects

4.1. Introduction

Shape acquisition of three-dimensional (3D) objects is of significant importance for various real-world applications including machine vision, reverse engineering, industrial inspections, and medical imaging. An economic reliable real-time technique that delivers such information is fringe projection imaging [76-78]. The imaging system comprises a projector-camera pair in which successive phase-shifted fringe patterns are projected onto objects, become distorted, and are then captured by the camera. These captured distorted fringes carry valuable information about the object's depth, which can be retrieved through phase shifting algorithms [76, 77].

However, conventional structured light imagers fail to recover depth data from objects of high dynamic range (HDR) where fringe visibility is greatly reduced in dark regions, bright areas, or over surfaces of large reflectivity variations. For instance, shiny metal objects reflect illuminating light specularly and saturate the camera without carrying any depth content.

Researchers have tackled these challenges through various approaches, achieving restricted success. Employing polarizers to filter the projected and captured fringe images results in a reconstruction trade-off between complementary reflecting surfaces. A crossed polarizer-analyzer pair [79] eliminates shiny areas, but with the cost of unresolving the dark zones. On the contrary, a parallel polarizer-analyzer alignment [80, 81] can maintain good fringe quality at the dark regions, but not in the bright places.

A different approach that better suits HDR cases without such a trade-off is to take multiple-shots of fringe images at various exposures [82] or automatically adapt the exposure times to fit the scene [83]. In these schemes, areas of dark intensities are picked from the long exposure patterns while regions of bright appearance are chosen from the short exposure ones. An alternative technique to avoid saturation and maintain good fringe quality is to adaptively adjust the projected fringe pattern intensities and then combine the captured fringes [84]. A more developed fringe acquisition approach [85] combines both different camera exposures and various fringe projection intensities to guarantee good depth recovery. However, taking multiple shots or adaptively adjusting either the exposure times or the projected fringe intensities, or both, may not permit fast capture of dynamic scenes, hence limiting the technique to slow or static scenes.

In this chapter, a single-shot multi-polarization fringe projection (MPFP) algorithm [86] that combines the advantages of most previous solutions is developed, allowing broader applications. Unlike prior techniques, the novelty of the developed approach is the use of snapshot multi-polarization measurements to process HDR dynamic scenes. Additionally, the MPFP algorithm can easily exploit combined-exposure measurements if further enhancement is desired.

4.2. Multi-Polarization Fringe Projection Imaging Algorithm

In the MPFP imaging system, shown in Fig. 4.1, the projected fringes are linearly polarized prior to incidence on the object and being captured after reflection through a multi-polarization camera. The employed camera has a pixelated polarizer array of four states ($P0^\circ$, $P45^\circ$, $P90^\circ$, and $P135^\circ$) attached to the

sensor [87-89]. Upon reflection, various object surfaces will modulate the polarized fringes differently, leading to dissimilar measurements in the multi-polarized channels.

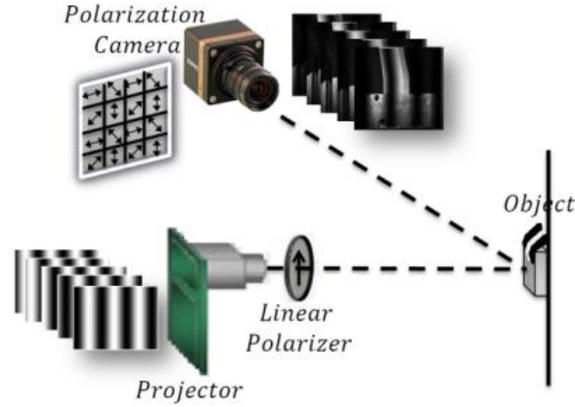


Fig. 4.1. Multi-polarization fringe projection (MPFP) imaging system.

The imaging equation of the proposed MPFP system can be mathematically described by

$$\vec{g}_{p,k} = H_p \vec{f} \vec{I}_k + \vec{n}_{p,k} \quad (1)$$

where p and k subscripts represents the polarization and fringe indices, respectively. Also, \vec{g} , \vec{f} , \vec{I} , and \vec{n} are the measurements, object, fringe, and noise vectors, respectively. Finally, H forms the impulse response projection matrix that is dependent on the captured polarization state.

The aim of this imaging technique is to render the shape of HDR objects \vec{f} from the captured multi-polarization distorted fringes $\vec{g}_{p,k}$ through the six steps presented in Fig. 4.2.

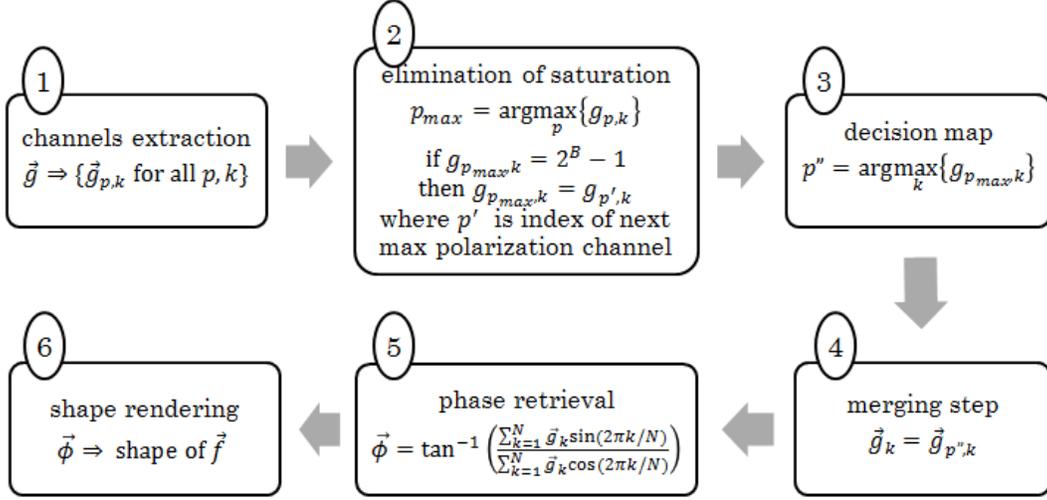


Fig. 4.2. Steps of multi-polarization fringe projection technique (MPFP) for HDR objects, vectors denote image-level operations while scalars denote pixel-level operations.

The algorithm starts by extracting M raw polarized images out of the B -bit sensor measurements for each of N sinusoidal fringe patterns. Next, steps 2 to 4 target eliminating any saturation and improving the fringes' quality by finding the best polarization channel for each pixel across all fringe images. For each pixel at the k^{th} fringe image, index of the maximum polarized channel is identified as $p_{max} = \arg \max_p \{g_{p,k} \text{ for } p = 1, \dots, M\}$. If $g_{p_{max},k} = 2^B - 1$ then the maximum polarized channel is saturated at that pixel belonging to fringe k . In this case, the saturated channel is replaced by the next largest channel, p' , across all fringe images: $g_{p_{max},k} = g_{p',k}$ for all $k = 1, \dots, N$ (can be repeated if there are multiple saturated channels as normally seen in specular reflection regions). If all channels are saturated, then the algorithm may simply pick any channel, and the fringes' distortions cannot be restored (here it is recommended to recapture the fringe images using shorter exposure time to avoid any loss). Once saturation has been eliminated, a maximization decision map of selected channel indices is found: p''

$= \arg \max_k \{g_{p_{max},k} \text{ where } k = 1, \dots, N\}$. This map identifies the maximum non-saturated channel for each pixel. Next, the decision map is utilized to merge the M modified polarized images located in the same fringe pattern into a single high-contrast image, $\vec{g}_k = \vec{g}_{p^n,k}$. This results in a total of N enhanced fringe images. It is important to note that when the algorithm picks a certain polarization channel at a particular pixel, this channel will be utilized at the corresponding pixel in all other fringes to maintain the sinusoidal modulation needed for accurate depth estimation.

Afterward, the phase ϕ can be retrieved from the N merged distorted fringes through a phase shifting algorithm given by the following arctan equation [80]:

$$\vec{\phi} = \tan^{-1} \left(\frac{\sum_{k=1}^N \vec{g}_k \sin(2\pi k/N)}{\sum_{k=1}^N \vec{g}_k \cos(2\pi k/N)} \right) \quad (2)$$

The final step is converting the phase information into depths that are calculated with respect to a reference plane determined at the calibration stage. Once depths are calculated, the object shape can be rendered in a 3D coordinate system.

4.3. Evaluation of Multi-Polarization Fringe Projection Algorithm with Various Objects

To validate the performance of the MPFP technique, it had been tested in four different scenarios using $N = 5$ successive equally phase-shifted sinusoidal fringe patterns, $M = 4$ polarization channels, and PolarCam micropolarizer camera [90] with $B = 8$ bits for the sensor resolution and pixelated sensor array of 1208×1608 pixels. The sinusoidal patterns are projected by a digital light projection (DLP

3000) device that has 768×1024 micro-mirrors. The experiments were carried in normal room light conditions where fringe illumination strength and image acquisition time were set to obtain good fringe contrast on the targeted objects.

4.3.1. Simple Object with Three Different Surfaces

A simple custom-made HDR object, shown in Fig. 4.3(a), is formed with three different material surfaces made of a black tape, a white tape, and a metal surface.

The four captured polarization fringe images carry different reflectance content that are presented in Fig. 4.3(b) for the first fringe pattern. To gain more insight on fringe contrast of the raw images, two cross-sections have been taken across the black-white-black tapes and the metal surface; the locations and profiles are illustrated in Fig. 4.3(b) and Fig. 4.3(c), respectively. The fringe profiles elucidate the dynamic range for each surface and the differences between the polarization channels especially at the saturated regions.

If utilized separately, the distorted fringes belonging to the same polarization channel may not be sufficient to reconstruct a complete object's shape, as shown by the rendered images in Fig. 4.3(d). Note that the specular reflections at the $P0^\circ$, $P90^\circ$, and $P135^\circ$ channels prevent any depth estimation since fringes are absent; however, this is not the case in the $P45^\circ$ channel.

When all polarizations are evaluated together, the proposed algorithm attempts to find the best fringe visibility existing in them. Fig. 4.4(a) shows a color-coded maximization decision map that delivers contrast-enhanced fringes after the merging process, presented in Fig. 4.4(b). As seen from the colored decision map and the MPFP cross-section profiles plotted in Fig. 4.3(c), the

technique succeeded in assigning the unsaturated polarization intensity ($P45^\circ$) to the bright metal surface while attaching the relatively high polarization intensity ($P135^\circ$) to the black tape region. Fig. 4.4(c) shows the phase image, which was calculated from the five enhanced fringes and used to find the depths of the targeted object. The whole 3D shape is successfully produced, as shown in Fig. 4.4(d), demonstrating the performance improvement of the MPFP technique over traditional fringe projection, which uses independent polarization channels, as shown in Fig. 4.4(d).

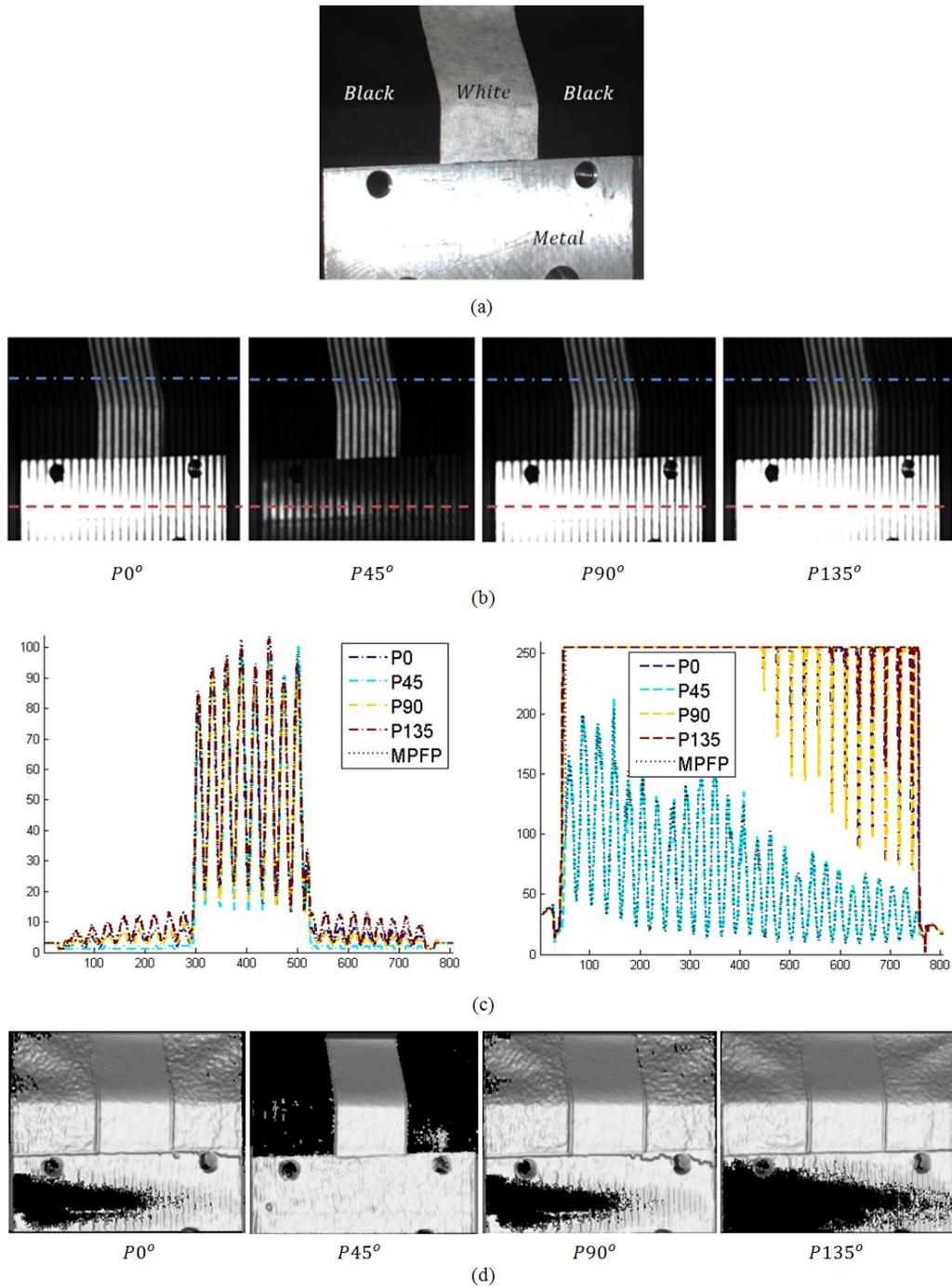


Fig. 4.3. Single-polarization fringe projection imaging of simple object. (a) Simple three-surface object captured by unpolarized camera. (b) Raw polarized data of first distorted fringes. (c) Fringe contrast of various polarization channels at two cross-sections of first distorted fringes (black-white-black tapes on left and metal surface on right). (d) Shape rendering of five fringe images at separate polarizations.

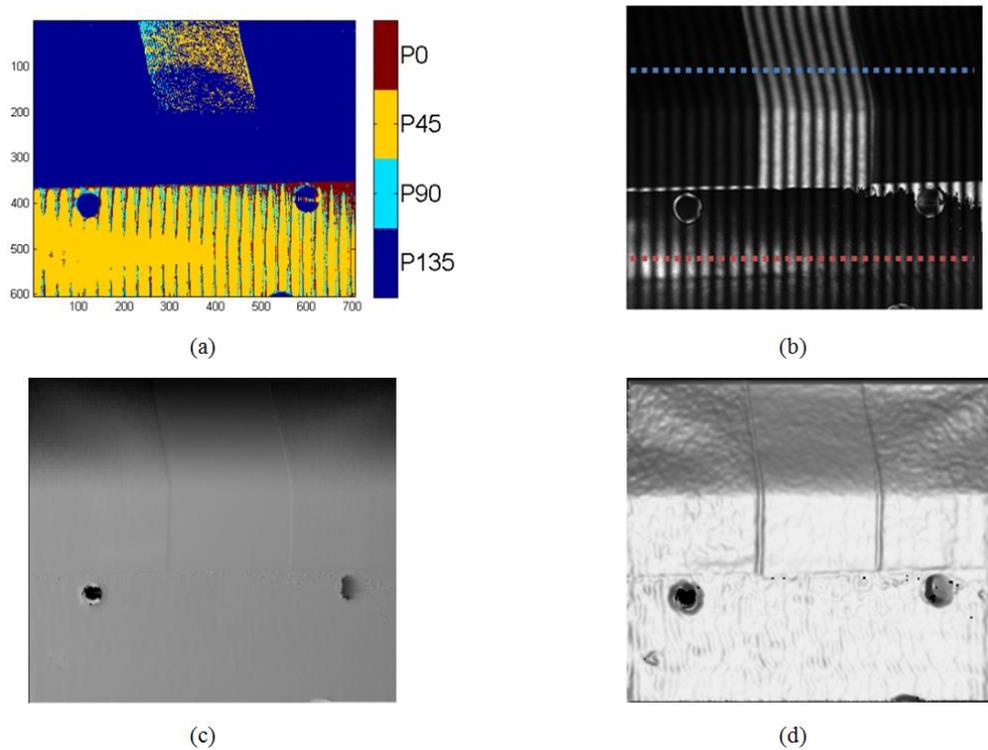


Fig. 4.4. Multi-polarization fringe projection imaging of simple object. (a) Multi-polarization decision map. (b) Merging results of first fringe images. (c) Phase retrieval. (d) Shape rendering of five enhanced fringe images.

The decision map is a result of a pixel-by-pixel maximization procedure after eliminating saturation. Hence, in regions where two or more different polarizers behave similarly (e.g., diffusing surfaces), the decision may alternate between the neighboring pixels, leading to a noise-like appearance in the decision map with negligible impact on the depth estimation, see the white tape area at the top-center of Fig. 4.4(a) where the decision alternates between $P45^{\circ}$, $P90^{\circ}$, and $P135^{\circ}$. Such noisy appearance of the decision map can be removed by making decisions for blocks of pixels instead of individual pixels; however, this may limit the depth resolution.

Also, note that the dynamic range of the merged fringe images has been extended when using four polarization channels as compared with two

perpendicular ones. However, the significance of having further polarization channels (more degrees of freedom) may be reduced as channels of adjacent polarization angles will show insignificant visual differences. Besides, such a setup will either sacrifice the spatial resolution or imply sequential measurements to account for all employed polarizers.

4.3.2. Microscopic Spatial Filter Stage as HDR Object

Another tested HDR scene was the microscopic spatial filter stage, shown in Fig. 4.5(b), which had microscopic objective causing specular reflections and black mounting stage leading to poor fringe visibility. The captured raw polarized images, presented in Fig. 4.5(a), reveal the different intensities acquired for the same fringe image where fringe visibility is the best at $P45^\circ$ channel for the saturated regions and at other polarization channels for the dark regions. The produced MPFP decision map shown in Fig. 4.5(c) reflects these trends by selecting the proper channels for the merging step so high quality fringe images are obtained as shown in Fig. 4.5(d) that enable accurate shape rendering shown in Fig. 4.5(e).

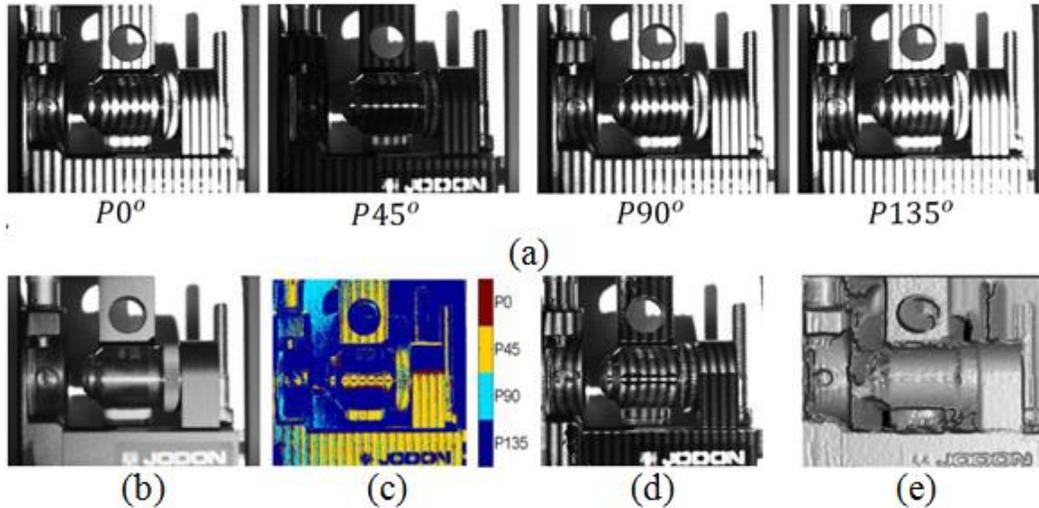


Fig. 4.5. Multi-polarization fringe projection imaging of microscopic spatial filter stage. (a) Raw polarized data of first distorted fringes. (b) Stage captured by regular camera. (c) Decision map. (d) Merging results of first fringe images. (e) Shape rendering of five enhanced fringe images.

4.3.3. Circuit Board of Various Intensities

The MPFP algorithm is further tested by imaging a circuit board, Fig. 4.6(a), as an example of real-world objects that contain a variety of intensity levels. Again, the MPFP decision map, Fig. 4.6(b), successfully assigned the high reflectivity components (metal ports and capacitor tops) to the unsaturated polarization channels and most of the dark background to the brightest channel ($P135^\circ$), producing high-contrast fringe images, Fig. 4.6(c). The rendered shape derived from the enhanced fringe images is presented in Fig. 4.6(d). The false shadows surrounding some components (e.g., capacitors) are not a shortcoming of the MPFP algorithm, but are produced by the components' shadows in the original image, which blocked the projected fringes from reaching these surfaces. A little roughness may still appear on some surfaces where saturation dominates all polarization channels (e.g., bottom metal plugs in Fig. 4.6(d)). This occurs due to the upper truncation of sinusoidal fringes which distorts the phase calculations.

Finally, the depth resolution can be further improved using a higher resolution camera lens and a denser sensor array.

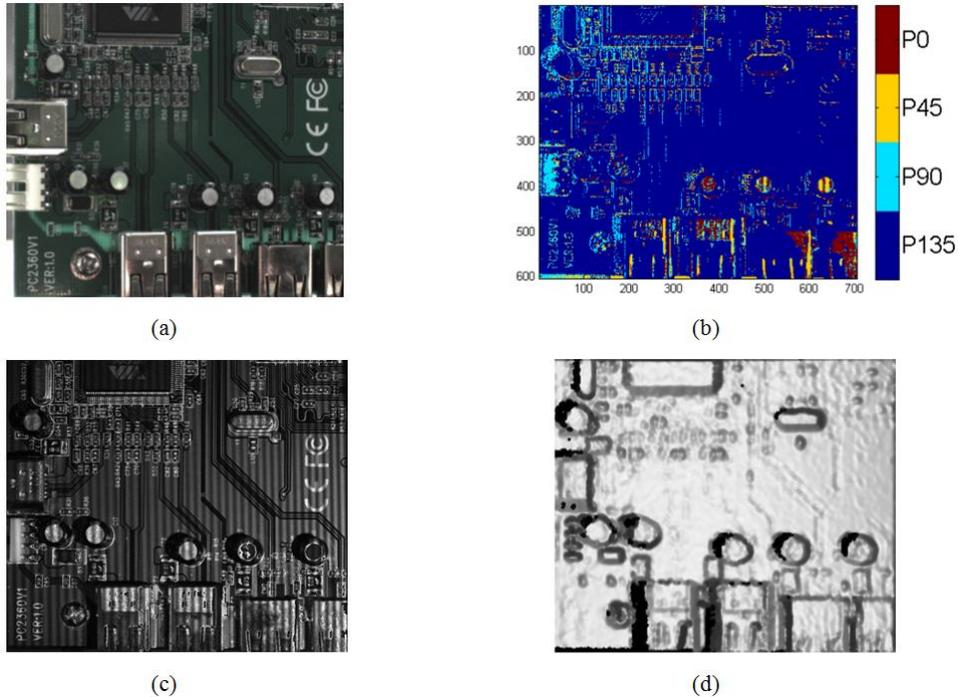


Fig. 4.6. Multi-polarization fringe projection imaging of circuit board object. (a) Circuit board captured by unpolarized camera. (b) Decision map. (c) Merging results shown for first fringe. (d) Shape rendering of five enhanced fringe images.

4.3.4. Object at Different Exposures

A further challenge occurs when the object has a very HDR or is unequally illuminated so that the rendered object's shape is only partially visible even when incorporating all polarization channels acquired under the same exposure. Here, the MPFP algorithm can easily be extended to include sufficient measurements captured under various exposure times and through different polarization channels. The algorithm obtains a more complete view of the object by selecting the exposure and polarization pairs that yield the maximum unsaturated pixels.

As an example, scissors were imaged, Fig. 4.7(a), using four polarization angles and two different exposure times. The left-side images in Fig. 4.7(b) and

Fig. 4.7(c) show that short exposure results in a poor decision map and imperfect rendering at the scissors' handles, whereas the middle images in Fig. 4.7(b) and Fig. 4.7(c) reveal that long exposure produces holes on the metal surfaces due to the limited decision options. When using measurements from both exposure times, MPFP selects the exposure-polarization pair that yields the maximum unsaturated value at each pixel (see Fig. 4.7(b) right), leading to more complete 3D shape rendering (see Fig. 4.7(c) right). This illustrates the combined polarization/exposure capability of MPFP in the very HDR or non-uniform illumination cases.

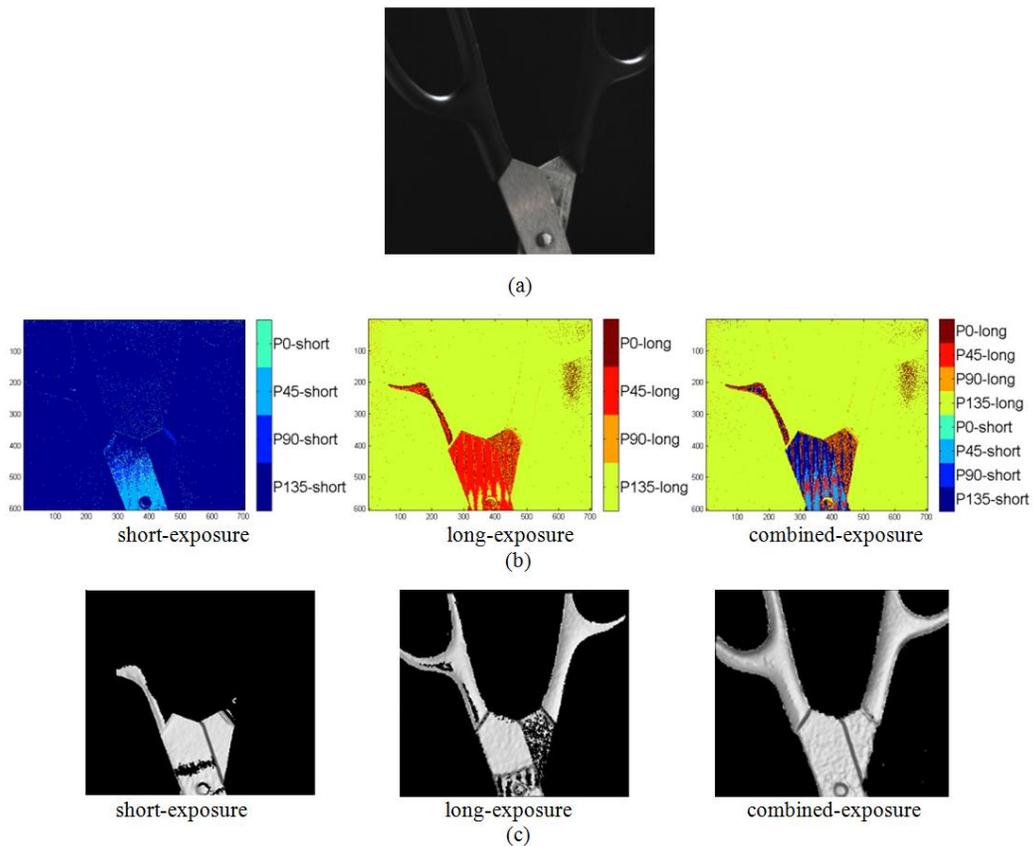


Fig. 4.7. Multi-polarization fringe projection imaging of an object under different exposures. (a) Scissors captured by unpolarized camera. (b) Decision map. (c) Shape rendering. The images shown in (b) and (c) are sorted left to right according to the utilized exposure time.

4.4. Conclusion

In summary, a multi-polarization fringe projection imaging technique is proposed to be capable of delivering complete depth estimation of HDR objects. The algorithm eliminates saturated or low-contrast fringe regions by selecting different polarization measurements, or the right combination of polarization angle and exposure time, in order to maintain good fringe visibility. This leads to greater coverage of the object in the shape rendering, better measurement of object topography, and thus a more accurate rendering of the object shape.

5. Summary

Computational imaging is a powerful concept that brings innovative imaging capabilities to the miniature applications. It enables the development of miniature cameras at a higher performance-to-complexity ratio where camera complexity can be formulated as a function of size, weight, and cost. Moreover, it provides flexibility in the imaging designing space where any trade-off to be made becomes more straightforward to be analyzed and quantified.

In computational imagers, the captured image is optically encoded and later computationally decoded given the knowledge of the forward imaging system to produce a new type of image or a richer representation of the scene. The computational decoding of such measurements can facilitate post-capture control of a variety of imaging parameters, including resolution, depth of field, and dynamic range.

5.1. Summarizing the Direct Superresolution Technique

A non-regularized direct superresolution technique has been developed to uniquely solve a set of linear equations representing the multi-shift image reconstruction problem with sufficient measurements to deliver realistic reconstructions without any inherent bias imposed by priors, regularizers, and optimization schemes. An adaptive frequency-based filtering scheme is introduced to gain robustness against poor focus, misestimated shift, boundary variations, and noisy scenarios while maintaining the merit of direct superresolution to achieve optimal restorations with minimal artifacts. Simulation results demonstrate that more fine features can be resolved with the developed

technique as compared with other superresolution techniques. Simulations also show that the retrieved high-resolution features are successfully transferred by the adaptive frequency-based filtering scheme to produce an artifact-suppressed high-resolution image under realistic imaging conditions.

5.2. Summarizing the Full-Focus Depth-Based Deconvolution

A computational depth-based deconvolution technique has been developed to equip ordinary cameras with an extended depth of field capability. This is achieved by deconvolving the captured scene with pre-calibrated depth-variant point spread function profiles to bring it into focus at different depth planes. Afterward, the focused features from the depth planes are stitched together to form a full-focus image.

The ringing artifacts are processed on three levels. First, the boundary artifacts are eliminated by adopting a block-tiling approach. Second, the sharp edges' ringing artifacts are adaptively controlled by reference maps working as local regularizers through an iterative deconvolution process. Finally, artifacts initiated by different depth surfaces are suppressed by a block-wise deconvolution or depth-based masking approach. The developed algorithm is demonstrated for planar objects and multi-depth objects scenarios.

5.3. Summarizing the Depth Acquisition for HDR Objects

A pixelated polarizer camera has been implemented in a fringe projection system to collect richer brightness information about HDR scene through its four polarization channels. In accordance with the hardware modification, a multi-polarization fringe projection algorithm has been developed to eliminate saturated

or low-contrast fringe regions. This is achieved by selecting different polarization measurements, or the right combination of polarization angle and exposure time according to a decision map to output enhanced fringe images. This leads to greater coverage of the object in the shape rendering, better measurement of object topography, and thus a more accurate rendering of the object shape.

5.4. Closing Thoughts

Although computational imaging techniques seem to be very promising, but there is always a compromise that has to be made in order to advance in other imaging dimensions. For instance, superresolution and fringe projection depth acquisition techniques lack the snapshot functionality as they require a set of measurements captured successively over time.

The design focus of conventional photographic imaging platforms is the collection optics. More and more lens elements are used in the optical design to correct for aberrations. The computational imaging shifts the focus to computation to acquire richer information with a simpler optics. Exceeding the limits of optical systems with minimal compromises will continue to motivate researchers to further innovate in the field of the computational imaging.

References

- [1] M. Shankar, *Sampling and Signal Estimation in Computational Optical Sensors*, Dissertation, ECE Department, Duke University, 2007.
- [2] G. Zheng, *Innovations in Imaging System Design: Gigapixel, Chip-Scale and Multi-Functional Microscopy*, Dissertation at California Institute of Technology, 2013.
- [3] A. W. Lohmann, R. G. Dorsch, D. Mendlovic, Z. Zalevsky, and C. Ferreira, "Space-bandwidth product of optical signals and systems," *J. Opt. Soc. Am. A*, vol. 13, no. 3, p. 470–473, 1996.
- [4] J. E. Greivenkamp, *Field Guide to Geometrical Optics*, vol. 1 (SPIE Press Bellingham, Washington, 2004).
- [5] R. H. Vollmerhausen and E. Jacobs, "The Targeting Task Performance (TTP) Metric A New Model for Predicting Target Acquisition Performance," *Center for Night Vision and Electro-Optics*, 2004.
- [6] S. K. Nayar, "Computational cameras: Approaches, benefits and limits," *Technical Report, Department of Computer Science, Columbia University CUCS-001-11*, 2011.
- [7] C. Zhou and S.K. Nayar, "Computational cameras: Convergence of optics and processing," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3322-3340, 2011.
- [8] S. K. Nayar, "Computational cameras: Redefining the image," *IEEE Computer Magazine, Special Issue on Computational Photography*, pp. 30-38, 2006.
- [9] J. Mait, R. Athale, and J. van der Gracht, "Evolutionary paths in imaging and recent trends," *Opt. Exp.*, vol. 11, no. 18, pp. 2093-2101, 2003.
- [10] K. P. Thompson and J. P. Rolland, "Will computational imaging change lens design?," in *Proc. SPIE 9293, International Optical Design Conference*, 2014.
- [11] T. Lukeš, *Super-Resolution Methods for Digital Image and Video Processing*, CZECH TECHNICAL UNIVERSITY IN PRAGUE, 2013.
- [12] Z. Zalevsky and D. Mendlovic, *Optical Superresolution*, Springer Series in Optical Sciences, Vol. 91, 2003.
- [13] Z. Zalevsky, *Super-Resolved Imaging: Geometrical and Diffraction Approaches*, Springer, 2011.
- [14] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equation*, New York: Dover, 1923.
- [15] A. Gilman and D. G. Bailey, "Near optimal non-uniform interpolation for image super-resolution from multiple images," in *Image and Vision Computing, New Zealand*, pp. 31-35., 2006.
- [16] A. N. Tikhonov and V. I. Arsenin, *Solutions of Ill-posed Problems*, Washington DC: V.H Winston and Sons, 1977.
- [17] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high resolution image from multiple degraded mis-registered low resolution images," *Proceedings of the IEEE International Conference on Image Processing*, pp. 553--557, 1994.
- [18] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621-1633, 1997.
- [19] S. Villena, M. Vega, D. Babacan, R. Molina, and A. Katsaggelos, "Bayesian combination of sparse and non-sparse priors in image super-resolution," in *Digital Signal Processing*, 23(2), 530-541, 2013.
- [20] P. Milanfar, *Super-resolution Imaging*, Digital Imaging and Computer Vision, CRC Press, 2010.
- [21] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays using

- convex projections," *J. Opt. Soc. Am.*, vol. A6, p. 1715–1726, 1989.
- [22] F. W. Wheeler, R. T. Hoxter, and E. B. Barrett, "Super-resolution image synthesis using projections onto convex sets in the frequency domain," *SPIE Symposium on Electronic Imaging, Conference on Computational Imaging*, vol. 5674, p. 479–490, 2005.
- [23] A. Zomet, A. Rav-Acha, and S. Peleg, "Robust super-resolution," in *Proceedings international conference on computer vision and pattern recognition (CVPR)*, 2001.
- [24] F. Šroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," in *IEEE Transactions on Image Processing*, (21)4, pp. 1687--1700, 2012.
- [25] B. Salahieh, J. J. Rodriguez, and R. Liang, "Direct superresolution for realistic image reconstruction," *Optics Express*, (Submitted for publication, 2015).
- [26] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nature Photonics*, vol. 7, p. 739–745, 2013.
- [27] S. Dong, P. Nanda, K. Guo, J. Liao, and G. Zheng, "Incoherent Fourier ptychographic photography using structured light," *Photon. Res.*, vol. 3, no. 1, p. 19–23, 2015.
- [28] S. Dong, J. Liao, K. Guo, L. Bian, J. Suo, and G. Zheng, "Resolution doubling with a reduced number of image acquisitions," *Biomedical Optics Express*, vol. 6, no. 8, 2015.
- [29] S. Pacheco, B. Salahieh, T. Milster, J. Rodriguez, and R. Liang, "Transfer function engineering in reflective Fourier ptychography," *Submitted for publication*, 2015.
- [30] NCAR Advanced Study Program, "Aliasing," [Online]. Available: <http://www.asp.ucar.edu/colloquium/1992/notes/part1/node62.html>. [Accessed 7 2015].
- [31] M. Woods and A. K. Katsaggelos, "A spatial frequency based metric for image super-resolution," *JOSA A*, 2015.
- [32] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on Applied Signal Processing (special issue on Superresolution)*, p. Article ID 71459, 2006.
- [33] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacement," *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pp. pp. 742-746, 1988.
- [34] J. N. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image registration by template matching using normalized cross-correlation," in *IEEE International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pp. 819-822, 2009.
- [35] M. Johansson, Image Registration with Simulated Annealing and Genetic Algorithms, Stockholm, Sweden: Thesis in Computer Science, Royal Institute of Technology, 2006.
- [36] S. Ram and J. J. Rodriguez, "Single image super-resolution using dictionary-based local regression," in *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, San Diego, CA, 2014.
- [37] E. Y. Lam and J. W. Goodman, "Iterative statistical approach to blind image deconvolution," *Journal of the Optical Society of America A*, vol. 17, no. 7, pp. pp. 1177-1184, 2000.
- [38] J. S. Lim, Two-Dimensional Signal and Image Processing, Englewood Cliffs, NJ, Prentice Hall, p. 548, 1990.
- [39] A. K. Kaw and E. E. Kalu, Numerical Methods with Applications, autarkaw, 2010, Chap. 4.
- [40] R. W. Farebrother, Linear Least Squares Computations, CRC Press, 1988.
- [41] B. Salahieh, J. J. Rodriguez, and R. Liang, "Direct superresolution technique for solving a miniature multi-shift imaging system," in *Imaging Systems and Applications, JW3A. 5*, 2015.
- [42] S. W. Hasinoff, "Variable-aperture photography," in *PhD Thesis, University of Toronto, Dept. of Computer Science*, 2008.
- [43] S. W. Hasinoff and K. N. Kutulakos, "Light-efficient photography," in *IEEE Trans. Pattern*

- Analysis and Machine Intelligence*, 33(11), pp. 2203-2214, 2011.
- [44] C. Zhou, D. Miao, and S.K. Nayar, Focal Sweep Camera for Space-Time Refocusing, in *Technical Report, Department of Computer Science, Columbia University*, 2012.
- [45] X. Lin, J. Suo, G. Wetzstein, Q. Dai, and R. Raskar, "Coded focal stack photography," in *IEEE International Conference on Computational Photography (ICCP)*, 1(9), pp. 19-21, 2013.
- [46] J. Sun, H. Y. Shum, L. Yuan and L. Quan, "Removing blur from an image", U.S. Patent, US 8750643 B2, 2014.
- [47] R. Tezaur, "Multiple phase method for image deconvolution", PCT patent application, Pub. No. WO2015017194 A1, 2015.
- [48] Y. Cao, S. Fang, and Z. Wang, "Digital multi-focusing from a single photograph taken with an uncalibrated conventional camera," *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3703-3714, 2013.
- [49] W. Zhang and W. K. Cham, "Single image focus editing," *IEEE 12th International Conference on Computer Vision (ICCV)*, pp. 1947-1954, 2009.
- [50] H. C. Tsai and J. L. Wu, "An improved adaptive deconvolution algorithm for single image deblurring," in *Mathematical Problems in Engineering*, Article ID 658915, 2014.
- [51] J. H. Lee, Y. S. Ho, "High-quality non-blind image deconvolution with adaptive regularization," *J. Vis. Commun. Image, Elsevier*, vol. R. 22, p. 653-663, 2011.
- [52] F. Weinhaus, "Fourier transform processing with ImageMagick - convolution and deconvolution," [Online]. Available: http://www.fmwconcepts.com/imagemagick/fourier_transforms/fourier.html#convolution_deconvolution. [Accessed 8 2015].
- [53] B. Salahieh, J. J. Rodriguez, and R. Liang, "Computational depth-variant deconvolution technique for full-focus imaging," in *Computational Optical Sensing and Imaging, CT3F.5*, 2015.
- [54] B. Salahieh, J. J. Rodriguez, S. Stetson, and R. Liang, "Single-image extended-focus reconstruction using depth-based deconvolution," *Submitted for Publication*, 2015.
- [55] R. Liu and J. Jia, "Reducing boundary artifacts in image deconvolution," in *IEEE International Conference on Image Processing ICIP*, pp.505-508, 2008.
- [56] F. Aguet, D. Van De Ville, and M. Unser, "Model-based 2.5-D deconvolution for extended depth-of-field in brightfield microscopy," *IEEE Trans. Image Processing*, vol. 17, no. 7, pp. 1144-1153, 2008.
- [57] A. Griffa, N. Garin, and D. Sage, "Comparison of deconvolution software in 3D microscopy. A user point of view," *Part I and Part II, G.I.T. Imaging & Microscopy*, vol. 1, pp. 43-45, 2010.
- [58] D. Biggs, "3D deconvolution microscopy," in *Current Protocols in Cytometry, Chapter 12*, John Wiley & Sons Inc., 2010.
- [59] A. Tikhonov, "On the stability of inverse problems," *Doklady Akademii Nauk SSSR*, vol. 39, no. 5, pp. 195-198, 1943.
- [60] C. V. Stewart, "Robust parameter estimation in computer vision," *SIAM Reviews*, vol. 41, no. 3, p. 513-537, 1999.
- [61] M. Wainwright and S. Simoncelli, "Scale mixtures of gaussians and the statistics of natural images," *NIPS*, p. 855-861, 1999.
- [62] A. Chambolle and P. L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, p. 167-188, 1997.
- [63] H. E. Fortunato and M. M. Oliveira, "Fast high-quality non-blind deconvolution using sparse adaptive priors," *The Visual Computer, Springer*, vol. 30, no. 6-8, pp. 661-671, 2014.
- [64] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Transactions on Graphics*, vol. 26, p. 70-77, 2007.

- [65] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," *Neural Information Processing Systems*, 2009.
- [66] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *PAMI*, vol. 4, p. 932–946, 1995.
- [67] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.
- [68] K. Anjyo, J. P. Lewis, and F. Pighin, "Scattered data interpolation for computer graphics," *SIGGRAPH Course Notes*, 2014.
- [69] M. Levoy, "Autofocus: contrast detection," 3 2012. [Online]. Available: <http://graphics.stanford.edu/courses/cs178/applets/autofocusCD.html>. [Accessed 8 2015].
- [70] B. Forster, D. Van De Ville, J. Berent, D. Sage, and M. Unser, "Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images," *Microsc. Res. Tech.*, vol. 65, no. 1-2, pp. 33–42, 2004.
- [71] S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, vol. 46, no. 5, pp. 1415–1432, 2013.
- [72] C. Huygens, "Traité de la Lumiere," (completed in 1678, published in 1690).
- [73] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, p. 600–612, 2004.
- [74] W. H. Richardson, "Bayesian-based iterative method of image restoration," *JOSA*, vol. 62, no. 1, p. 55–59, 1972.
- [75] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astronomical Journal*, vol. 79, no. 6, p. 745–754, 1974.
- [76] S. Zhang, High-Resolution, Three-Dimensional Shape Measurement, in *Ph.D. Thesis, Stony Brook University, Stony Brook, NY*, 2005.
- [77] N. Karpinsky and S. Zhang, "High-resolution, real-time 3-D imaging with fringe analysis," in *Real Time Image Processing*, 7(1) 55-66, 2012.
- [78] X. Su and Q. Zhang, "Dynamic 3-D shape measurement method: A review," *J. Optics and Lasers in Eng.*, vol. 48, no. 2, p. 191–204, 2010.
- [79] L. Wolff, "Using polarization to separate reflection components," in *In Proceedings of CVPR*, pages 363–369, 1989.
- [80] T. Chen, H. P. A. Lensch, C. Fuchs, and H. P. Seidel, "Polarization and phase-shifting for 3D scanning of translucent objects," in *Computer Vision and Pattern Recognition, IEEE Conference on Volume pp1-8*, 2007.
- [81] R. Liang, "Short wavelength and polarized phase shifting fringe projection imaging of translucent objects," in *Opt. Eng.* 53(1), 014104, 2014.
- [82] S. Zhang and S.T. Yau, "High dynamic range scanning technique," in *Optical Engineering* 48(3), 2009.
- [83] L. Ekstrand and S. Zhang, "Autoexposure for three-dimensional shape measurement using a digital-light-processing projector," in *Optical Engineering* 50, no. 12, 2011.
- [84] C. Waddington and J. Kofman, "Saturation avoidance by adaptive fringe projection in phase-shifting 3D surface-shape measurement," in *In Optomechatronic Technologies (ISOT), 2010 International Symposium on*, pp. 1-4. IEEE, 2010.
- [85] H. Jiang, H. Zhao, and X. Li, "High dynamic range fringe acquisition: a novel 3-D scanning technique for high-reflective surfaces," in *Opt Lasers Eng.* 50 (17), 2012.
- [86] B. Salahieh, Z. Chen, J. J. Rodriguez, and R. Liang, "Multi-polarization fringe projection imaging for high dynamic range objects," *Optics express*, vol. 22, no. 8, 2014.
- [87] N. J. Brock, B. T. Kimbrough, and J. E. Millerd, "A pixelated polarizer-based camera for

- instantaneous interferometric measurements," in *Proc. of SPIE 8160, 81600W*, 2011.
- [88] T. Kiire, S. Nakadate, M. Shibuya, and T. Yatagai, "Three-dimensional displacement measurement for diffuse object using phase-shifting digital holography with polarization imaging camera," in *Applied Optics, Vol. 50, No. 34*, 2011.
- [89] Z. Chen, B. Salahieh, X. Wang, and R. Liang, "Multichannel micropolarizer camera as a three-dimensional imager for fast and high dynamic range objects," in *Imaging Systems and Applications, IMIA. 2*, 2015.
- [90] 4D Technology, PolarCam Polarization Camera, in Available at <http://www.4dtechnology.com/>.