Super-resolution imaging : Markov chain Monte Carlo and state-space approaches

Tian, Jing

2009

Tian, J. (2009). Super-resolution imaging : Markov chain Monte Carlo and state-space approaches. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/46816>

<https://doi.org/10.32657/10356/46816>

Nanyang Technological University

Downloaded on 26 Apr 2023 15:40:48 SGT

Super-resolution Imaging: Markov Chain Monte Carlo and State-space Approaches

Tian Jing

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2009

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original re-
search and has not been submitted for a higher degree to any other University or
Institution.

.....

Date

.....

Tian Jing

Acknowledgments

First and foremost, I wish to express my sincerest gratitude to my advisor, Prof. Kai-Kuang Ma. He has directed my research works throughout my Ph.D. study, and his advices have been most valuable for the completion of this thesis. At every stage of my work, his insights and his incisive questioning have been crucial in bringing my efforts to fruition. He always gave me strong support and sincere encouragement throughout the tough times of the research, particularly when my paper got rejection due to unfair judgements.

The first time I contacted with Prof. Ma was in 2001. I was in China and I sent an email to him, expressing my intentions of joining his group. Within one day, I got his prompt reply, telling me that he would consider my application. Since then, I have started my journey of conducting fundamental and challenging research under his supervision. After I joined NTU in 2003, I realized how lucky I was really to be a member of his research group. In the past five years, I have learnt not only specialized knowledge, but also the spirit to explore science, from Prof. Ma himself, as well as from all senior members in this academic family.

I also wish to take this opportunity to thank School of EEE, Nanyang Technological University for offering a research scholarship to me in my first three-year's studying, and providing research facilities for conducting research. It is also my great pleasure to thank my colleagues in Media Technology Lab, NTU. I have been greatly benefited from the collaboration and discussion with them.

Acknowledgments

Acknowledgments

Lastly but most important, I wish to dedicate this thesis to my family, whose love and support are always the greatest inspiration for me in my life. Thank my parents and parents-in-law, for their care, understanding and encouragement during my tough studying. Thank my wife, Zheng Jie, and my son, Tian Zi-Zheng, for the very special life they have been giving to me. Without them, nothing could be truly fun.

Contents

Acknowledgments	i
List of Figures	vii
List of Tables	xviii
Summary	xxi
List of Acronyms	xxiv
1 Introduction	1
1.1 Motivation	1
1.2 Contributions	3
1.3 Organization	7
2 Background and Related Works	9
2.1 Super-resolution Image Reconstruction	9
2.1.1 Observation Model	10
2.1.2 Frequency-domain-based SR Image Approach	12
2.1.3 Registration-interpolation-restoration-based SR Image Approach	15
2.1.4 Bayesian-based SR Image Approach	19
2.1.5 Learning-based SR Image Approach	21

CONTENTS	CONTENTS
2.2 Super-resolution Video Reconstruction	25
2.2.1 Observation Model	26
2.2.2 Sliding-window-based SR Video Approach	26
2.2.3 Simultaneous SR Video Approach	28
2.2.4 Learning-based SR Video Approach	29
3 Markov Chain Monte Carlo Super-resolution Image Reconstruc-	
tion	30
3.1 Introduction	30
3.2 Proposed MCMC Super-resolution Approach	34
3.2.1 Observation Model	34
3.2.2 Bayesian Inference	35
3.2.3 Proposed MCMC SR Approach	38
3.2.4 Experimental Results	41
3.3 Derivation of The Burn-in Parameter of the Proposed MCMC Super-	
resolution Process	55
3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter	58
3.4.1 Detection	58
3.4.2 Filtering	60
3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov	
Chain Monte Carlo Approach	66
3.5.1 Introduction	66
3.5.2 Proposed Edge-directed MCMC SR Approach	67
3.5.3 Experimental Results	71
3.6 Conclusions	74
Appendices	75
Appendix 3.A Derivation of (3.16)	75

CONTENTS

CONTENTS

Appendix 3.B Proof of (3.24)	77
Appendix 3.C Derivation of (3.26)	78
Appendix 3.D Derivation of (3.40)	83
4 A State-space Super-resolution Video Approach	86
4.1 Introduction	86
4.2 Observation Model	89
4.3 Proposed State-space Super-resolution Imaging Framework	90
4.3.1 State-space Representations	90
4.3.2 Conventional Two-equation-based Kalman Filtering for SR Video Reconstruction	94
4.3.3 Proposed Three-equation-based State-space SR Approach	95
4.3.4 Relationships to Other SR Approaches	99
4.3.5 Computational Complexity	104
4.4 Experimental results	106
4.4.1 Implementation	106
4.4.2 Simulation Results	110
4.5 Conclusions	133
Appendices	134
Appendix 4.A Derivation of (3.19)	134
Appendix 4.B Derivations of (3.34) and (3.36)	139
5 Conclusions and Future Works	142
5.1 Conclusions	142
5.2 Future Works	144
A Grid Computing Enabled Super-resolution Imaging	154
B Markov Chain Monte Carlo	165

CONTENTS	CONTENTS
C Markov Random Field	170
Author's Publications	172

List of Figures

1.1	A conceptual framework of super-resolution imaging [1–19].	2
1.2	Various versions of the <i>Text</i> image: (a) original image (50×50); (b) enlarged image (200×200) by applying pixel duplication on the image (a); (c) enlarged image (200×200) by applying super-resolution algorithm on the image (a).	3
1.3	An overview of the proposed stochastic SR image approach using the Markov chain Monte Carlo technique.	4
1.4	The proposed three-equation-based state-space representation for representing the relationships among the low-resolution frames and high-resolution frames.	6
2.1	The relationship between the original high-resolution image and the observed low-resolution images. The observed low-resolution images are the warped, blurred, down-sampled and noisy version of the original high-resolution image.	11
2.2	The registration-interpolation-restoration-based approach for super-resolution image reconstruction.	17
2.3	A conceptual framework of learning-based SR image approach.	22
2.4	The sliding-window-based super-resolution video scheme.	27
2.5	The simultaneous multiple high-resolution frames reconstruction scheme.	28

LIST OF FIGURES

LIST OF FIGURES

3.1 Various reconstructed results of *Boat* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation approach, followed by deblurring (PSNR = 25.80 dB); (e) Vandewalle *et al.*'s algorithm [20] (PSNR = 25.97 dB); (g) Pham *et al.*'s algorithm [21] (PSNR = 26.61 dB); (i) Farsiu *et al.*'s algorithm [22] (PSNR = 27.45 dB); (k) the proposed algorithm with the automatically-updated hyperparameter (PSNR = 28.02 dB); images (d), (f), (h), (j) and (l) are the absolute difference between image (a) and images (c), (e), (g), (i) and (k), respectively. Furthermore, the intensity values of these absolute difference images are scaled up three times for demonstration. 49

3.2 Various reconstructed results of *Text* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation approach, followed by deblurring (PSNR = 17.62 dB); (e) Vandewalle *et al.*'s algorithm [20] (PSNR = 18.32 dB); (g) Pham *et al.*'s algorithm [21] (PSNR = 17.74 dB); (i) Farsiu *et al.*'s algorithm [22] (PSNR = 19.29 dB); (k) the proposed algorithm with the automatically-updated hyperparameter (PSNR = 20.17 dB); images (d), (f), (h), (j) and (l) are the absolute difference between image (a) and images (c), (e), (g), (i) and (k), respectively. 50

LIST OF FIGURES

LIST OF FIGURES

3.3 Various reconstructed results of *Boat* image: (a) the original image; (b) the bi-cubic spline interpolation, followed by deblurring approach (PSNR = 25.80 dB); (c) the proposed method with the fixed-valued hyperparameter $\lambda = 0.1$ to demonstrate the *under*-smoothing case (PSNR = 24.65 dB); (d) the proposed method with the fixed-valued hyperparameter $\lambda = 300$ to demonstrate the *over*-smoothing case (PSNR = 24.85 dB); (e) the proposed method with the best hyperparameter $\lambda = 20$, which is obtained by exhaustively experimenting individual integer value over the interval $[1, 50]$ to demonstrate the *optimal* smoothing case (PSNR = 28.13 dB); and (f) the proposed algorithm with the *adaptive* hyperparameter (PSNR = 28.02 dB), which is fairly close to the optimal smoothing case as presented in (e). 51

3.4 Various reconstructed results of *Text* image: (a) the original image; (b) the bi-cubic spline interpolation, followed by deblurring approach (PSNR = 17.62 dB); (c) the proposed method with the fixed-valued hyperparameter $\lambda = 0.1$ to demonstrate the *under*-smoothing case (PSNR = 17.73 dB); (d) the proposed method with the fixed-valued hyperparameter $\lambda = 300$ to demonstrate the *over*-smoothing case (PSNR = 15.92 dB); (e) the proposed method with the best hyperparameter $\lambda = 8$, which is obtained by exhaustively experimenting individual integer value over the interval $[1, 50]$, to demonstrate the *optimal* smoothing case (PSNR = 20.36 dB); and (f) the proposed algorithm with the *adaptive* hyperparameter (PSNR = 20.17 dB), which is fairly close to the optimal smoothing case as presented in (e). 52

3.5 Various reconstructed results of *Boat* image of the proposed MCMC SR approach. 53

LIST OF FIGURES

LIST OF FIGURES

3.6	Various reconstructed results of <i>Text</i> image of the proposed MCMC SR approach.	54
3.7	A comparison of the burn-in parameter T computed via the proposed closed-form bound (3.27) versus the number of burn-in samples (iterations) obtained in simulations: (a) <i>Boat</i> image; (b) <i>Text</i> image.	59
3.8	Various reconstructed results of <i>Boat</i> image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation, followed by image deblurring approach (PSNR = 25.80 dB); (d) the proposed MCMC algorithm using $N = 3 \times T$ image samples (PSNR = 25.32 dB); (e) post-processing result (PSNR = 26.27 dB) by applying the bilateral filter [23] on the picture (d); (f) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 26.92 dB); (g) the proposed MCMC algorithm using $N = 10 \times T$ image samples (PSNR = 28.02 dB); (h) post-processing result (PSNR = 28.12 dB) by applying the bilateral filter [23] on the picture (g); and (i) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 28.21 dB).	64

LIST OF FIGURES

LIST OF FIGURES

3.9 Various reconstructed results of *Text* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation, followed by image deblurring approach (PSNR = 17.62 dB); (d) the proposed MCMC algorithm using $N = 3 \times T$ image samples (PSNR = 17.88 dB); (e) post-processing result (PSNR = 18.81 dB) by applying the bilateral filter [23] on the picture (d); (f) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 19.43 dB); (g) the proposed MCMC algorithm using $N = 10 \times T$ image samples (PSNR = 20.17 dB); (h) post-processing result (PSNR = 20.34 dB) by applying the bilateral filter [23] on the picture (g); and (i) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 20.40 dB). 65

3.10 The image sample, which is generated at each sample generation step of the proposed MCMC SR approach, is enhanced by adding a positive scaling factor β that is multiplied with a linear high-pass filtered version of the image sample. 71

3.11 Three test images used in the experiments: (a) 256×256 *Airplane*; (b) 256×256 *Boat*; and (c) 256×256 *Bridge*. 72

LIST OF FIGURES

LIST OF FIGURES

3.12 Various results of *Airplane* image (2×2 enlargement): (a) the original image; (b) the bi-cubic spline interpolation approach (PSNR = 25.69 dB); (c) Vandewalle *et al.*'s algorithm developed in [20] (PSNR = 27.49 dB); (d) the MCMC SR approach developed in [24] (PSNR = 28.81 dB); (e) the proposed edge-directed MCMC SR approach using the edge map computed from the observed low-resolution image ($\beta = 4$, PSNR = 29.09 dB); (f) the proposed edge-directed MCMC SR approach using the edge map computed from the ground truth high-resolution image ($\beta = 4$, PSNR = 29.53 dB); and images (g)-(l) are the part of the images of (a)-(f), respectively. 73

4.1 Two conventional super-resolution video reconstruction methodologies: (a) the sliding-window-based super-resolution video scheme; and (b) the simultaneous multiple high-resolution frames reconstruction scheme. 87

4.2 The relationship between the original high-resolution image $\mathbf{X}_n$ and the observed low-resolution image $\mathbf{y}_n$. The observed low-resolution image $\mathbf{y}_n$ is a warped, blurred and down-sampled version of the original high-resolution image $\mathbf{X}_n$, with the additive white Gaussian noise. 89

4.3 The relationship among the low-resolution frames $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ and the high-resolution frames $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ using: (a) the conventional two-equation-based Kalman filtering; (b) the proposed three-equation-based state-space filtering. 92

LIST OF FIGURES

LIST OF FIGURES

4.4 Four versions of the 84-th frame of the *Foreman* image sequence: (a) the original image X (144×176); (b) a blurred image X_B , after convolving X with a 3×3 uniform averaging filter; (c) the down-sampled image X_D , after down-sampling X_B by a factor of two in both the horizontal and the vertical directions; and (d) the simulated low-resolution image, after adding the zero-mean white Gaussian noise to X_D . (For demonstration, images (c) and (d) are scaled up to the same size as that of X through pixel duplication.) 121

4.5 Four versions of the 94-th frame of the *Bus* image sequence: (a) the original image X (144×176); (b) a blurred image X_B , after convolving X with a 3×3 uniform averaging filter; (c) the down-sampled image X_D , after down-sampling X_B by a factor of two in both the horizontal and the vertical directions; and (d) the simulated low-resolution image, after adding the zero-mean white Gaussian noise to X_D . (For demonstration, images (c) and (d) are scaled up to the same size as that of X through pixel duplication.) 122

4.6 Various 2×2 enlarged results of the 84-th frame of the *Foreman* image sequence using the artificially-degraded image sequence: (a) the bi-cubic spline interpolation approach (PSNR = 27.49 dB); (b) Vandewalle *et al.*'s algorithm [20] (PSNR = 28.19 dB); (c) the conventional two-equation-based Kalman filtering (PSNR = 29.20 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 29.76 dB). 123

LIST OF FIGURES

LIST OF FIGURES

4.7 Various 2×2 enlarged results of the 94-th frame of the *Bus* image sequence using the artificially-degraded image sequence: (a) the bi-cubic spline interpolation approach (PSNR = 21.64 dB); (b) Vandewalle *et al.*'s algorithm [20] (PSNR = 21.38 dB); (c) the conventional two-equation-based Kalman filtering (PSNR = 23.86 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 24.39 dB). 124

4.8 Various 2×2 enlarged results (in CIF) of the 84-th frame of the *Foreman* image sequence using the original image sequence (QCIF frame size): (a) the bi-cubic spline interpolation approach; (b) Vandewalle *et al.*'s algorithm [20]; (c) the conventional two-equation-based Kalman filtering; and (d) the proposed three-equation-based state-space filtering. 125

4.9 Various 2×2 enlarged results (in CIF) of the 94-th frame of the *Bus* image sequence using the original image sequence (QCIF frame size): (a) the bi-cubic spline interpolation approach; (b) Vandewalle *et al.*'s algorithm [20]; (c) the conventional two-equation-based Kalman filtering; and (d) the proposed three-equation-based state-space filtering. 126

4.10 Various 2×2 enlarged results of the 4-th frame of the *License-Plate* image sequence: (a) original image (80×80); (b) the conventional two-equation-based Kalman filtering; (c) the proposed three-equation-based state-space filtering; image (d)-(f) are the zoomed parts of images (a)-(c) using pixel duplication, respectively. (For demonstration, images (a) and (d) are scaled up through pixel duplication.) 127

LIST OF FIGURES

LIST OF FIGURES

4.11 Various 2×2 enlarged results of the 6-th frame of the *EEE-Board* image sequence: (a) original image (100×160); (b) the conventional two-equation-based Kalman filtering; (c) the proposed three-equation-based state-space filtering; image (d)-(f) are the zoomed parts of images (a)-(c) using pixel duplication, respectively. (For demonstration, images (a) and (d) are scaled up through pixel duplication.) 128

4.12 Various deblurred results of the 84-th frame of the *Foreman* image sequence: (a) the original image; (b) the simulated blurring image; (c) the conventional two-equation-based Kalman filtering (PSNR = 30.63 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 31.24 dB). 129

4.13 Various deblurred results of the 94-th frame of the *Bus* image sequence: (a) the original image; (b) the simulated blurring image; (c) the conventional two-equation-based Kalman filtering (PSNR = 25.32 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 26.59 dB). 130

4.14 Various deblurred results of the 7-th frame of the *License-Plate* image sequence: (a) original image; (b) the conventional two-equation-based Kalman filtering; and (c) the proposed three-equation-based state-space filtering. 131

4.15 Various deblurred results of the 3-rd frame of the *EEE-Board* image sequence: (a) original image; (b) the conventional two-equation-based Kalman filtering; and (c) the proposed three-equation-based state-space filtering. 131

LIST OF FIGURES

LIST OF FIGURES

4.16	The relationship between the states $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ and the observations $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$: (a) at the discrete-time step n ; (b) at the discrete-time step $n - 1$	134
5.1	A conceptual framework of simultaneous resolution enhancement in both the spatial domain and the temporal domain.	150
A.1	An overview of image samples generation procedure.	156
A.2	A general framework of the grid computing structure.	158
A.3	(a) A single MCMC SR task that needs to sequentially exploit all the hyperparameter values permissible in the hyperparameter set chosen. (b) A set of independent MCMC SR tasks that are dispatched in the grid computing environment for parallel computing, in which each task runs under one fixed, but different, hyperparameter value. . . .	159
A.4	An overview of the proposed grid-based MCMC SR framework. . .	161
A.5	Four test images used in the experiments: (a) Airplane; (b) Boat; (c) Bridge; and (d) Text.	162
A.6	The work-flow of the proposed grid-based MCMC SR framework. . .	163
A.7	The comparison of computational time.	164
B.1	The comparison between the distribution defined in (B.1) and the histogram of the samples generated by the M-H method according to this distribution. The histogram of the samples become more consistent to their distribution when a larger number of samples are generated: (a) the distribution defined in (B.1); (b) the histogram of 50 samples generated by the M-H method; (c) the histogram of 500 samples generated by the M-H method; and (d) the histogram of 5000 samples generated by the M-H method.	167

LIST OF FIGURES

LIST OF FIGURES

B.2 The samples generated by the Gibbs sampler according to the distribution defined in (B.2): (a) 10 samples starting from the position (10, 10); (b) 200 samples starting from the position (10, 10); (c) 1000 samples starting from the position (10, 10); and (d) 1000 samples starting from the position (−10, −10). 169

List of Tables

3.1	The PSNR performance comparison.	45
3.2	The PSNR performance comparison of 2×2 SR computation of the proposed MCMC SR approach using four low-resolution observations (including one severely degraded low-resolution observation).	45
3.3	The run time comparison.	46
3.4	The run time of the burn-in period comparison of the proposed MCMC SR approach using two different starting image samples.	46
3.5	The PSNR performance comparison.	62
3.6	The PSNR performance comparison of the proposed outlier-sensitive bilateral filtering with the incorporation of the mean filter defined in (3.33) and that with the incorporation of the median filter defined in (3.32).	63
3.7	The PSNR performance comparison. Proposed method I represents the proposed edge-directed MCMC SR using the edge map computed from the low-resolution observations, while the proposed method II represents the proposed edge-directed MCMC SR using the edge map computed from the original high-resolution image.	74
4.1	The absolute difference d_n (see (4.67)) between $p(\mathbf{V}_n, \mathbf{U}_n)$ and $p(\mathbf{V}_n) \times p(\mathbf{U}_n)$	116

LIST OF TABLES

LIST OF TABLES

4.2 The average PSNR performance comparison on the SR computation using the artificially-degraded image sequence. 118

4.3 The average SSIM [25] performance comparison on the SR computation using the artificially-degraded image sequence. 118

4.4 The average variance [26] performance comparison on the SR computation using the original image sequence. 118

4.5 The average NRQA [27] performance comparison on the SR computation using the original image sequence. 119

4.6 The average variance [26] performance comparison on the SR computation using the real-life image sequence captured by a built-in camera of a handphone. 119

4.7 The average NRQA [27] performance comparison on the SR computation using the real-life image sequence captured by a built-in camera of a handphone. 119

4.8 The average PSNR performance comparison on the deblurring computation using the artificially degraded image sequence. 120

4.9 The average SSIM [25] performance comparison on the deblurring computation using the artificially degraded image sequence. 120

4.10 The average variance [26] performance comparison on the deblurring computation using the real-life video sequence captured by a built-in camera of a handphone. 120

4.11 The average NRQA [27] performance comparison on the deblurring computation using the real-life video sequence captured by a built-in camera of a handphone. 120

4.12 The run time (averaged over all 100 frames) comparison on the SR computation using the artificially-degraded image sequence. 132

LIST OF TABLES

LIST OF TABLES

4.13 The run time (averaged over all 100 frames) comparison on the SR
computation on the SR computation using the original image sequence.132

4.14 The run time (averaged over all 9 frames) comparison on the SR
computation using the real-life image sequence captured by a built-
in camera of a handphone. 132

Summary

The key objective of *super-resolution* (SR) imaging is to reconstruct a single higher-resolution image based on a set of images, acquired from the same scene and denoted as ‘low-resolution’ images, to overcome the limitation and/or ill conditions of the image acquisition process for facilitating better visualization and content recognition. This thesis investigates two fundamental approaches—Bayesian inference via Markov chain Monte Carlo (stochastic approach) and state-space estimation (deterministic approach).

For the development of stochastic SR imaging, it consists of two milestone stages. First, a Bayesian inference framework is first mathematically formulated, which is based on both the observed low-resolution images and the prior model of unknown high-resolution image. Second, the *Markov chain Monte Carlo* (MCMC) technique is novelly exploited with four fundamental issues insightfully observed, as follows. First, since the prior image model’s hyperparameter value controls the degree of regularization and intimately affects the quality of the reconstructed high-resolution image, how to determine an optimal hyperparameter value without calling for the exhaustive search plays a very crucial role in stochastic SR imaging. For that, a novel approach is proposed in this thesis by allowing the value of the hyperparameter being simultaneously updated via each generation of image sample throughout the entire MCMC process. Second, the samples generated during the so-called *burn-in* period (measured in terms of the number of image samples initially generated) of

the MCMC process are considered unreliable and should be discarded. To dynamically determine how long the burn-in period for each set of low-resolution input images, a closed-form solution is mathematically derived. Third, image artifacts could be incurred in the reconstructed high-resolution image, if the number of image samples (counting after the burn-in period) is generated for the MCMC process is insufficient. For that, a new bilateral filtering is proposed as a ‘complementary’ post-processing scheme, meaning that it contributes appreciable performance improvement only when the number of image samples is not enough. Fourth, an edge-directed MCMC SR approach is proposed to enhance the edge information of the high-resolution image by exploiting an edge map, which is generated by performing interpolation on the edge map obtained from one of the low-resolution input images.

For the state-space estimation approach, the conventional *two*-equation-based Kalman filter is studied and observed that it only considers the information from the previously-reconstructed high-resolution frame and the currently-observed low-resolution frame for producing each higher-resolution frame. It has been further observed that the information inherited in the previously-observed low-resolution frame could be beneficial on the reconstruction of the SR video. For that, an extra state-space equation based on the useful information extracted from adjacent low-resolution frames is incorporated into the framework of the conventional two-equation-based Kalman filtering to establish a *three*-equation-based state-space approach as a more generalized state-space framework. The closed-form solution is mathematically derived, and extensive simulations using both artificially-degraded and real-life image sequences are conducted to demonstrate its superior performance. Furthermore, a unified theoretical analysis is provided to analyze the

relationship between the proposed framework and two existing SR approaches—the sliding-window-based Bayesian estimation approach and the conventional two-equation-based Kalman filtering, respectively.

List of Acronyms

ABBREVIATIONS	FULL EXPRESSIONS
ANN	Artificial neural network
CCD	Charge coupled device
CFA	Color filter array
CIF	Common intermediate format
DCT	Discrete cosine transform
DFD	Displaced frame difference
DFT	Discrete Fourier transform
DWT	Discrete wavelet transform
EM	Expectation-maximization
IBP	Iterated back-projection
MAP	Maximum a posterior
MCMC	Markov chain Monte Carlo
M-H	Metropolis-Hastings
ML	Maximum likelihood
MRF	Markov random field
POCS	Projection onto convex sets
PSF	Point spread function

List of ACRONYMS

List of ACRONYMS

PSNR	Peak signal-to-noise ratio
QCIF	Quarter common intermediate format
SNR	Signal-to-noise ratio
SR	Super-resolution
SVM	Support vector machine

Chapter 1

Introduction

1.1 Motivation

Super-resolution (SR) image processing [1–19] aims to overcome the limitation of the image acquisition device/system and/or possibly ill acquisition conditions to produce a higher-resolution image based on a set of images, acquired from the same scene as illustrated in Figure 1.1. Likewise, for SR video, higher-resolution frames are produced from their lower-resolution counterparts. With rapid development and deployment of image processing for visual communications, there is a strong demand for providing the viewer with high-resolution imaging not only for better visualization (fidelity issue) but also for additional information details (recognition issue). For examples, a high-resolution image is beneficial to achieve a better classification of regions in a multi-spectral image or to assist radiologist for making correct diagnosis. In video conferencing, the resolution of the video usually needs to be improved for effective communication. And, in video surveillance systems, high-resolution images are usually required to recognize criminals.

Figure 1.1. A conceptual framework of super-resolution imaging [1–19].

The most direct way to achieve higher-resolution images is to improve the image acquisition device/system by reducing the pixel size in its sensor (e.g., *charge-coupled device* (CCD)) or increasing its chip size. However, there is a limitation in reducing the sensor’s pixel size in the sensor manufacturing technologies. More importantly, the captured image quality decreases if the sensor’s pixel size becomes small, due to the noise remains roughly the same with the sensor’s pixel size reduction, while the signal power decreases proportional to the sensor’s pixel size reduction. On the other hand, there is a high commercial cost to increase the chip size. Owing to the above-mentioned inherent limitations in the sensor manufacturing technologies, SR imaging becomes a promising alternative.

The SR image approaches reconstruct a single high-resolution image from a set of given low-resolution images; likewise, the SR video approaches reconstruct an image sequence with a high-resolution from a group of adjacent low-resolution uncompressed image frames or compressed image frames. Based on multiple low-resolution images, each of which provides a unique information on the scene, it is possible to exploit the information provided by all the captured low-resolution images to produce an image with a higher resolution than that of each captured low-resolution image. The above-mentioned multiple low-resolution images can either have relative geometric transformations among each other, or have different blurs [28–31] or zoomed factors [32–35] that are incurred during image acquisition. Furthermore,

1.2 Contributions

Figure 1.2. Various versions of the *Text* image: (a) original image (50×50); (b) enlarged image (200×200) by applying pixel duplication on the image (a); (c) enlarged image (200×200) by applying super-resolution algorithm on the image (a).

rather than using a single image acquisition device (as illustrated in Figure 1.1), a few attempts have been done on multiple image acquisition devices, such as a CCD image sensor array consisting of a rectangular array of sensing elements, each of which is shifted relative to each other by user-controlled values [36–41].

To understand SR imaging, there are several fundamental concepts need to be clarified. Note that an image’s *resolution* is fundamentally different from its *size*. In our context, the objective of SR imaging is to produce an image with a clearer content from its low-resolution counterpart (e.g., produce Figure 1.2 (c) from Figure 1.2 (a)), rather than achieving larger size only (e.g., produce Figure 1.2 (b) by applying pixel duplication on Figure 1.2 (a)). Furthermore, in our context, the term *resolution* of SR refers to the *spatial* resolution of the image, not the *temporal* resolution of the image sequence, which refers to the number of frames captured per second and is also commonly known as the *frame rate*.

1.2 Contributions

The major contributions of this thesis are on the fundamental investigations and development of two new SR approaches—Bayesian inference via *Markov chain Monte Carlo* (MCMC) and state-space estimation.

Figure 1.3. An overview of the proposed stochastic SR image approach using the Markov chain Monte Carlo technique.

For the development of stochastic SR imaging, it consists of two milestone stages. First, a Bayesian inference framework is mathematically formulated based on both the observed low-resolution images and the prior model of unknown high-resolution image. Second, the MCMC technique is exploited to produce a high-resolution image through stochastic simulations, as illustrated in Figure 1.3. The proposed image samples generation process starts from a randomly-generated image sample (in our case, each image sample is a two-dimensional image). Then a fairly large number of image samples are generated according to the joint posterior distribution of the unknown high-resolution image and the prior image model's hyperparameter given the observed low-resolution images. Finally, those initially-generated image samples (within the burn-in period) are considered as unreliable and should be discarded, then the remaining reliable image samples are averaged to obtain the

1.2 Contributions

final high-resolution image result. Furthermore, there are several fundamental issues inherent with the proposed MCMC SR approach required to be addressed as follows.

The first issue is that the prior image model's hyperparameter value controls the degree of regularization and intimately affects the quality of the reconstructed high-resolution image. To determine the optimal hyperparameter value without calling for the exhaustive search, a novel approach by simultaneously updating the hyperparameter value throughout the proposed MCMC process. More specifically, the *Metropolis-Hastings* (M-H) algorithm [42] is exploited to simultaneously update the hyperparameter value, along with the updating of the image sample using the *Gibbs sampler* [43], at each sample-generation step throughout the proposed MCMC process. The second issue is that the samples generated during the *burn-in* period of the MCMC SR process are considered unreliable and should be discarded. Determination of how many initially-generated image samples are discarded is equivalent to the determination of the length of the burn-in period. For that, a closed-form bound is mathematically derived. The third critical issue is that image artifacts could be inevitably yielded in the reconstructed high-resolution image, if the number of image samples generated for the reconstruction process is insufficient. For that, a new bilateral filtering is proposed to further enhance the reconstructed high-resolution image. Fourth, an edge-directed MCMC SR approach is proposed to enhance the edge information of the reconstructed high-resolution image by exploiting an edge map, which is generated by performing interpolation on the edge map obtained from one of the low-resolution input images.

For the state-space estimation approach, a three-equation-based state-space model is proposed by incorporating an extra observation equation based on the

Figure 1.4. The proposed three-equation-based state-space representation for representing the relationships among the low-resolution frames and high-resolution frames.

useful information extracted from adjacent low-resolution frames into the framework of the conventional two-equation-based Kalman filtering, as depicted in Figure 1.4. The proposed approach exploits the information from the previously-reconstructed high-resolution frame, the currently-observed low-resolution frame as well as the previously-observed low-resolution frame for sequentially producing the high-resolution frames. This is in contrast to the conventional two-equation-based Kalman filtering, in which only the previously-reconstructed high-resolution frame and the currently-observed low-resolution frame are exploited. Note that the proposed equation is *not* obtained from any combination of two equations of conventional Kalman filtering; that is, it is not redundant. Therefore, the proposed extra equation can exploit new and useful information from the input data, which are not available in any one of two equations of conventional Kalman filtering, or both of them two. Furthermore, a unified theoretical analysis is provided to analyze

1.3 Organization

the fundamental relationship between the proposed framework and two existing SR approaches—the sliding-window-based Bayesian estimation approach and the conventional two-equation-based Kalman filtering, respectively.

1.3 Organization

The organization of this thesis is outlined as follows.

In Chapter 2, essential background knowledge on the image acquisition process and the mathematical formulation of the SR problem are provided first to establish the foundation for the development of all algorithms presented in the thesis. Next, a comprehensive review of existing SR imaging techniques is provided in two categories: the SR image reconstruction and the SR video reconstruction.

In Chapter 3, a stochastic SR image approach is proposed by exploiting Bayesian inference and the MCMC technique. Furthermore, the bound for the length of the burn-in period of the proposed MCMC SR process is mathematically derived. Moreover, a variation-sensitive bilateral filtering is proposed as a complementary post-processing stage for addressing the potential image artifacts issue. In addition, an edge-directed MCMC SR approach is proposed to further enhance the edge information in the reconstructed high-resolution image.

In Chapter 4, a new state-space SR video approach is proposed by incorporating an extra observation equation into the framework of the conventional two-equation-based Kalman filtering to establish a *three*-equation-based state-space approach as a more generalized framework. Its closed-form solution is mathematically derived, followed by conducting extensive experiments using both artificially-degraded and real-life image sequences. Furthermore, a unified theoretical analysis is provided to analyze the relationship between the proposed framework and two existing super-resolution approaches, the sliding-window-based Bayesian estimation approach and

1.3 Organization

the conventional two-equation-based Kalman filtering, respectively.

Finally, Chapter 5 concludes this thesis and discusses future research directions.

Chapter 2

Background and Related Works

The SR imaging research has grown very rapidly, after it was firstly addressed by Tsai and Huang [44] in 1984, due to its potential in overcoming the inherent limitations of existing image acquisition hardwares. Various developed SR research works have been reported in a number of books [1–3], journal special issues [4–9], and survey papers [10–19]. In view of this, the purpose of this chapter is to provide a comprehensive survey for the literature.

2.1 Super-resolution Image Reconstruction

The objective of SR image reconstruction is to fuse a set of images captured on the same scene to produce an image with a higher resolution. The essential issue of SR computation is on how to recover the information that have been degraded or lost during the phase of image acquisition. In general, the SR image techniques can be classified into two classes: the *reconstruction*-based approach and the *learning*-based approach. The former approach gets the high-resolution image result from low-resolution input images only, while the latter approach achieves the same

2.1 Super-resolution Image Reconstruction

objective by further exploit the information provided by an image database. Furthermore, the reconstruction-based approach can be further classified into three categories: i) *frequency-domain*-based approach, ii) *interpolation-restoration*-based approach, and iii) *Bayesian*-based approach. In the following, an observation model of the SR image reconstruction will be presented, followed by a discussion on the existing SR image approaches from the above-mentioned aspects.

2.1.1 Observation Model

The first step to address the SR problem is to establish observation models that describe the image acquisition process of the LR images, since the aim of the SR computation is to overcome the limitation of the image acquisition system. Given an original image, it is considered as the high-resolution ground truth, which is to be compared with the reconstructed image from a set of low-resolution images for performance evaluation. The image acquisition process is usually modeled by the following four stages, as illustrated in Figure 2.1: i) geometric transformation, ii) blurring, iii) down-sampling and iv) noise adding. The geometric transformation includes translation, rotation and scaling. Various blurrings (such as motion blur and out-of-focus blur) are usually modeled by convolving the image with a low-pass filter, which is modeled by a *point spread function* (PSF). The noise introduced in the image acquisition process is usually assumed to be additive, zero-mean white Gaussian noise.

To summarize, the above-mentioned image acquisition process is that the original high-resolution image (with a size of $M_1 \times M_2$) is warped, blurred, down-sampled and added with white Gaussian noise to generate the observed low-resolution image (with a size of $L_1 \times L_2$, where $L_1 = M_1/q_1$ and $L_2 = M_2/q_2$). Here, the downsampling ratio is $q_1 \times q_2$. Hence, the observation model for such image acquisition

2.1 Super-resolution Image Reconstruction

Figure 2.1. The relationship between the original high-resolution image and the observed low-resolution images. The observed low-resolution images are the warped, blurred, down-sampled and noisy version of the original high-resolution image.

process can be mathematically formulated as

$$\mathbf{y}^{(k)} = \mathbf{D}^{(k)}\mathbf{P}^{(k)}\mathbf{W}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}, \quad (2.1)$$

$$= \mathbf{H}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}, \quad (2.2)$$

where $\mathbf{y}^{(k)}$ and $\mathbf{X}$ represent the k -th $L_1 \times L_2$ low-resolution image and the $M_1 \times M_2$ high-resolution image, respectively. Both of these two images are represented in the lexicographic-ordered vector form, with a size of $L_1 L_2 \times 1$ and $M_1 M_2 \times 1$, respectively. With the above matrix notation, each $L_1 \times L_2$ image can be transformed into a $L_1 L_2 \times 1$ column vector, obtained by ordering the image by rows (lexicographic ordering). $\mathbf{D}^{(k)}$ is the decimation matrix with a size of $L_1 L_2 \times M_1 M_2$, $\mathbf{P}^{(k)}$ is the blurring matrix of size $M_1 M_2 \times M_1 M_2$, and $\mathbf{W}^{(k)}$ is the warping matrix of size $M_1 M_2 \times M_1 M_2$. $\mathbf{H}^{(k)} = \mathbf{D}^{(k)}\mathbf{P}^{(k)}\mathbf{W}^{(k)}$ is an $L_1 L_2 \times M_1 M_2$ matrix, representing the above-mentioned warping (i.e., the matrix $\mathbf{W}^{(k)}$ in (2.1), convolving (i.e., the matrix $\mathbf{P}^{(k)}$ in (2.1) and downsampling process (i.e., the matrix $\mathbf{D}^{(k)}$ in (2.1). Additionally, $\mathbf{V}^{(k)}$ is a $L_1 L_2 \times 1$ vector, representing the white Gaussian noise (with a zero mean and a variance σ_k^2) encountered during the image acquisition process. Furthermore, $\mathbf{V}^{(k)}$ is assumed to be independent of $\mathbf{X}$.

2.1 Super-resolution Image Reconstruction

Over a period of time, one can capture a set of (say, ρ) observations $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(\rho)}\}$. With such establishment, the goal of the SR image reconstruction is to produce a single high-resolution image $\mathbf{X}$ based on a few low-resolution observations $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(\rho)}\}$.

Note that, in the literature SR research works [45], the other observation model is also used. It has different image acquisition operations order with that of (2.1). More specifically, the original image is blurred first, then warped, decimated and added with white Gaussian noise to yield low-resolution images. That is, (2.1) needs to be rewritten as $\mathbf{y}^{(k)} = \mathbf{D}^{(k)}\mathbf{W}^{(k)}\mathbf{P}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}$ accordingly. However, this model is unable to capture large rotations and scale variations incurred in the aerial imagery problem [46]. The determination of mathematical model, which is used for addressing SR problems, should coincide with the imaging physics (i.e., the physical process to capture low-resolution images from the original high-resolution ones) [2].

2.1.2 Frequency-domain-based SR Image Approach

DFT-domain-based SR Image Approach

The first frequency-domain SR algorithm was developed by Tsai and Huang [44], where they considered the SR computation for the noise-free low-resolution images. They proposed to first transform the low-resolution image data into the *discrete Fourier transform* (DFT) domain and combined them according to the relationship between the aliased DFT coefficients of the observed low-resolution images and that of the unknown high-resolution image. This combined data is then transformed back to the spatial domain where the new image could have a higher resolution than that of the input images. In view of only noise-free image data is considered in [44], Kim *et al.* addressed the SR computation for the noisy and blurred low-resolution image data by developing a recursive algorithm [47, 48]. Rhee and Kang

2.1 Super-resolution Image Reconstruction

[49] further reduced the computational load of the method [47] by exploiting the *Discrete cosine transform* (DCT), rather than exploiting the DFT used in [47]. More recent frequency-domain-based SR work has been carried out by Woods *et al.* in [50], where they presented an iterative *expectation maximization* (EM) algorithm for simultaneously performing the registration, blind deconvolution and interpolation operations.

The frequency-domain-based SR approaches has a number of advantages. Firstly, it is an intuitive way to enhance the details (usually the high-frequency information) of the images by extrapolating the high-frequency information presented in the low-resolution images. Secondly, these frequency-domain-based SR approaches have low computational complexity. However, the frequency-domain-based SR methods are insufficient to handle the real-world applications, since they require that there only exist global displacement between the observed images and the linear space-invariant blur during the image acquisition process.

Wavelet-domain-based SR Image Approach

Motivated by that the challenge of the SR computation is to recover the detailed information (usually the high-frequency information) that is lost or degraded during the image acquisition process, many researchers have begun to investigate the use of the wavelet transform for addressing the SR problem, since the wavelet transform provides a powerful and efficient multi-scale representation of the image for recovering the high-frequency information. Commonly, the existing wavelet SR methods can be classified into two categories: the *interpolation*-based approach and the *reconstruction*-based approach.

2.1 Super-resolution Image Reconstruction

- The *interpolation*-based wavelet-domain SR approaches typically treat the observed low-resolution images as the low-pass filtered subbands of the unknown wavelet-transformed high-resolution image. The aim is to estimate the finer scale subband coefficients, followed by applying the inverse wavelet transform, to produce the high-resolution image. To be more specific, take the 2×2 SR computation as an example. The low-resolution images are viewed as the representation of wavelet coefficients after several level (say, 3 levels) decompositions. Then the high-resolution image can be produced by estimating the 4-th scale wavelet coefficients, followed by applying the inverse wavelet decomposition. Several methods have been developed for estimating the wavelet coefficients of the high-pass filtered subbands based on the coefficients in the low-pass filtered subbands, such as the linear interpolation of the coefficients in the low-pass filtered subbands [51].
- The *reconstruction*-based wavelet-domain SR approaches address the SR computation as image reconstruction in the wavelet domain. Ei-Khamy *et al.* [52, 53] proposed to firstly register the multiple low-resolution images in the wavelet domain, then fuse the registered low-resolution wavelet coefficients to obtain a single image, followed by performing interpolation to get a high-resolution image. Ji and Fermuller [54] proposed a robust wavelet SR approach to handle the error incurred in both the registration computation and the blurring identification computation. Chappalli and Bose incorporated a denoising stage into the conventional wavelet-domain SR framework to develop a simultaneous denoising and SR reconstruction approach in [55]. Nguyen and Milanfar [56] proposed to apply the wavelet transform to reduce the computational cost by utilizing the sparse structure of the wavelet coefficients. Chan

2.1 Super-resolution Image Reconstruction

et al. [57] analyzed the SR problem from the wavelet point of view to investigate the case when some of the observed low-resolution images are missing, possibly due to the sensor failure. Shankar and Neifeld [58] proposed to impose the sparse constraints for the desired wavelet coefficients. On the other hand, rather than exploiting the traditional decimated wavelet transforms, the lifting-based wavelet scheme [55, 59, 60] and the biorthogonal wavelet [61, 62] have also been investigated for addressing the SR problem.

2.1.3 Registration-interpolation-restoration-based SR Image Approach

The basic idea of the *registration-interpolation-restoration*-based SR approach is to construct a high-resolution image by projecting all the observed low-resolution images to the reference image, then fuse together all the information available from each image, due to the fact that each low-resolution image provides a amount of additional information about the scene, and finally deblurs the image. Note that the single image interpolation algorithm can not handle the SR problem well, since they can not produce the high-frequency components lost during the image acquisition process. The quality of the interpolated image generated by any of the single input image interpolation algorithms is inherently limited by the amount of information available in the single input image only.

The registration-interpolation-restoration-based SR approach usually consists the following three stages, as depicted in Figure 2.2: i) the registration stage for aligning the input low-resolution images, ii) the interpolation stage for producing a high-resolution image, and iii) the deblurring stage for enhancing the reconstructed high-resolution image in the step ii). It can be further classified into two classes: i) *non-iterative* approach and ii) *iterative* approach.

Non-iterative Approach

The simplest interpolation algorithm is the so-called nearest neighbor algorithm, where each unknown pixel is given the intensity value of the pixel closest to it. But this method tends to produce images with a blocky appearance. More satisfactory results can be obtained with bi-linear interpolation or by using kernel cubic convolution techniques [63].

Ur and Gross [64] performed a nonuniform interpolation of a set of spatially-shifted low-resolution images by utilizing the generalized multichannel sampling theorem. The key advantage of this approach is that it has low computational load; thus, it is quite suitable for real-time applications. However, the optimality of the entire reconstruction process is not guaranteed, since the interpolation errors are not taken into account. Prasad [65] extended the method [64] into the spatial-frequency domain, and incorporate the regularization term to handle the SR computation with arbitrary numbers of LR images.

Chiang and Boulton [66] proposed a SR approach via image warping, which has the following four stages: i) choose one of the images as the reference image, and compute the motion field between all the images and the reference image; ii) scale up the reference image, and then warp all the images to the reference image based on the motion field computed in the previous step and the chosen scale; iii) obtain a SR image by fusing all the images together; and iv) deblur the resulting SR image.

Lin *et al.* [67] exploited a fuzzy decision system built from the characteristics of the *human visual system* (HVS) to classify pixels of the input image into human perception non-sensitive class and sensitive class. Then, the bilinear interpolation is used to interpolate the nonsensitive regions and a neural network is exploited to interpolate the sensitive regions along edge directions.

2.1 Super-resolution Image Reconstruction

Figure 2.2. The registration-interpolation-restoration-based approach for super-resolution image reconstruction.

Bose and Ahuja [68] exploited the *moving least square* (MLS) method to estimate the intensity value at each pixel position of the high-resolution image via a polynomial approximant using the pixels in a defined neighborhood of the pixel position under consideration. Furthermore, the coefficients and the order of the polynomial approximation are adaptively changed for each pixel position.

Lertrattanapanich and Bose [69] developed a Delaunay triangulation based approach, which was further improved by Li [70] using a *partial differential equation* (PDE) based nonlinear deblurring to accommodate the blindness and variation of sensor PSFs.

Iterative Approach

Irani and Peleg [71] proposed an *iterated back-projection* (IBP) algorithm, where the high-resolution image is estimated by iteratively back projecting the difference between the observed low-resolution images and the simulated low-resolution images, which are simulated by applying blurring on the estimated high-resolution image. However, this method has no unique solution due to the ill-posed nature of the SR problem. Jiang *et al.* [72] further improved the method [71] by considering the spatially-variant PSF using the *elliptical weighted area* (EWA) filter. Chen *et al.* [73] improved [71] by considering the initial interpolation, robust image registration, automatic image selection and image enhancement post-processing issues. Lu and Minoru developed the IBP method using a pyramid structure in [74–76]. Laplacian pyramid [77] is also exploited to predict the high-resolution image based on the Gaussian/Laplacian pyramid structure.

A *Projection onto Convex Sets* (POCS) was proposed by Patti and Tekalp [78] to develop a set-theoretic algorithm to produce the high-resolution image that is consistent with the information arising from the observed low-resolution images and the prior knowledge about the unknown high-resolution image. These information are associated with the constraint sets in the solution space; the intersection of these sets represents the space of acceptable solutions. By projecting an initial estimate of the unknown high-resolution image onto these constraint sets iteratively, a fairly good solution can be obtained. This kind of method is easy to be implemented, however it does not guarantee uniqueness of the solution. Furthermore, the computational cost of this algorithm is very high.

2.1 Super-resolution Image Reconstruction

2.1.4 Bayesian-based SR Image Approach

Motivated by the fact that the SR computation is, in essence, an ill-posed inverse problem [79], numerous Bayesian inference-based SR algorithms have been developed for addressing this issue. The basic idea of these Bayesian-based SR approaches is to use the regularization strategy to resolve the ill-posed nature of the SR computation by incorporating the prior knowledge of the unknown high-resolution image. From the Bayesian point of view, the information that can be extracted from observations (i.e., the low-resolution images) about the unknown signal (i.e., the high-resolution image) is contained in the probability distribution of the unknown. Then the unknown high-resolution image can be estimated via some statistics of a probability distribution of the unknown high-resolution image, which is usually established by applying Bayesian inference to exploit the information provided by both the observed low-resolution images and the prior knowledge of the unknown high-resolution image.

Two most popular Bayesian-based SR approaches are: *maximum likelihood* (ML) estimation approach and *maximum a posterior* (MAP) estimation approach.

i) **ML estimation approach:** An ML estimation approach is proposed by Tom and Katsaggelos in [80], where the aim is to find the ML estimation of the high-resolution image (denoted as $\hat{\mathbf{X}}^{ML}$) by

$$\hat{\mathbf{X}}^{ML} = \arg \max_{\mathbf{X}} P(\mathbf{Y}|\mathbf{X}). \quad (2.3)$$

The key of the above ML estimation approach is to establish the conditional pdf $p(\mathbf{Y}|\mathbf{X})$ in (2.3). Firstly, since the low-resolution images are obtained independently from the original (high-resolution) image, the conditional pdf $p(\mathbf{Y}|\mathbf{X})$ can be expressed as

$$p(\mathbf{Y}|\mathbf{X}) = \prod_{k=1}^{\rho} p(\mathbf{y}^{(k)}|\mathbf{X}). \quad (2.4)$$

2.1 Super-resolution Image Reconstruction

Furthermore, according to (2.2), the conditional pdf $p(\mathbf{y}^{(k)}|\mathbf{X})$ can be expressed as

$$\begin{aligned} p(\mathbf{y}^{(k)}|\mathbf{X}) &= p((\mathbf{H}^{(k)}\mathbf{X} + \mathbf{V}^{(k)})|\mathbf{X}) = p(\mathbf{V}^{(k)}|\mathbf{X}) \\ &= p(\mathbf{V}^{(k)}) \propto \exp\left(-\frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right). \end{aligned} \quad (2.5)$$

Substituting (2.5) into (2.4), the conditional pdf $p(\mathbf{Y}|\mathbf{X})$ can be obtained as

$$\begin{aligned} p(\mathbf{Y}|\mathbf{X}) &\propto \prod_{k=1}^{\rho} \exp\left(-\frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right) \\ &= \exp\left(-\sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right). \end{aligned} \quad (2.6)$$

Substituting (2.6) into (2.3), we get

$$\hat{\mathbf{X}}^{ML} = \arg \max_{\mathbf{X}} \exp\left(-\sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right). \quad (2.7)$$

ii) **MAP estimation approach:** In contrast to that the ML SR approach only considers the relationship among the observed low-resolution images and the original high-resolution image, the MAP SR approach further incorporates the prior image model to reflect the expectation of the unknown high-resolution image. These approaches provide a flexible framework for the inclusion of prior knowledge of the unknown high-resolution image, as well as for modeling the dependence between the observed low-resolution images and the unknown high-resolution image. The aim of the MAP SR approach is to find the MAP estimation of the high-resolution image (denoted as $\hat{\mathbf{X}}^{MAP}$) via

$$\hat{\mathbf{X}}^{MAP} = \arg \max_{\mathbf{X}} P(\mathbf{X}|\mathbf{Y}). \quad (2.8)$$

To evaluate the above pdf, the Bayes rule is applied to rewrite $p(\mathbf{X}|\mathbf{Y})$ as

$$p(\mathbf{X}|\mathbf{Y}) = \frac{p(\mathbf{X}, \mathbf{Y})}{p(\mathbf{Y})} \propto p(\mathbf{X}, \mathbf{Y}) = p(\mathbf{Y}|\mathbf{X}) p(\mathbf{X}). \quad (2.9)$$

2.1 Super-resolution Image Reconstruction

Now, the computation of (2.9) boils down to evaluating each product term, where $p(\mathbf{Y}|\mathbf{X})$ is the conditional pdf of the low-resolution images given the high-resolution image, $p(\mathbf{X})$ is the prior pdf of the high-resolution image, respectively.

The prior image model plays a key role for determining the reconstructed high-resolution image. For that, several models have been exploited, such as the *Huber-Markov random field* model (HMRF) [81], the *conditional random field* (CRF) [82], the *discontinuity adaptive MRF* (DAMRF) [83], the prior image model with the incorporation of the segmentation information [84–86], and the prior image model particularly for text images [87] or face images [88, 89]. On the other hand, various methods haven been developed to derive the relationship between the low-resolution images and the high-resolution image using the graph cut [90], manifold learning [91], dynamic tree [92], and tensor voting [93].

2.1.5 Learning-based SR Image Approach

Recently, learning-based techniques are proposed to solve the SR problem, where the new information required for predicting the high-resolution image can be further obtained from a set of training images, rather than from the low-resolution observations only [94, 95]. Commonly, the learning-based SR approaches fall into the following two categories: i) exploit the training set only, and ii) jointly consider the information provided by the training set and the low-resolution observations. The first approach aims to enhance the high-frequency information of the given single low-resolution image, by retrieving the most likely high-frequency information from the given training image samples based on the local features of the input low-resolution image. The second approach jointly exploits the information learnt

2.1 Super-resolution Image Reconstruction

Figure 2.3. A conceptual framework of learning-based SR image approach.

from a given high-resolution training data set, as well as that provided by multiple low-resolution observations.

Exploit The Training Set Only

Hertzmann [96] proposed an image analogy method to *create* the high-frequency details for the observed low-resolution images from a training image database, as illustrated in Figure 2.3. It contains two stages: an off-line training stage and a SR reconstruction stage. In the off-line training stage, pairs of low-resolution patches and the corresponding (ground truth) high-frequency patches are collected. In the SR reconstruction stage, the patches extracted from the input low-resolution images are compared with those stored in the database. Then the best matching patches are selected according to some criteria (e.g., *nearest distance*) as the corresponding high-frequency patches used for producing the high-resolution image.

Candocia and Principe [97, 98] proposed to produce the high-resolution image by exploiting the local correlations among the neighbors within the input low-resolution image. First, a feature extraction is performed on local neighborhood information from the low-resolution image. These features are then used to classify the neighborhoods into various sets, for which an optimal family kernels are trained offline to map these neighborhoods to their high-resolution counterparts. Finally, a

2.1 Super-resolution Image Reconstruction

high-resolution image is obtained through convolving the low-resolution image with the above-determined family of kernels.

Freeman [99] described an image interpolation algorithm which uses a database of training images to create the unknown high-frequency details in zoomed images. By blurring and down-sampling sharply defined images they construct a training set of sharp and blurred images, which are then incorporated into a Markov network to learn their relationship. [99] was improved to handle the LR image with unknown PSF parameters by Begin and Ferrie [100], and to handle the unknown blurring parameters by Qiao *et al.* [101].

Pickup *et al.* [102] addressed the SR problem for texture images. Given a small patch around any particular pixel position of the low-resolution image, their method learns the intensity distribution for the central pixel by examining the values at the centers of similar patches available in the training data, then determines the intensity value at this pixel position according to the above-estimated distribution.

In contrast to that the generation of a high-resolution image patch depends on only one of the nearest neighbors in the training set as used in the above-mentioned SR approaches [96–99, 102], Chang *et al.* [103] argued that the generation of the high-resolution image patch depends simultaneously on multiple nearest neighbors in the training set in a way similar to the concept of manifold learning methods, particularly the *locally linear embedding* (LLE) method. By this way, this method requires fewer training examples than that of other learning-based SR methods [96–99, 102].

Exploit The Training Set Plus Observed Low-resolution Images

A learning and wavelet based approach was proposed by Jiji *et al.* in [104]. The wavelet coefficients at finer scales of the unknown high-resolution image are learned from a set of training images, then the learned image in the wavelet domain is used

2.1 Super-resolution Image Reconstruction

for regularization while super-resolving the input observed low-resolution images. The prior image model has two terms: i) the smoothness term, which ensures the spatial correlation among the pixels; and ii) the learning term, which chooses the best edges from the training set. Their method was further improved by Jiji and Chaudhuri in [105] by exploiting the contourlet transform.

Li [106] proposed to reconstruct the high-resolution images by alternating the projections onto two constraint sets: the observation constraint defined by the given low-resolution image and the prior constraint derived from the training high-resolution data set. His work is built upon both the parametric wavelet-based texture synthesis and the non-parametric example-based SR approach [99]. Furthermore, two types of prior constraints are considered in his algorithm: the spatially homogeneous constraint is suitable for texture images, while the patch-based inhomogeneous constraint is suitable for generic images.

Datsenko and Elad [107] tackled the SR problem specifically for the scanned document image containing written texts and graphics. Their algorithm firstly assigned several candidate high-quality patches at each pixel position in the observed low-resolution image; these are found as the nearest-neighbors in an image-database that contains pairs of corresponding low-resolution and high-resolution image patches. These found patches are used as the prior image model, and are then merged into a MAP cost function to arrive at the closed-form solution of the desired high-resolution image.

Many works have been attempted to interpolate the low-resolution image based on an interpolation kernel, which is adaptively changed according to the input image and is learnt from the training data set. This interpolation kernel can be determined via an *artificial neural network* (ANN) [108], a look-up table algorithm [109], a probabilistic neural network [110], or the *support vector regression* (SVR) algorithm [111]. On the other hand, the regularization term is determined by the

2.2 Super-resolution Video Reconstruction

training set, and then is incorporated into a Bayesian framework to produce the high-resolution image [112, 113]. There has also been some efforts on applying an output feedback while performing the SR computation, particularly for face images [114] and text images [115], where the partially super-resolved image is first matched to a database to extract the correct match and then this information can be used to enhance the prior image model for further improving the image quality.

2.2 Super-resolution Video Reconstruction

The SR image approaches reconstruct a single high-resolution image from a set of given low-resolution images; likewise, the SR video approaches reconstruct an image sequence with a higher resolution from a group of adjacent low-resolution uncompressed image frames or compressed image frames. The major challenges in SR video reconstruction lie in several parts: i) The first challenge is how to exploit the additional information provided by the temporal correlated low-resolution images to improve the quality of the desired high-resolution images. ii) The second challenge is how to consider the uncertainty (i.e., estimation error) incurred in motion estimation of the SR computation. The reliable motion estimation among the observed images is most essential for high quality SR reconstruction. The SR algorithms are usually required to be robust with respect to these errors.

The existing SR video approaches can be classified into the following three categories: i) the sliding-window-based SR video approach, ii) the simultaneous SR video approach, and iii) the sequential SR video approach. In the following, an observation model of the SR video reconstruction will be presented, followed by a review on the above-mentioned three kinds of SR video approaches.

2.2.1 Observation Model

At each temporal instance n , the relationship between the low-resolution image and the high-resolution image can be mathematically formulated as

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{X}_n + \mathbf{V}_n, \quad (2.10)$$

where $\mathbf{y}_n$ and $\mathbf{X}_n$ represent the $L_1 \times L_2$ low-resolution image and the $M_1 \times M_2$ high-resolution image in the temporal instance n , respectively. Both of these two images are represented in the lexicographic-ordered vector form, with a size of $L_1 L_2 \times 1$ and $M_1 M_2 \times 1$, respectively. $\mathbf{H}_n$ is defined as in (2.2). Additionally, $\mathbf{V}_n$ is modeled as a zero-mean white Gaussian noise vector (with a size of $L_1 L_2 \times 1$), encountered during the image acquisition process. Furthermore, $\mathbf{V}_n$ is assumed to be independent of $\mathbf{X}_n$.

In summary, the SR video reconstruction is an *inverse problem*, and its goal is to make use of a set of low-resolution images $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ with a dimension of $L_1 \times L_2$ each to compute the high-resolution ones $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$ with a dimension of $M_1 \times M_2$ each (i.e., the enlarged resolution ratio is $q_1 \times q_2$).

2.2.2 Sliding-window-based SR Video Approach

A direct way to apply the SR image techniques for conducting SR video reconstruction, as what have been done by most of the existing SR approaches for video signals, is based on the *sliding-window*-based approach [11], as illustrated in Figure 2.4. The sliding window selects a set of consecutive low-resolution frames for producing one high-resolution image frame; that is, the window is moved across the input frames to produce successive high-resolution frames sequentially. The major drawback of this approach is that the temporal correlations among the consecutively-reconstructed high-resolution images are not considered.

2.2 Super-resolution Video Reconstruction

Figure 2.4. The sliding-window-based super-resolution video scheme.

Suresh *et al.* [116] proposed a SR method for the license plate video, where the desired high-resolution image is estimated from the observations by a non-convexity optimization procedure. Furthermore, a discontinuity adaptive regularization is used to preserve the edges in the reconstructed high-resolution image for improved visual image quality. Narayanan *et al.* [117] proposed to use the *partition-based weighted sum* (PWS) filter to simultaneously perform nonuniform interpolation and deconvolution of the PSF in the image acquisition process. Ng *et al.* [118] exploited the *total variation* (TV) regularization. Moreover, they addressed a SR computation case when there exist some missing or corrupted pixels in the observed low-resolution images. For that, they exploited the image inpainting algorithm to perform image inpainting and SR reconstruction simultaneously. Patanavijit and Jitapunkul [119] proposed a Bayesian MAP estimation framework, where the *Lorentzian error norm* is used for measuring the relationship between the reconstructed high-resolution image and each low-resolution image; moreover, the *Lorentzian-Tikhonov* regularization is used to remove artifacts from the reconstructed high-resolution image and accelerate the convergence of the iterative SR computation process.

2.2 Super-resolution Video Reconstruction

Figure 2.5. The simultaneous multiple high-resolution frames reconstruction scheme.

Several SR algorithms have been developed for the color video SR reconstruction, where the aim is to estimate the high-resolution color frames from a low-resolution compressed color video sequence. With more color components available, the motion estimation performance can be improved by considering the information provided by all of the color channels [120,121].

2.2.3 Simultaneous SR Video Approach

In contrast to that the temporal correlations among the consecutively-reconstructed high-resolution images are not considered in the sliding-window-based SR video approaches, many works have been done to simultaneously reconstruct multiple high-resolution images at one time, as illustrated in Figure 2.5. Borman and Steven-son [122] proposed an algorithm to reconstruct multiple high-resolution images simultaneously by imposing temporal smoothness constraint on the prior image model, while Zibetti and Mayer [123–125] reduced its computational complexity by only considering the temporally-smooth low-resolution frames in the observa-tion model. Alvarez *et al.* [126] proposed a multi-channel SR approach for the compressed image sequence, which applies the sliding window to select a set of

2.2 Super-resolution Video Reconstruction

low-resolution frames to reconstruct another set (same number of frames) of high-resolution frames simultaneously. However, all these three algorithms require large memory storage, since these multiple high-resolution and low-resolution images are required available at the same time during the reconstruction process. Moreover, how many high-resolution frames should be exploited for simultaneous reconstructions is another important issue, which has not been addressed.

2.2.4 Learning-based SR Video Approach

Bishop *et al.* [127] proposed a learning-based SR video method based on the principle of example-based SR still image approach developed in [99]. Their method uses a learnt data set of image patches capturing the relationship between the low-frequency and the high-frequency bands of natural images and uses an appropriate prior image model over such patches. Furthermore, their method uses the previously-enhanced frame to provide part of the training set for producing the current high-resolution frame.

Chapter 3

Markov Chain Monte Carlo

Super-resolution Image

Reconstruction

3.1 Introduction

The aim of SR image reconstruction is to fuse a set of similar low-resolution images acquired from the same scene to produce a single higher-resolution image with more details. Motivated by the fact that the SR image reconstruction computation is, in essence, an ill-posed inverse problem [79], numerous Bayesian-based approaches [81, 128, 129] have been developed for addressing this issue. The basic idea of these approaches is to exploit a regularization strategy for solving the ill-posed nature of the SR image reconstruction computation by incorporating the prior knowledge of unknown high-resolution image. The unknown high-resolution image can be estimated via some statistics (e.g., mean) according to the probability distribution of unknown high-resolution image, which is usually established by applying Bayesian inference to exploit the information provided by both the observed low-resolution

3.1 Introduction

images and the prior knowledge of unknown high-resolution image. The major challenge of these Bayesian-based approaches is that the formulated probability distribution of unknown high-resolution image is usually very complicate; thus, its statistics is quite difficult to compute. To tackle this challenge, the *Markov chain Monte Carlo* (MCMC) method [130] is considered in this thesis as a promising approach, since it provides an extremely powerful and efficient way to compute the statistics of any complicate distribution.

The Monte Carlo method, which is placed among the top ten scientific computing algorithms that have had the greatest impact on the development and practice of science and engineering in the 20-th century [131], firstly became popular in physics [42], and it has been applied to numerous applications [130]. Despite its rather long history, the exploitation of the Monte Carlo method in signal processing is fairly recent [43, 85, 86, 132–138]. The basic idea of the Monte Carlo method is to provide an accurate estimation of the unknown target (in our case, the high-resolution image) through stochastic simulations. For that, a fairly large number of samples are generated according to the target probability distribution, followed by discarding those initially-generated unreliable samples such that more accurate statistical measurement of the target distribution can be obtained. By further incorporating the first-order Markov chain property (that is, the generation of the current sample only depends on the previous one) into the Monte Carlo method [42] for generating sufficiently large number of reliable samples as mentioned above, the established MCMC process [130] is exploited for the development of our proposed stochastic method. Furthermore, these reliable samples are two-dimensional image each and generated according to the information provided by both the observed low-resolution images and the prior knowledge of unknown high-resolution image. Finally, the high-resolution image is produced by taking average of the above-mentioned image samples. However, there are certain fundamental

issues required to be addressed for the proposed MCMC approach, as follows.

The first issue is how to determine the optimal hyperparameter value of the chosen prior image model. It is important to note that the hyperparameter controls the degree of regularization and intimately affects the quality of the reconstructed high-resolution image. To tackle this problem, some works have been developed to automatically determine the prior image model's hyperparameter value during the process of estimating the unknown high-resolution image [139–142], rather than applying an experimentally-determined fixed value [143]. Tipping and Bishop [139] exploited the *expectation-maximization* (EM) algorithm for estimating the hyperparameter value by maximizing its marginal likelihood function. However, the EM algorithm used in this method yields tremendous computational cost, and more importantly, the EM algorithm does not always converge to the global optimum [144]. Pickup *et al.* [140] improved the method [139] by reducing its computational load and further modified the prior image model by considering illumination changes among the captured low-resolution images. He and Kondi [141] proposed a method to jointly estimate both the hyperparameter value and the high-resolution image by optimizing a single cost function. However, this method assumes that the displacements among the low-resolution images are restricted to integer-valued shifts on the high-resolution grid; obviously, this assumption often mismatches the real-world image acquisition. In this chapter, the prior image model's hyperparameter value is updated through each sample-generation iteration. More specifically, the *Metropolis-Hastings* (M-H) algorithm [42] is exploited to simultaneously update the hyperparameter value, along with the updating of the image sample using the *Gibbs sampler* [43], at each sample-generation step throughout the proposed MCMC process.

The second fundamental issue is the determination of when the sample generated starts becoming reliable for producing the high-resolution image. Typically,

3.1 Introduction

the MCMC process starts from an arbitrarily-generated sample as the seed, and based on which a fairly large number of samples are generated through iterations. The period before samples become reliable is known as the *burn-in* period, and the samples generated during this period are considered unreliable and should be discarded. Only the follow-up generated samples (i.e., after the burn-in period) are viewed reliable and used to produce the high-resolution image. In this chapter, a mathematical analysis is conducted to derive a bound for determining the length of the burn-in period of the proposed MCMC process.

The third critical issue is that image artifacts could appear in the reconstructed high-resolution image, if the image was estimated based on insufficient number of samples. Like any stochastic simulation task, the MCMC-based approach also requires a fairly large number of samples generated for estimating the unknown high-resolution image. Depending on the image content, the number of samples generated for the MCMC process could be insufficient for certain areas, such as edges and textures. This could lead to the presence of image artifacts in the reconstructed high-resolution image. To tackle this problem, our strategy is to incorporate a newly-developed post-processing method, called the *variation-sensitive* bilateral filtering, into the proposed MCMC approach for enhancing the reconstructed high-resolution image. The proposed artifact removal method contains two sequential stages: i) a *detection* stage to check whether the phenomenon of insufficient number of samples has incurred at *each* pixel position; and then ii) a *filtering* stage to remove image artifacts from the reconstructed high-resolution image, by applying either the conventional bilateral filter [23] (for “sufficient number” case) or its modified version (for “insufficient number” case), according to the binary decision result obtained at the detection stage. For the modified bilateral filter, a median pre-processing operation is performed for replacing the original intensity values of the estimated high-resolution image.

3.2 Proposed MCMC Super-resolution Approach

The forth significant issue is how to further enhance the edge and detail information in the reconstructed high-resolution image, an edge-directed MCMC SR approach is proposed. Steered by an edge map of the desired high-resolution image, the proposed method can adaptively enhance the edge information at the edge pixel positions while exploiting the conventional MCMC SR at the rest pixel positions.

The chapter is organized as follows. In Section 3.2, a Bayesian formulation of the SR image reconstruction is established, and an MCMC algorithm is developed. The bound for the length of the burn-in period of the proposed MCMC process is derived in Section 3.3. The variation-sensitive bilateral filtering is proposed in Section 3.4 as a post-processing stage for addressing the potential image artifacts issue. An edge-directed MCMC SR approach is proposed in Section 3.5 to further enhance the edge information presented in the reconstructed high-resolution image. Extensive simulation experiments are conducted to validate our derived theoretical results throughout the chapter. Finally, Section 3.6 concludes this chapter.

3.2 Proposed MCMC Super-resolution Approach

3.2.1 Observation Model

Given an image, it is considered as the ‘ground truth’ of the high-resolution image to be constructed from a set of low-resolution images for objective performance evaluation. To simulate the observed low-resolution observations, the given image is warped, convolved with a *point spread function* (PSF), downsampled to a lower resolution, and finally added with a zero-mean white Gaussian noise. These operations can be mathematically expressed as

$$\mathbf{y}^{(k)} = \mathbf{H}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}, \quad (3.1)$$

3.2 Proposed MCMC Super-resolution Approach

where $\mathbf{y}^{(k)}$ and $\mathbf{X}$ represent the k -th $L_1 \times L_2$ low-resolution image and the $M_1 \times M_2$ high-resolution image represented in the lexicographic-ordered vector form, with a size of $L_1 L_2 \times 1$ and $M_1 M_2 \times 1$, respectively. $\mathbf{H}^{(k)}$ is an $L_1 L_2 \times M_1 M_2$ matrix, representing the above-mentioned warping (e.g., shift and rotation), convolving and downsampling operations, and $\mathbf{V}^{(k)}$ is an $L_1 L_2 \times 1$ vector, representing the above-mentioned additive white Gaussian noise with a zero-mean and a variance σ_k^2 .

With such establishment, the goal of SR image reconstruction is to produce a single high-resolution image $\mathbf{X}$ based on a set (say, cardinality ρ) of low-resolution observations $\mathbf{Y} = \{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(\rho)}\}$ with a size of $L_1 L_2 \times 1$ in each column vector.

3.2.2 Bayesian Inference

In this section, a Bayesian inference approach is exploited to establish the joint posterior conditional *probability density function* (pdf) of the unknown high-resolution image (denoted as $\mathbf{X}$) and the prior image model's hyperparameter (denoted as λ) given the observed low-resolution images (denoted as $\mathbf{Y}$); that is $p(\mathbf{X}, \lambda | \mathbf{Y})$. To evaluate the above joint pdf, the Bayes rule is applied to rewrite $p(\mathbf{X}, \lambda | \mathbf{Y})$ as

$$p(\mathbf{X}, \lambda | \mathbf{Y}) = \frac{p(\mathbf{X}, \lambda, \mathbf{Y})}{p(\mathbf{Y})} \propto p(\mathbf{X}, \lambda, \mathbf{Y}). \quad (3.2)$$

Furthermore, (3.2) can be rewritten as

$$p(\mathbf{X}, \lambda, \mathbf{Y}) = p(\mathbf{Y} | \mathbf{X}, \lambda) p(\mathbf{X}, \lambda) = p(\mathbf{Y} | \mathbf{X}, \lambda) p(\mathbf{X} | \lambda) p(\lambda). \quad (3.3)$$

Substituting (3.3) into (3.2), we have

$$p(\mathbf{X}, \lambda | \mathbf{Y}) \propto p(\mathbf{Y} | \mathbf{X}, \lambda) p(\mathbf{X} | \lambda) p(\lambda). \quad (3.4)$$

Now, the computation of (3.4) boils down to evaluating each product term, where $p(\mathbf{Y} | \mathbf{X}, \lambda)$ is the conditional pdf of the low-resolution images given the high-resolution image and the prior image model's hyperparameter, $p(\mathbf{X} | \lambda)$ is the prior pdf of the

3.2 Proposed MCMC Super-resolution Approach

high-resolution image given the prior image model's hyperparameter, and $p(\lambda)$ is the prior pdf of the hyperparameter, respectively.

Firstly, since the low-resolution images are obtained independently from the original (high-resolution) image and the hyperparameter λ does not play any role in this stage, the conditional pdf $p(\mathbf{Y}|\mathbf{X}, \lambda)$ can be expressed as

$$p(\mathbf{Y}|\mathbf{X}, \lambda) = p(\mathbf{Y}|\mathbf{X}) = \prod_{k=1}^p p(\mathbf{y}^{(k)}|\mathbf{X}). \quad (3.5)$$

Furthermore, according to (3.1), the conditional pdf $p(\mathbf{y}^{(k)}|\mathbf{X})$ can be expressed as

$$\begin{aligned} p(\mathbf{y}^{(k)}|\mathbf{X}) &= p((\mathbf{H}^{(k)}\mathbf{X} + \mathbf{V}^{(k)})|\mathbf{X}) = p(\mathbf{V}^{(k)}|\mathbf{X}) \\ &= p(\mathbf{V}^{(k)}) \propto \exp\left\{-\frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right\}, \end{aligned} \quad (3.6)$$

Substituting (3.6) into (3.5), the conditional pdf $p(\mathbf{Y}|\mathbf{X}, \lambda)$ can be obtained as

$$p(\mathbf{Y}|\mathbf{X}, \lambda) \propto \prod_{k=1}^p \exp\left(-\frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right) = \exp\left\{-\sum_{k=1}^p \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\|^2\right\}. \quad (3.7)$$

Secondly, since the image is locally smooth; hence, the *Markov random field* (MRF) [145] is considered as a useful prior image model in this chapter, which bears the following form [145]

$$p(\mathbf{X}|\lambda) = \frac{1}{Z(\lambda)} \exp\{-\lambda U(\mathbf{X})\}, \quad (3.8)$$

where λ is the so-called *hyperparameter*, and $Z(\lambda)$ is a normalizing factor, which is called the *partition function* and defined as [145]

$$Z(\lambda) = \int_{\chi} \exp\{-\lambda U(\mathbf{X})\} d\mathbf{X}, \quad (3.9)$$

where χ represents all possible configurations of the unknown high-resolution image $\mathbf{X}$, and the *energy function* $U(\mathbf{X})$ is made of [145]

$$U(\mathbf{X}) = \sum_{c \in \mathcal{C}} V_c(\mathbf{X}), \quad (3.10)$$

3.2 Proposed MCMC Super-resolution Approach

in which $V_c(\mathbf{X})$ is a *potential function* of a local group of pixels c (called the *clique*) and its value depends on the local configuration on the clique c . Hence, $U(\mathbf{X})$ is obtained by computing the potential function over all possible cliques (denoted as the set C). In our works, the *Gaussian Markov random field model* [146] is selected for formulating the prior image model, which has been proved to be effective for SR image reconstruction [128]; hence, (3.8) can be rewritten as [146]

$$p(\mathbf{X}|\lambda) = \frac{1}{Z(\lambda)} \exp \left\{ -\frac{1}{2} \lambda \mathbf{X}^T \mathbf{C} \mathbf{X} \right\}, \quad (3.11)$$

where $\mathbf{C}$ is an $M_1 M_2 \times M_1 M_2$ matrix, whose entries are given by [146]

$$C_{i,j} = \begin{cases} \gamma, & i = j; \\ -1, & i \text{ and } j \text{ are adjacent in the } \gamma\text{-neighborhood}; \\ 0, & \text{otherwise.} \end{cases} \quad (3.12)$$

Note that according to (3.12), matrix $\mathbf{C}$ is symmetrical with all the off-diagonal entries being negative.

Thirdly, the uniform distribution is chosen as the prior pdf of the hyperparameter $p(\lambda)$ in (3.4), which has been proved to be an effective assumption for image restoration problem [147], [148]; that is,

$$p(\lambda) = \frac{1}{\lambda_{max} - \lambda_{min}}, \quad \text{for } \lambda \in [\lambda_{min}, \lambda_{max}], \quad (3.13)$$

where $[\lambda_{min}, \lambda_{max}]$ represents the dynamic range of the value of λ .

Finally, substituting (3.7), (3.11) and (3.13) into (3.4), $p(\mathbf{X}, \lambda|\mathbf{Y})$ can be arrived at

$$\begin{aligned} p(\mathbf{X}, \lambda|\mathbf{Y}) &\propto p(\mathbf{Y}|\mathbf{X}, \lambda) p(\mathbf{X}|\lambda) p(\lambda) \\ &\propto \exp \left\{ -\sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)} \mathbf{X}\|^2 \right\} \frac{1}{Z(\lambda)} \exp \left\{ -\frac{1}{2} \lambda \mathbf{X}^T \mathbf{C} \mathbf{X} \right\} \frac{1}{\lambda_{max} - \lambda_{min}} \\ &= \frac{1}{Z(\lambda)} \exp \left\{ -\sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)} \mathbf{X}\|^2 - \frac{1}{2} \lambda \mathbf{X}^T \mathbf{C} \mathbf{X} \right\} \frac{1}{\lambda_{max} - \lambda_{min}}. \end{aligned} \quad (3.14)$$

3.2 Proposed MCMC Super-resolution Approach

Obviously, the direct assessment of (3.14) is extremely difficult, since the computation of $Z(\lambda)$ needs to integrate over all possible configurations of the unknown high-resolution image $\mathbf{X}$ (i.e., over the space χ in (3.9)). For example, given an N -bit gray-level image with a size of $M_1 \times M_2$, the space χ consists of $N^{M_1 \times M_2}$ configurations. Furthermore, the function $Z(\lambda)$ depends on the hyperparameter λ , whose optimal value is difficult to determine. For that, the MCMC method has been observed as a potential approach for tackling this problem.

3.2.3 Proposed MCMC SR Approach

The goal of the proposed MCMC algorithm is to simultaneously generate N sets of samples for estimating the high-resolution image $\mathbf{X}$ and the hyperparameter λ based on $p(\mathbf{X}, \lambda | \mathbf{Y})$ [24, 149]; that is, the image samples, $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(N)}$ and their associated hyperparameter samples, $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(N)}$, where N is large enough to guarantee the convergence of the MCMC process. Note that all image samples have the same size as that of $\mathbf{X}$. Furthermore, the initially-generated T image samples during the *burn-in* period are considered unreliable and thus should be discarded. Finally, the high-resolution image is estimated by computing the mean of the follow-up generated $(N - T)$ reliable image samples; that is,

$$\hat{\mathbf{X}} = \frac{1}{N - T} \sum_{i=T+1}^N \mathbf{X}^{(i)}. \quad (3.15)$$

The determination of T in (3.15) is mathematically derived in this chapter and presented in Section 3.3.

The sample-generation process is described in detail as follows. The proposed MCMC process starts from the randomly-initialized sample set: $\mathbf{X}^{(1)}$ and $\lambda^{(1)}$. Then, at each follow-up sample-generation step, the *Gibbs sampler* [43] is applied to generate a new image sample, and the *Metropolis-Hastings* (M-H) algorithm [42] is used to generate a new hyperparameter value. Note that the generation

3.2 Proposed MCMC Super-resolution Approach

of $\mathbf{X}^{(n)}$ and $\lambda^{(n)}$ only depend on $\mathbf{X}^{(n-1)}$ and $\lambda^{(n-1)}$. This process is repeated as such to generate a fairly large N sets of samples. Note that the M-H algorithm is computationally efficient enough to generate the one-dimensional samples that meet the target distribution, while the Gibbs sampler is suitable to handle high-dimensional sample-generation [150]; in our case, each sample is a two-dimensional image. More details are described as follows.

At each sample-generation step, the Gibbs sampler randomly selects one pixel location of the current image sample, followed by updating the intensity value of this pixel, according to its conditional distribution given the remaining pixels of the current image sample, the hyperparameter in the prior image model and the observed low-resolution images. To be more specific, let X_n denote the n -th component of the lexicographic-order $M_1 M_2 \times 1$ vector $\mathbf{X}$, and $\mathbf{X}_{-n}$ denote the same vector of $\mathbf{X}$ except with the n -th component being excluded; that is, $\mathbf{X}_{-n} = \{X_1, X_2, \dots, X_{n-1}, X_{n+1}, \dots, X_{M_1 M_2}\}$. According to (3.14), the probability density function of X_n , conditional on the remaining components of the high-resolution image $\mathbf{X}_{-n}$, the hyperparameter in the prior image model λ , and the observed low-resolution images $\mathbf{Y}$, can be expressed as

$$p(X_n | \mathbf{X}_{-n}, \lambda, \mathbf{Y}) \propto \exp \left\{ -\frac{1}{2\tau_{X_n}} (X_n - \mu_{X_n})^2 \right\}, \quad (3.16)$$

where

$$\mu_{X_n} = \frac{\tau_{X_n}}{2} \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^L \frac{\psi_i^{(k)}(X_n) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(X_n) \right), \quad (3.17)$$

$$\tau_{X_n} = \frac{1}{\sum_{k=1}^{\rho} \sum_{i=1}^L \frac{(H_{i,n}^{(k)})^2}{\sigma_k^2} + \lambda C_{n,n}}, \quad (3.18)$$

in which $\psi_i^{(k)}(X_n)$ and $\phi(X_n)$ are defined as

$$\psi_i^{(k)}(X_n) = y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} X_j, \quad (3.19)$$

$$\phi(X_n) = \sum_{s \neq n} X_s C_{s,n} + \sum_{t \neq n} C_{n,t} X_t. \quad (3.20)$$

3.2 Proposed MCMC Super-resolution Approach

The detailed proof of (3.16) is provided in Section 3.6.

For updating the hyperparameter λ at each sample-generation step, the M-H algorithm [130] first generates a candidate sample (denoted as $\lambda^{(t+1)*}$) according to a Gaussian distribution with the mean $\lambda^{(t)}$ and the variance $\alpha \times \lambda^{(t)}$, where $\alpha = 0.5$ is empirically determined in this chapter. Depending on the *acceptance probability* measured, which is defined by $A = \min\{1, \frac{p(\mathbf{X}^{(t)}, \lambda^{(t+1)*} | \mathbf{Y})}{p(\mathbf{X}^{(t)}, \lambda^{(t)} | \mathbf{Y})}\}$, the above-mentioned candidate sample $\lambda^{(t+1)*}$ will be either accepted as the new hyperparameter sample (i.e., $\lambda^{(t+1)} = \lambda^{(t+1)*}$) or rejected (i.e., $\lambda^{(t+1)} = \lambda^{(t)}$). The proposed MCMC approach can be summarized as follows.

- Randomly generate $\mathbf{X}^{(1)}$ and $\lambda^{(1)}$ as the initialized sample set.
- For the sample index $t = 1 : N - 1$,
 - Select one pixel n from the current image sample $\mathbf{X}^{(t)}$, where n is randomly drawn from the discrete uniform distribution over the interval $[1, M_1 M_2]$.
 - Only the intensity value of the above-selected pixel n gets updated for establishing the next image sample $\mathbf{X}^{(t+1)}$; that is, $\mathbf{X}^{(t+1)} = \{X_1^{(t)}, X_2^{(t)}, \dots, X_{n-1}^{(t)}, X_n^{(t+1)}, X_{n+1}^{(t)}, \dots, X_{M_1 M_2}^{(t)}\}$. The new intensity value at the n -th component $X_n^{(t+1)}$ is generated according to the distribution $p(X_n^{(t+1)} | \mathbf{X}_{-n}^{(t)}, \lambda^{(t)}, \mathbf{Y})$, which is defined in (3.16).
 - Generate a candidate sample $\lambda^{(t+1)*}$ according to a Gaussian distribution with the mean $\lambda^{(t)}$ and the variance $0.5 \times \lambda^{(t)}$, where the value 0.5 is empirically determined in this chapter.
 - Whether the candidate sample $\lambda^{(t+1)*}$ can be accepted as the hyperparameter sample $\lambda^{(t+1)}$ or not is decided according to the acceptance

3.2 Proposed MCMC Super-resolution Approach

probability A , which is given by

$$A = \min \left\{ 1, \frac{p(\mathbf{X}^{(t)}, \lambda^{(t+1)*} | \mathbf{Y})}{p(\mathbf{X}^{(t)}, \lambda^{(t)} | \mathbf{Y})} \right\}. \quad (3.21)$$

Note that if $\lambda^{(t+1)*}$ is rejected, $\lambda^{(t)}$ will be remained to be used; that is, $\lambda^{(t+1)} = \lambda^{(t)}$.

- Discard the initial T samples and use the follow-up generated $(N - T)$ samples only to estimate the high-resolution image $\hat{\mathbf{X}}$ via (3.15).

Furthermore, there are two critical issues required to be solved for the proposed MCMC process. The first one is how many initially-generated image samples are considered *unreliable* and thus they should be discarded. This is equivalent to the determination of the value of the burn-in parameter T in (3.15). To tackle this issue, a closed-form bound has been mathematically derived and presented in Section 3.3. The second critical issue is that image artifacts could be inevitably yielded in the reconstructed high-resolution image, if the number of image samples generated for the reconstruction process is not sufficient; that is, when $(N - T)$ in (3.15) is small. For that, the *variation-sensitive* bilateral filtering is proposed to further enhance the reconstructed high-resolution image and presented in Section 3.4.

3.2.4 Experimental Results

Simulation Setup

To explore the performance of the proposed MCMC approach, a 256×256 *Boat* and a 200×200 *Text* images are used as the test images in the experiments. In our simulation, four low-resolution images are independently generated from each original image by applying a series of operations as mathematically formulated in (3.1). In detail, we first apply a shift, which is drawn from a continuous uniform distribution over the interval $(-2, 2)$ in the units of pixels, followed by applying a

3.2 Proposed MCMC Super-resolution Approach

rotation, which is drawn from a continuous uniform distribution over the interval $[-4, 4]$ in the units of degrees. Next, the processed images are convolved with a point spread function, which is a Gaussian low-pass filter with a size of 4×4 and the standard deviation of 2, followed by a down-sampling operation with a decimation factor of two in both the horizontal and the vertical directions, respectively. Lastly, the processed image is added with a zero-mean white Gaussian noise with a standard deviation 8 to yield a noisy image.

Implementation

The proposed algorithm produces the high-resolution image as follows. Firstly, the initial image sample $\mathbf{X}^{(1)}$ is randomly generated; that is, all the pixel intensities are independently drawn from the discrete uniform distribution over the interval $[\theta_{min}, \theta_{max}]$ (for example, $\theta_{min} = 0$, and $\theta_{max} = 255$ for an 8-bit monochrome image). Alternatively, one can use an interpolated version of any one of the observed low-resolution input images (using, say, pixel duplication or the bi-cubic interpolation method) as the initial image sample $\mathbf{X}^{(1)}$ to shorten the burn-in period of the proposed MCMC process. Meanwhile, the initial hyperparameter sample $\lambda^{(1)}$ is drawn from the continuous uniform distribution over the interval $[\lambda_{min}, \lambda_{max}]$, where λ_{min} and λ_{max} are experimentally set to be 1 and 50, respectively.

Next, the burn-in parameter T is determined via (3.27). In our simulation, the user-defined tolerance ϵ is set to be 10, and the neighborhood's size γ is set to be 4. The shift and rotation parameters for the observed low-resolution images are assumed to be known; otherwise, it can be estimated using the existing image registration methods such as [20]. The blurring parameter is assumed to be known; this assumption has been used in several existing SR image reconstruction experiments [128, 129]. The MCMC method is then conducted through $N = k \times T$ iterations, and the image sample plus the hyperparameter value get updated in

3.2 Proposed MCMC Super-resolution Approach

each iteration. The multiplicity parameter k can be chosen by the user so that N is large enough to guarantee that not only the statistics of the target distribution is met, but also image artifacts are absent from the final reconstructed high-resolution image. Our experimental simulations have suggested that $k = 10$. Note that, in practice, the above parameter k can be determined under the criterion that if the currently-updated image sample $\mathbf{X}^{(i)}$ no longer makes any appreciable change on estimating $\mathbf{X}$ via (3.15), then the entire MCMC process could be stopped.

Finally, the first T image samples are discarded, since they are viewed as unreliable, and the remaining $(N - T)$ reliable image samples are used to produce the high-resolution image via (3.15).

Experimental Results

In the first experiment, we compare the proposed MCMC algorithm with four SR approaches: a *bi-cubic interpolation* approach, followed by a deblurring presented in Section II.E in [22], three SR approaches developed in [20–22], respectively. All the parameters involved in the above approaches [20–22] are set as the same as that are suggested in the papers. The results of the above-mentioned five methods are presented in Table 3.1 (for the PSNR performance), Figures 3.1 (for *Boat*) and 3.2 (for *Text*), which clearly show that our proposed algorithm yields the best subjective image quality.

In our second experiment, we demonstrate that the choice of the prior image model’s hyperparameter is critical to the quality of the reconstructed high-resolution image. Furthermore, we show that the proposed MCMC algorithm can automatically update the prior image model’s hyperparameter. In the following, we compare the performance of exploiting the *fixed*-valued hyperparameter versus that of exploiting the *adaptive*-valued hyperparameter. Figures 3.3 (c) and (d) demonstrate the reconstructed high-resolution *Boat* image with two hyperparameters: $\lambda = 0.1$

3.2 Proposed MCMC Super-resolution Approach

and $\lambda = 300$, respectively. One can see that if the hyperparameter value is too small, the high-resolution image will be *under*-regularized (i.e., an under-smoothing case). On the other hand, if the hyperparameter is too large, it will cause the high-resolution image *over*-regularized (i.e., an over-smoothing case). Moreover, Figure 3.3 (e) presents the results obtained by the best hyperparameter, which is obtained by exhaustively experimenting individual integer value over the interval $[1, 50]$. Comparing Figure 3.3 (e) with Figure 3.3 (f), which is resulted from the proposed MCMC method with the incorporation of simultaneous adaptation of the prior image model's hyperparameter, one can see that the performance of the proposed method is fairly close to the case when the best hyperparameter value is exploited. Likewise, the above-concluded comments can be consistently applied to the results of the test image *Text*, as shown in Figure 3.4.

The third experiment is to explore the performance of our method for handling the severely degraded low-resolution observations in the following two cases: (i) severely noisy low-resolution observation; and (ii) severely blurred low-resolution observation. These severely degraded low-resolution images are generated from the original image (ground truth) by applying a series of operations, which are as the same as that used in the previous section (Simulation Setup) except that: (i) an additive white Gaussian noise with a zero mean and a standard derivation 20 is used for generating the severely noisy low-resolution observation; and (ii) a Gaussian low-pass filter with a size of 9×9 and a standard deviation of 5 is used for generating the severely-blurred low-resolution observation. The proposed MCMC SR approach is used to conduct 2×2 SR computation using four low-resolution images, including three low-resolution images generated as the previous section (Simulation Setup), plus one severely degraded low-resolution image. The experimental results are presented in Table 3.2 (for the PSNR performance), plus Figures 3.5 (for *Boat*) and 3.6 (for *Text*). As seen from the above Table and

3.2 Proposed MCMC Super-resolution Approach

Table 3.1
THE PSNR PERFORMANCE COMPARISON.

Test image	Bi-cubic spline interpolation + image deblurring	Vandewalle <i>et al.</i> 's approach [20]	Pham <i>et al.</i> 's approach [21]	Farsiu <i>et al.</i> 's approach [22]	Proposed approach
<i>Boat</i>	25.80 dB	25.97 dB	26.61 dB	27.45 dB	28.02 dB
<i>Text</i>	17.62 dB	18.32 dB	17.74 dB	19.29 dB	20.17 dB

Table 3.2
THE PSNR PERFORMANCE COMPARISON OF 2×2 SR COMPUTATION OF THE PROPOSED MCMC SR APPROACH USING FOUR LOW-RESOLUTION OBSERVATIONS (INCLUDING ONE SEVERELY DEGRADED LOW-RESOLUTION OBSERVATION).

Test image	Four low-resolution observations including one severely noisy image (Figure 3.5 (a) for <i>Boat</i> , and Figure 3.6 (a) for <i>Text</i>)	Four low-resolution observations including one severely blurred image (Figure 3.5 (c) for <i>Boat</i> , and Figure 3.6 (c) for <i>Text</i>)
<i>Boat</i>	27.16 dB	26.97 dB
<i>Text</i>	19.26 dB	19.41 dB

Figures, our method can still produce fairly good high-resolution image to handle the severely degraded low-resolution observations.

Computational Time

Experiments are conducted to compare the run tim of the proposed method with that of other above-mentioned four SR methods. The proposed method is implemented using the *C* programming language; the other four SR methods are implemented using the *Matlab* programming language. One hundred experiments are conducted for each SR method on a PC with a Pentium 2.4 GHz CPU and a 512 MB RAM, and their respective average run times are compared in Table 3.3.

Note that, to further shorten the run time of the proposed MCMC approach, one effective way is to use a bi-cubic interpolated version of any one of the low-resolution input images as the initial image sample, rather than starting with a

3.2 Proposed MCMC Super-resolution Approach

Table 3.3
THE RUN TIME COMPARISON.

Test image	Bi-cubic spline interpolation + image deblurring	Vandewalle <i>et al.</i> 's approach [20]	Pham <i>et al.</i> 's approach [21]	Farsiu <i>et al.</i> 's approach [22]	Proposed approach ($N = 10 \times T$)
<i>Boat</i>	3.89 seconds	4.06 seconds	20.58 seconds	1.89 seconds	648 seconds
<i>Text</i>	2.09 seconds	2.32 seconds	11.81 seconds	0.96 seconds	392 seconds

Table 3.4
THE RUN TIME OF THE BURN-IN PERIOD COMPARISON OF THE PROPOSED MCMC SR APPROACH USING TWO DIFFERENT STARTING IMAGE SAMPLES.

Test image	Randomly-generated image	A bi-cubic interpolated image (from any one of the low-resolution input images)
<i>Boat</i>	57.60 seconds	11.52 seconds
<i>Text</i>	30.15 seconds	4.52 seconds

randomly-generated image. Experiments are conducted to present how different of their resulted run time required for the burn-in period. One hundred experiments are conducted for each of the above-mentioned two implementation cases, and their average run times of the burn-in period are compared in Table 3.4. In these experiments, the burn-in period is viewed as expired if the *mean absolute difference* (MAD) between the currently-updated image sample $\mathbf{X}^{(i)}$ and the high-resolution image (obtained via the proposed MCMC approach, in the $N = 10 \times T$ case, as described in Section 3.3) is smaller than a user-defined tolerance τ ($\tau = 5$ in our experiments). As seen from Table 3.4, the burn-in period of the proposed MCMC process can be effectively reduced by using a bi-cubic interpolated version of any one of the low-resolution input images as the initial image sample.

3.2 Proposed MCMC Super-resolution Approach

Memory Requirement

In this section, the memory requirement of the proposed framework and that of the above-mentioned four SR approaches are compared. Throughout the following analysis, it is assumed that an $M_1 \times M_2$ high-resolution image needs to be reconstructed by performing 2×2 SR computation using four $L_1 \times L_2$ low-resolution image, where $L_1 = M_1/2$ and $L_2 = M_2/2$.

A study of the memory requirements of these five methods is provided as follows:

- The *bi-cubic interpolation* approach requires storing the low-resolution images (i.e., four $L_1 \times L_2$ matrices). Furthermore, the deblurring approach presented in Section II.E in [22] requires storing the low-resolution image (i.e., four $L_1 \times L_2$ matrices), and a temporarily-reconstructed high-resolution image (i.e., an $M_1 \times M_2$ matrix), plus four matrices $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \mathbf{H}^{(3)}, \mathbf{H}^{(4)}$. Assume the kernel of the PSF in $\mathbf{H}^{(k)}$ to be $p \times p$; therefore, each $\mathbf{H}^{(k)}$ is a sparse matrix, in which each column contains only $p^2 L_1 L_2 / M_1 M_2 = p^2 / 4$ non-zero entries.
- The SR approach [20] requires storing the low-resolution images (i.e., four $L_1 \times L_2$ matrices).
- The SR approach [21] requires storing the low-resolution images (i.e., four $L_1 \times L_2$ matrices).
- The SR approach [22] requires storing the low-resolution images (i.e., four $L_1 \times L_2$ matrices), a temporarily-reconstructed high-resolution image (i.e., an $M_1 \times M_2$ matrix) and four matrices $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \mathbf{H}^{(3)}, \mathbf{H}^{(4)}$, each of which is a sparse matrix, where each column contains only $p^2 L_1 L_2 / M_1 M_2 = p^2 / 4$ non-zero entries.
- The Proposed approach requires storing the low-resolution images (i.e., four $L_1 \times L_2$ matrices), and four matrices $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \mathbf{H}^{(3)}, \mathbf{H}^{(4)}$ (each of which is a

3.2 Proposed MCMC Super-resolution Approach

sparse matrix, where each column contains only $p^2 L_1 L_2 / M_1 M_2 = p^2 / 4$ non-zero entries), plus the high-resolution image samples (each of which has a size of $M_1 \times M_2$). There are two ways to store these image samples. The first way is to store all of the $N - T$ reliable image samples $\mathbf{X}^{(T+1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(N)}$, and then compute the final high-resolution image via $\hat{\mathbf{X}} = \frac{1}{N-T} \sum_{i=T+1}^N \mathbf{X}^{(i)}$. The second way is to only store a temporarily-reconstructed high-resolution image (denoted as $\hat{\mathbf{X}}^{(i)}$ at the i -th sample generation step). Then the temporarily-reconstructed high-resolution image is updated at each sample generation step by adding the current image sample $\mathbf{X}^{(i)}$ multiplied with a factor $\frac{1}{N-T}$; that is $\hat{\mathbf{X}}^{(i+1)} = \hat{\mathbf{X}}^{(i)} + \frac{1}{N-T} \mathbf{X}^{(i)}$. Comparing the above two ways, one can see that the second way has a lower memory requirement than that of the first way.

3.2 Proposed MCMC Super-resolution Approach

Figure 3.1. Various reconstructed results of *Boat* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation approach, followed by deblurring (PSNR = 25.80 dB); (e) Vandewalle *et al.*'s algorithm [20] (PSNR = 25.97 dB); (g) Pham *et al.*'s algorithm [21] (PSNR = 26.61 dB); (i) Farsiu *et al.*'s algorithm [22] (PSNR = 27.45 dB); (k) the proposed algorithm with the automatically-updated hyperparameter (PSNR = 28.02 dB); images (d), (f), (h), (j) and (l) are the absolute difference between image (a) and images (c), (e), (g), (i) and (k), respectively. Furthermore, the intensity values of these absolute difference images are scaled up three times for demonstration.

3.2 Proposed MCMC Super-resolution Approach

Figure 3.2. Various reconstructed results of *Text* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation approach, followed by deblurring (PSNR = 17.62 dB); (e) Vandewalle *et al.*'s algorithm [20] (PSNR = 18.32 dB); (g) Pham *et al.*'s algorithm [21] (PSNR = 17.74 dB); (i) Farsiu *et al.*'s algorithm [22] (PSNR = 19.29 dB); (k) the proposed algorithm with the automatically-updated hyperparameter (PSNR = 20.17 dB); images (d), (f), (h), (j) and (l) are the absolute difference between image (a) and images (c), (e), (g), (i) and (k), respectively.

3.2 Proposed MCMC Super-resolution Approach

Figure 3.3. Various reconstructed results of *Boat* image: (a) the original image; (b) the bi-cubic spline interpolation, followed by deblurring approach (PSNR = 25.80 dB); (c) the proposed method with the fixed-valued hyperparameter $\lambda = 0.1$ to demonstrate the *under-smoothing* case (PSNR = 24.65 dB); (d) the proposed method with the fixed-valued hyperparameter $\lambda = 300$ to demonstrate the *over-smoothing* case (PSNR = 24.85 dB); (e) the proposed method with the best hyperparameter $\lambda = 20$, which is obtained by exhaustively experimenting individual integer value over the interval $[1, 50]$ to demonstrate the *optimal* smoothing case (PSNR = 28.13 dB); and (f) the proposed algorithm with the *adaptive* hyperparameter (PSNR = 28.02 dB), which is fairly close to the optimal smoothing case as presented in (e).

3.2 Proposed MCMC Super-resolution Approach

Figure 3.4. Various reconstructed results of *Text* image: (a) the original image; (b) the bi-cubic spline interpolation, followed by deblurring approach (PSNR = 17.62 dB); (c) the proposed method with the fixed-valued hyperparameter $\lambda = 0.1$ to demonstrate the *under-smoothing* case (PSNR = 17.73 dB); (d) the proposed method with the fixed-valued hyperparameter $\lambda = 300$ to demonstrate the *over-smoothing* case (PSNR = 15.92 dB); (e) the proposed method with the best hyperparameter $\lambda = 8$, which is obtained by exhaustively experimenting individual integer value over the interval $[1, 50]$, to demonstrate the *optimal* smoothing case (PSNR = 20.36 dB); and (f) the proposed algorithm with the *adaptive* hyperparameter (PSNR = 20.17 dB), which is fairly close to the optimal smoothing case as presented in (e).

3.2 Proposed MCMC Super-resolution Approach

Figure 3.5. Various reconstructed results of *Boat* image of the proposed MCMC SR approach.

3.2 Proposed MCMC Super-resolution Approach

Figure 3.6. Various reconstructed results of *Text* image of the proposed MCMC SR approach.

3.3 Derivation of The Burn-in Parameter of the Proposed MCMC Super-resolution Process

3.3 Derivation of The Burn-in Parameter of the Proposed MCMC Super-resolution Process

Summarizing the previous sections, the proposed MCMC process starts from an arbitrarily-initialized image sample and then runs for a fairly large number of iterations so that not only the statistics of the target distribution is met, but also undesirable image artifacts could be avoided in the final reconstructed high-resolution image. The initially-generated T samples (where T is defined as the *burn-in* parameter) are considered unreliable and discarded. Only the follow-up generated samples (i.e., from the $(T + 1)$ -th sample onwards) are viewed reliable and used to produce the high-resolution image.

In this section, a closed-form bound for computing the value of the burn-in parameter T of the proposed MCMC process is mathematically derived. For that, our goal is to derive a bound to ensure that regardless what kind of initialized image sample is used, any chain of image samples generated for the proposed MCMC process will converge to the same result within the user-defined tolerance ϵ after T sample-generation steps. A lemma provided in [151] is served as the foundation of our derivation, which has been successfully applied for the image restoration problem [151], and it is succinctly described as follows. The MCMC method is applied to assess the target distribution on a bounded space $\mathcal{X}$. In the context of image processing, the bounded space means the dynamic range of the image's gray-scale levels (e.g., $[0, 255]$ for an 8-bit image). For any two independently-generated chains of samples (denoted as $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \dots$ and $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots$, respectively, and each of these samples has a size of $M_1 M_2 \times 1$) being generated for the MCMC process. Both $\mathbf{f}^{(1)}$ and $\mathbf{g}^{(1)}$ are arbitrarily-initialized samples. Then at each sample-generation step

3.3 Derivation of The Burn-in Parameter of the Proposed MCMC Super-resolution Process

t , there exists a constant $c \in (0, 1)$ such that [151]

$$E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \leq c \cdot d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)}), \quad (3.22)$$

where $E[\cdot]$ is the conditional expectation, and $d(\cdot)$ is the pixel-wise distance measurement [151]; that is,

$$d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)}) = \sum_{n=1}^{M_1 M_2} |f_n^{(t)} - g_n^{(t)}|, \quad (3.23)$$

in which $f_n^{(t)}$ and $g_n^{(t)}$ are the n -th component of the vectors $\mathbf{f}^{(t)}$ and $\mathbf{g}^{(t)}$, respectively. The MCMC process is considered “converged” after T sample-generation steps in the sense that the distance between the above two chains of samples is no larger than a user-defined tolerance ϵ , and T can be arrived at [151]

$$T = \frac{\ln(\epsilon/D(\mathcal{X}))}{\ln c}, \quad (3.24)$$

where c is defined in (3.22), ϵ is the user-defined tolerance, which is in terms of the distance defined as that in (3.23), and $D(\mathcal{X})$ is the largest distance of any two image samples in the space $\mathcal{X}$; that is, $D(\mathcal{X}) = \max_{\mathbf{f}, \mathbf{g} \in \mathcal{X}} d(\mathbf{f}, \mathbf{g})$. More specifically, suppose that the dimension of the high-resolution image is $M_1 \times M_2$ and the dynamic range of the image’s gray-scale levels is $[\theta_{min}, \theta_{max}]$, then $D(\mathcal{X})$ is

$$D(\mathcal{X}) = M_1 M_2 (\theta_{max} - \theta_{min}). \quad (3.25)$$

The key point of (3.24) is that any chain of image samples (i.e., $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \dots$ or $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots$ as mentioned above) will converge towards to each other along the way, since their distance decreases after each sample-generation step (see (3.22)). Furthermore, to justify (3.24), a mathematical analysis is provided in Section 3.6 to prove that the burn-in parameter T computed via (3.24) ensures the proposed MCMC process converging after T sample-generation steps.

According to (3.24), the determination of the burn-in parameter T boils down to the determination of the parameter c in (3.24) for the proposed MCMC process.

3.3 Derivation of The Burn-in Parameter of the Proposed MCMC Super-resolution Process

Firstly, the detailed mathematical derivation of the parameter c is documented in Section 3.6, and its result is provided below

$$c = \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda_{max} \gamma}{\frac{\rho H_{min}^2}{\sigma_v^2} + \lambda_{max} \gamma} \right), \quad (3.26)$$

where $M_1 \times M_2$ is the dimension of the high-resolution image, λ_{max} represents the permissible largest value of λ (see (3.13)), and γ are the parameters used in the prior image model, ρ is the number of low-resolution observations, σ_v^2 is the largest value of the variances (i.e., various σ_k^2) of the additive noise in the image acquisition process, and H_{min} represents the minimum absolute value of all non-zero entries of the matrices $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \dots, \mathbf{H}^{(\rho)}$, which are defined in (3.1), respectively. Finally, substituting (3.25) and (3.26) into (3.24), the burn-in parameter T of the proposed MCMC process can be obtained as

$$T = \frac{\ln \left(\frac{\epsilon}{M_1 M_2 (\theta_{max} - \theta_{min})} \right)}{\ln \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda_{max} \gamma}{\frac{\rho H_{min}^2}{\sigma_v^2} + \lambda_{max} \gamma} \right)}. \quad (3.27)$$

From the viewpoint of physical meaning, it is intuitive to observe that the burn-in parameter T will become large in the following three cases: i) the desired high-resolution image is expected to be very smooth (equivalently, λ_{max} in (3.27) is large); ii) the size of the desired high-resolution image is large (equivalently, $M_1 M_2$ in (3.27) is large); and iii) the dynamic range of the gray-scale levels of the desired high-resolution image is large (equivalently, $(\theta_{max} - \theta_{min})$ in (3.27) is large).

To evaluate the performance of the proposed closed-form bound (3.27) for the burn-in parameter T , experiments are conducted to compare the burn-in parameter T computed via (3.27) with the number of burn-in samples (iterations) obtained in our proposed stochastic simulations. One hundred experiments are conducted on the test images used in Section 3.2.4. In each simulation, the proposed MCMC method is applied to generate two chains of image samples, independently, where

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

each chain begins with a randomly-initialized image sample. Next, the number of burn-in samples (iterations), which is required for the two chains converging towards to each other and falling within the user-defined tolerance ϵ , is recorded. In our simulations, the dynamic range of the gray-scale levels of the desired high-resolution image is normalized to the range of $[0, 1]$; that is, θ_{min} and θ_{max} are set to be 0 and 1, respectively. The user-defined tolerance ϵ is set to be 10, and the permissible dynamic range of prior image model's hyperparameter value is set to be integer values in the range of $[1, 10]$. Figure 3.7 demonstrates the above-mentioned comparison, from which one can observe that the burn-in parameter T computed via (3.27) is always bounded above the experimental results.

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

Like any stochastic simulation task, the proposed MCMC approach also requires a “sufficiently” large number of image samples generated for estimating the unknown high-resolution image; otherwise, in the context of our work, image artifacts could be yielded in the reconstructed high-resolution image. To address this concern, the *variation-sensitive* bilateral filtering, which is a modified version of bilateral filtering [23], is proposed as a post-processing scheme to effectively reduce the above-mentioned artifacts. The proposed method contains two sequential stages: (i) detection and (ii) filtering, as follows.

3.4.1 Detection

In this stage, a binary decision is conducted at each pixel location to determine whether the estimated pixel intensity value using (3.15) is reliable or not. Since

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

(a)

(b)

Figure 3.7. A comparison of the burn-in parameter T computed via the proposed closed-form bound (3.27) versus the number of burn-in samples (iterations) obtained in simulations: (a) *Boat* image; (b) *Text* image.

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

the *mean* of $(N - T)$ reliable image samples is used as the reconstructed high-resolution image (see (3.15)), the *variance* of these samples is exploited as the decision criterion. If the value of the calculated variance is larger than a user-defined tolerance τ , the pixel position under consideration will be flagged accordingly to signify that the number of samples generated at this pixel position is insufficient; hence, the estimated pixel intensity value at this position is considered *unreliable* and should be replaced by a post-filtering value.

3.4.2 Filtering

The bilateral filter [23] performs low-pass filtering by replacing the intensity value of the pixel under consideration with a weighted average of intensity values of the pixels from its neighborhood. The weighting factor is a function of both the spatial distance measured with respect to the center pixel under consideration (called the *domain filter*) and the pixel intensity difference (called the *range filter*). More specifically, let $\mathbf{u}$ be the location of the pixel under consideration, and let Ω be a square region of support consisting of the pixels in a neighborhood of $\mathbf{u}$. The bilateral filter replaces the intensity value at $\mathbf{u}$ (denoted as $f(\mathbf{u})$) by

$$h(\mathbf{u}) = \frac{\sum_{\mathbf{v} \in \Omega} f(\mathbf{v}) c(\mathbf{v}, \mathbf{u}) s(f(\mathbf{v}), f(\mathbf{u}))}{\sum_{\mathbf{v} \in \Omega} c(\mathbf{v}, \mathbf{u}) s(f(\mathbf{v}), f(\mathbf{u}))}, \quad (3.28)$$

where $f(\mathbf{v})$ is the intensity value of the pixel $\mathbf{v}$. The domain filter $c(\mathbf{v}, \mathbf{u})$ and the range filter $s(f(\mathbf{v}), f(\mathbf{u}))$ are commonly modeled using Gaussian functions as

$$c(\mathbf{v}, \mathbf{u}) = \exp \left\{ -\frac{|\mathbf{v} - \mathbf{u}|^2}{2\sigma_d^2} \right\}, \quad (3.29)$$

$$s(f(\mathbf{v}), f(\mathbf{u})) = \exp \left\{ -\frac{|f(\mathbf{v}) - f(\mathbf{u})|^2}{2\sigma_r^2} \right\}, \quad (3.30)$$

where the parameters σ_d^2 and σ_r^2 are the variances of the respective functions.

Intuitively, pixels that are far away from $\mathbf{u}$ should have lower influence on $\mathbf{u}$,

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

hence the output of the domain filter $c(\mathbf{v}, \mathbf{u})$ decreases as the spatial distance between the pixels $\mathbf{u}$ and $\mathbf{v}$ increases, and consequently, the blurring phenomenon can be effectively reduced. Another intuition is that the pixels with significantly different intensity values with that of $\mathbf{u}$ (e.g., edges) should carry lower influence on $\mathbf{u}$, hence the output of the range filter $s(f(\mathbf{v}), f(\mathbf{u}))$ decreases as the intensity value difference between the pixels $\mathbf{u}$ and $\mathbf{v}$ increases; hence, edges can be better kept. In this light, one can see that for those pixel positions contributing image artifacts, the intensity values are usually either much larger or much smaller than that of their respective spatial neighboring pixels. Consequently, the conventional bilateral filter fails to remove image artifacts from the image, since it will assign much smaller weighting factors (i.e., the output of the range filter) to the pixels in its neighborhood [152]. To overcome this problem, the idea is to replace any unreliable reference pixel $f(\mathbf{u})$ used in the range filter by a reliable estimator; for that, the *median* value (denoted as $g(\mathbf{u})$) computed based on the neighborhood of the pixel $\mathbf{u}$ is used. This leads to the modified bilateral filter as

$$h(\mathbf{u}) = \frac{\sum_{\mathbf{v} \in \Omega} f(\mathbf{v}) c(\mathbf{v}, \mathbf{u}) s(f(\mathbf{v}), g(\mathbf{u}))}{\sum_{\mathbf{v} \in \Omega} c(\mathbf{v}, \mathbf{u}) s(f(\mathbf{v}), g(\mathbf{u}))}, \quad (3.31)$$

where $g(\mathbf{u})$ is determined by

$$g(\mathbf{u}) = \text{median} \{f(\mathbf{v})\}_{\mathbf{v} \in \Omega}. \quad (3.32)$$

Based on the binary decision obtained in the detection stage, a bilateral filter with the incorporation of a median pre-processing (i.e., (3.31)) will be performed to those pixel positions where the pixel sample values generated have high variance values, while the conventional bilateral filter (3.28) will be applied at the remaining pixel positions.

To compare the proposed variation-sensitive bilateral filter with the conventional bilateral filter [23], they are respectively applied to the high-resolution image that

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

Table 3.5
THE PSNR PERFORMANCE COMPARISON.

Test image	The number of image samples generated for the proposed MCMC SR approach	Proposed MCMC SR approach	Proposed MCMC SR approach + conventional bilateral filter [23]	Proposed MCMC SR approach + proposed variation-sensitive bilateral filter
<i>Boat</i>	$N = 3 \times T$	25.32 dB	26.27 dB	26.92 dB
	$N = 6 \times T$	27.10 dB	27.35 dB	27.76 dB
	$N = 10 \times T$	28.02 dB	28.12 dB	28.21 dB
<i>Text</i>	$N = 3 \times T$	17.88 dB	18.81 dB	19.43 dB
	$N = 6 \times T$	19.44 dB	19.80 dB	20.14 dB
	$N = 10 \times T$	20.17 dB	20.34 dB	20.40 dB

is estimated by the proposed MCMC algorithm developed in Section 3.2.3. For conducting the simulation experiments, the following parameters are set: $\sigma_d = 4$, $\sigma_r = 20$, $\Omega = 7 \times 7$ and $\tau = 1000$. To demonstrate how image artifacts appeared and then being reduced by these two post-processing methods under comparison, three different total numbers of image samples are individually set as $N = 3 \times T$, $N = 6 \times T$ and $N = 10 \times T$ (where T is defined as in Section 3.3). Their resulted PSNR performance measurements are presented in Table 3.6. For the subjective performance evaluation, Figures 3.8, 3.9 present the reconstructed high-resolution image for the cases $N = 3 \times T$ and $N = 10 \times T$. As can be seen from Figures 3.8, 3.9, the proposed variation-sensitive bilateral filter constantly yields superior performance to that of the conventional bilateral filter, particularly in the case when a smaller number of image samples are generated for the MCMC process.

Note that the proposed variation-sensitive bilateral filter adaptively switches between the conventional bilateral filter (3.28) and its median-value-substituted version (3.31). In fact, “no filtering” can be exploited to replace the former for further reducing the processing time. Experimental results show that the proposed post-processing method still yields superior performance (in terms of both PSNR and visual quality) to that of the conventional bilateral filter.

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

Table 3.6

THE PSNR PERFORMANCE COMPARISON OF THE PROPOSED OUTLIER-SENSITIVE BILATERAL FILTERING WITH THE INCORPORATION OF THE MEAN FILTER DEFINED IN (3.33) AND THAT WITH THE INCORPORATION OF THE MEDIAN FILTER DEFINED IN (3.32).

Test image	The number of image samples generated for the proposed MCMC SR approach	The proposed variation-sensitive bilateral filtering with the incorporation of the mean filter defined in (3.33)	The proposed variation-sensitive bilateral filtering with the incorporation of the median filter defined in (3.32)
<i>Boat</i>	$N = 3 \times T$	26.52 dB	26.92 dB
	$N = 6 \times T$	27.56 dB	27.76 dB
	$N = 10 \times T$	28.17 dB	28.21 dB
<i>Text</i>	$N = 3 \times T$	18.80 dB	19.43 dB
	$N = 6 \times T$	19.84 dB	20.14 dB
	$N = 10 \times T$	20.04 dB	20.40 dB

The second experiment is to compare the performance of the proposed variation-sensitive bilateral filter with the incorporation of the median filter (i.e., (3.32)) and that with the incorporation of the mean filter as follows. In this experiment, the median filter defined in (3.32) is replaced by a mean filter defined as

$$g(\mathbf{u}) = \text{mean} \{f(\mathbf{v})\}_{\mathbf{v} \in \Omega}. \quad (3.33)$$

Then the above mean filter (3.33) is further incorporated into the proposed outlier-sensitive bilateral filter to develop a new filter. The above two filters are respectively applied to the high-resolution image that is estimated by the proposed MCMC algorithm developed in Section 3.2.3. The following parameters are set: $\sigma_d = 4$, $\sigma_r = 20$, $\Omega = 7 \times 7$ and $\tau = 1000$. Table 3.6 presents the PSNR performance of the above two filters for three cases $N = 3 \times T$, $N = 6 \times T$ and $N = 10 \times T$ (where T is defined as in Section 3.3). As seen from Table 3.6, the proposed variation-sensitive bilateral filter with the incorporation of the median filter (defined in (3.32)) always outperforms that with the incorporation of the mean filter (defined in (3.33)).

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

Figure 3.8. Various reconstructed results of *Boat* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation, followed by image deblurring approach (PSNR = 25.80 dB); (d) the proposed MCMC algorithm using $N = 3 \times T$ image samples (PSNR = 25.32 dB); (e) post-processing result (PSNR = 26.27 dB) by applying the bilateral filter [23] on the picture (d); (f) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 26.92 dB); (g) the proposed MCMC algorithm using $N = 10 \times T$ image samples (PSNR = 28.02 dB); (h) post-processing result (PSNR = 28.12 dB) by applying the bilateral filter [23] on the picture (g); and (i) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 28.21 dB).

3.4 Proposed Post-Processing: Variation-Sensitive Bilateral Filter

Figure 3.9. Various reconstructed results of *Text* image: (a) the original image; (b) the simulated quarter-sized low-resolution image; (c) the bi-cubic spline interpolation, followed by image deblurring approach (PSNR = 17.62 dB); (d) the proposed MCMC algorithm using $N = 3 \times T$ image samples (PSNR = 17.88 dB); (e) post-processing result (PSNR = 18.81 dB) by applying the bilateral filter [23] on the picture (d); (f) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 19.43 dB); (g) the proposed MCMC algorithm using $N = 10 \times T$ image samples (PSNR = 20.17 dB); (h) post-processing result (PSNR = 20.34 dB) by applying the bilateral filter [23] on the picture (g); and (i) the proposed MCMC algorithm with the incorporation of the developed variation-sensitive bilateral filtering (PSNR = 20.40 dB).

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

3.5.1 Introduction

In the previous sections, the *Markov chain Monte Carlo* (MCMC) technique has been developed for addressing the SR problem with attractive performance. The developed algorithm provides a flexible framework for the incorporation of the prior knowledge of the unknown high-resolution image, as well as for the modeling of the dependence among the observed low-resolution images and the unknown high-resolution image.

However, the MCMC SR method developed in [24] usually generates a fairly smooth high-resolution image, due to the fact that the prior image model used for the MCMC SR approach only reflects the expectation that the image is locally smooth. To overcome the above problem, our aim is to incorporate an edge enhancement term into the prior image model used for the MCMC SR framework for adaptively enhancing the edge information according to the local statistics of the image. For that, Charbonnier *et al.* proposed to impose an edge-preserving regularization term on the reconstructed image [153]. Vanzella *et al.* [154] proposed to adaptively adjust the smoothing regularization parameter according to the local statistics of the image. In [155], Mignotte proposed a segmentation-based regularization method by exploiting various regularization terms based on the segmentation map of the image. Choi *et al.* [156] proposed a region-based SR approach by segmenting the image into smooth, intermediate, and edge regions, then applying various diffusion-based regularization terms and regularization parameters

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

accordingly. Furthermore, for regularization in edge regions in particular, a directional smoothing based on adaptive anisotropic diffusion, is applied to restore the continuity and sharpness of edges.

In this chapter, to further *enhance* the edge information in the reconstructed high-resolution image, an edge enhancement component is further incorporated into our research works developed in [24] to propose an edge-directed MCMC SR approach. The proposed approach exploits an edge map of the desired high-resolution image to adaptively enhance the edge information presented in each image sample, which is generated during the proposed MCMC SR process, and intimately enhance the quality of the reconstructed high-resolution image.

3.5.2 Proposed Edge-directed MCMC SR Approach

The visual appearance of the image can be significantly improved by emphasizing its high-frequency contents to enhance the edge and detail information inside. To achieve the above goal, it is necessary to raise the gray level of the brighter side of an edge while lowering the gray level of the darker side. For that, a linear high-pass filtered version of the image to be enhanced (denoted as $\mathbf{X}$) is first multiplied with a positive scaling factor β and then added to the input image [157] to obtain a filtered image result (denoted as $\mathbf{X}'$) with enhanced edge information as

$$\mathbf{X}' = \mathbf{X} + \beta \mathbf{g} \mathbf{X} = \beta \left(\frac{1}{\beta} \mathbf{I} + \mathbf{g} \right) \mathbf{X} = \beta \mathbf{G} \mathbf{X}, \quad (3.34)$$

where the matrix $\mathbf{g}$ represents the linear high-pass filter, whose entries are given by $g_{i,j} = 1$, if $i = j$; $g_{i,j} = -\frac{1}{\gamma}$, if i and j both fall in the γ -neighbourhood; otherwise, $g_{i,j} = 0$. In addition, β is the positive scaling factor that controls the degree of enhancement, $\mathbf{I}$ is an $M_1 M_2 \times M_1 M_2$ identify matrix, $\mathbf{G}$ is an $M_1 M_2 \times M_1 M_2$

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

matrix, whose entries are given by

$$G_{i,j} = \begin{cases} 1 + \frac{1}{\beta}, & i = j; \\ -\frac{1}{\gamma}, & i \text{ and } j \text{ are adjacent in the } \gamma\text{-neighbourhood}; \\ 0, & \text{otherwise.} \end{cases} \quad (3.35)$$

The above process is illustrated in Figure 3.10.

Next, according to (3.11) and (3.34), the prior image model on the high-resolution image $p(\mathbf{X})$ with the expectation that the edge information should be enhanced at the edge pixel positions while the image is locally smooth at the rest pixel positions can be expressed as

$$p(\mathbf{X}) \propto \exp \left\{ -\frac{1}{2} (\lambda \mathbf{X}^T \mathbf{C} \mathbf{X} + \beta \|\mathbf{G}(\mathbf{E} * \mathbf{X})\|_1) \right\}, \quad (3.36)$$

where $*$ is the element-by-element product function, $\|\cdot\|_1$ is the L_1 norm function, and the $M \times 1$ vector $\mathbf{E}$ is the edge map of the high-resolution image $\mathbf{X}$; that is, the entries of $\mathbf{E}$ are given by

$$E_n = \begin{cases} 1, & \text{if } X_n \text{ is determined as the edge pixel;} \\ 0, & \text{otherwise.} \end{cases} \quad (3.37)$$

Finally, substituting (3.7) and (3.36) into (3.4), the posterior distribution of the unknown high-resolution image given the low-resolution observations $p(\mathbf{X}|\mathbf{Y})$ can be obtained as

$$p(\mathbf{X}|\mathbf{Y}) \propto \exp \left\{ -\sum_{k=1}^P \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)} \mathbf{X}\|^2 - \frac{1}{2} (\lambda \mathbf{X}^T \mathbf{C} \mathbf{X} + \beta \|\mathbf{G}(\mathbf{E} * \mathbf{X})\|_1) \right\}. \quad (3.38)$$

Comparing (3.14) and (3.38), one can explicitly see that the novelty of the proposed edge-directed MCMC SR method is that it can adaptively enhance the edge information at the edge pixel positions (determined by the edge map matrix $\mathbf{E}$) while exploiting the conventional MCMC SR [24] at the rest pixel positions.

The goal of the proposed edge-directed MCMC SR algorithm is to estimate the high-resolution image by applying the MCMC technique to generate N samples

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

from $p(\mathbf{X}|\mathbf{Y})$ (i.e., (3.38)); these samples are denoted as $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(N)}$ (each of them has the same size as that of $\mathbf{X}$), where N is large enough to guarantee the convergence of the MCMC SR process. Then, the image samples generated up to the sample number T are considered unreliable and thus should be discarded. Finally, each component (say, the n -th) of the high-resolution image (denoted as $\hat{X}_n$) can be estimated via the mean of the respective n -th components of the rest $(N - T)$ image samples; that is,

$$\hat{X}_n = \frac{1}{N - T} \sum_{i=T+1}^N X_n^{(i)}, \quad (3.39)$$

where the parameter T was derived in [24] as

$$T = \frac{\ln \left(\frac{\epsilon}{M_1 M_2 (\theta_{max} - \theta_{min})} \right)}{\ln \left(1 - \frac{1}{M_1 M_2} \frac{\rho H_{min}^2}{\rho H_{min}^2 + \max(\lambda, \beta) \gamma \sigma_v^2} \right)}, \quad (3.40)$$

in which ϵ is the user-defined tolerance, θ_{min} and θ_{max} are the minimum and maximum gray values, respectively. $M_1 \times M_2$ is the dimension of the high-resolution image, λ , β and γ are the parameters used in the prior image model, ρ is the number of low-resolution observations, $\max(\lambda, \beta)$ represents the maximum value between λ and β , σ_v^2 is the largest values of the variances (i.e., various σ_k^2) of the additive Gaussian white noise and H_{min} represents the minimum value of all non-zero entries of the matrices $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \dots, \mathbf{H}^{(\rho)}$.

The proposed edge-directed MCMC SR process starts from randomly generating the first image sample for initialization. Next, at each sample generation step, the Gibbs sampler is applied to update one single component of the image sample, according to its conditional distribution given the remaining components and the observations. The above-mentioned processing steps will be repeated to generate N samples. Let X_n denote the n -th component of $M \times 1$ vector $\mathbf{X}$, and $\mathbf{X}_{-n}$ denote the vector of $\mathbf{X}$ with the n -th component being omitted; that is, $\mathbf{X}_{-n} = \{X_1, X_2, \dots, X_{n-1}, X_{n+1}, \dots, X_{M_1 M_2}\}$. According to (3.38), the distribution

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

function of X_n conditional on $\mathbf{X}_{-n}$ and $\mathbf{Y}$ can be expressed as

$$p(X_n | \mathbf{X}_{-n}, \mathbf{Y}) \propto \exp \left\{ -\frac{1}{2} \tau_{X_n}^{-1} (X_n - \mu_{X_n})^2 \right\}, \quad (3.41)$$

where

$$\mu_{X_n} = \frac{\tau_{X_n}}{2} \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(X_n) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(X_n) \right), \quad (3.42)$$

$$\tau_{X_n} = \frac{1}{\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{(H_{i,n}^{(k)})^2}{\sigma_k^2} + \lambda C_{n,n}}, \quad (3.43)$$

in which $\psi_i^{(k)}(X_n)$ and $\phi(X_n)$ are defined as follows, respectively,

$$\psi_i^{(k)}(X_n) = y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} X_j, \quad (3.44)$$

$$\phi(X_n) = \lambda \sum_{s \neq n} X_s C_{s,n} + \lambda \sum_{t \neq n} C_{n,t} X_t + \beta \sum_{l=1}^{M_1 M_2} G_{l,n} E_n. \quad (3.45)$$

The detailed proof of (3.41) is provided in Section 3.6.

To summarize, the proposed edge-directed MCMC SR process can be stated as below.

- Randomly generate $\mathbf{X}^{(1)}$ for initialization.
- For $t = 1 : N - 1$,
 - Update the n -th component of $\mathbf{X}^{(t+1)}$ by sampling $X_n^{(t+1)} \sim p(X_n^{(t+1)} | \mathbf{X}_{-n}^{(t)}, \mathbf{Y})$ (defined as in (3.41)), and n is randomly selected; that is, n is drawn from the discrete uniform distribution over the interval $[1, M_1 \times M_2]$.
 - The remaining components of $\mathbf{X}^{(t+1)}$ are the same as that of $\mathbf{X}^{(t)}$; that is, $\mathbf{X}^{(t+1)} = \left\{ X_1^{(t)}, X_2^{(t)}, \dots, X_{n-1}^{(t)}, X_n^{(t+1)}, X_{n+1}^{(t)}, \dots, X_{M_1 M_2}^{(t)} \right\}$.
- Compute $\hat{\mathbf{X}}$ via (3.39).

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

Figure 3.10. The image sample, which is generated at each sample generation step of the proposed MCMC SR approach, is enhanced by adding a positive scaling factor β that is multiplied with a linear high-pass filtered version of the image sample.

3.5.3 Experimental Results

Three test images are used in the experiments, as demonstrated in Figure 3.11. In our simulation, the low-resolution images are generated from the original image by applying a series of operations, as follows. We first apply a shift, which is drawn from a continuous uniform distribution over the interval $(-2, 2)$ in the units of pixels, followed by applying a rotation, which is drawn from a continuous uniform distribution over the interval $[-4, 4]$ in the units of degrees. Next, the processed images are convolved with a point spread function, which is a Gaussian low-pass filter with a size of 5×5 and the standard deviation of 1, followed by a down-sampling operation in both the horizontal and the vertical directions, respectively. Lastly, the processed image is added with a zero-mean white Gaussian noise to yield a 20 dB *signal-to-noise ratio* (SNR) image.

The proposed SR algorithm produces the high-resolution image as follows. Firstly, the initial image sample is randomly generated. The convergence bound T is computed via (3.40). In our simulation, the parameter ϵ is set to 10, γ is set to 4; that is, we apply 4-neighborhood MRF model. In addition, the parameter λ and β are experimentally selected to be 20 and 4, respectively. The MCMC technique is then applied to generate N samples, where the parameter N is defined by the user and

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

Figure 3.11. Three test images used in the experiments: (a) 256×256 Airplane; (b) 256×256 Boat; and (c) 256×256 Bridge.

it should be large enough to meet the statistics of the target distribution. In our simulation, N is selected to be $10 \times T$. Finally, the first T samples are removed, and the remaining $9 \times T$ samples are used to produce the high-resolution image via (3.39). On the other hand, the edge map of the high-resolution image can be computed by two ways: (i) applying the *Canny* edge detector [158] on the original ground truth high-resolution image; (ii) applying the *Canny* edge detector on the observed low-resolution image, followed by applying the bi-cubic interpolation technique on the low-resolution edge map.

Experiments are conducted to compare the performance of the proposed edge-directed MCMC SR algorithm with that of the *bi-cubic interpolation* approach and the SR algorithm developed in [20], plus that of the MCMC SR method developed in [24]. Figure 3.12 presents the results obtained from the test images *Airplane*. Figure 3.12 (e) and (f) present the results of the proposed edge-directed MCMC SR approach using the edge map estimated from the observed low-resolution image and that obtained from the ground truth high-resolution image, respectively. Furthermore, part of the above six images (i.e., Figure 3.12 (a)-(f)) are demonstrated in Figure 3.12 (g)-(l), respectively. Table 3.7 compares the PSNR performance of the above methods. As seen from Figure 3.12 and Table 3.7, the proposed method can effectively adaptively enhance the edge information of the reconstructed high-resolution image for yielding superior visual quality.

3.5 Edge-directed Super-resolution Image Reconstruction Using A Markov Chain Monte Carlo Approach

Figure 3.12. Various results of *Airplane* image (2×2 enlargement): (a) the original image; (b) the bi-cubic spline interpolation approach (PSNR = 25.69 dB); (c) Vandewalle *et al.*'s algorithm developed in [20] (PSNR = 27.49 dB); (d) the MCMC SR approach developed in [24] (PSNR = 28.81 dB); (e) the proposed edge-directed MCMC SR approach using the edge map computed from the observed low-resolution image ($\beta = 4$, PSNR = 29.09 dB); (f) the proposed edge-directed MCMC SR approach using the edge map computed from the ground truth high-resolution image ($\beta = 4$, PSNR = 29.53 dB); and images (g)-(l) are the part of the images of (a)-(f), respectively.

Table 3.7

THE PSNR PERFORMANCE COMPARISON. PROPOSED METHOD I REPRESENTS THE PROPOSED EDGE-DIRECTED MCMC SR USING THE EDGE MAP COMPUTED FROM THE LOW-RESOLUTION OBSERVATIONS, WHILE THE PROPOSED METHOD II REPRESENTS THE PROPOSED EDGE-DIRECTED MCMC SR USING THE EDGE MAP COMPUTED FROM THE ORIGINAL HIGH-RESOLUTION IMAGE.

Test image	Enlarge ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Tian and Ma's method [24]	Proposed method I	Proposed method II
<i>Airplane</i>	2×2	25.69 dB	27.49 dB	28.81 dB	29.09 dB	29.53 dB
<i>Boat</i>		24.55 dB	25.97 dB	26.89 dB	27.00 dB	27.40 dB
<i>Bridge</i>		23.01 dB	23.21 dB	25.18 dB	25.45 dB	25.68 dB
<i>Airplane</i>	4×4	20.48 dB	22.92 dB	26.02 dB	26.45 dB	27.07 dB
<i>Boat</i>		20.21 dB	23.57 dB	24.78 dB	25.04 dB	25.27 dB
<i>Bridge</i>		19.19 dB	22.10 dB	22.93 dB	23.25 dB	23.52 dB

3.6 Conclusions

In this chapter, a stochastic SR image approach is proposed by exploiting the Bayesian inference and the MCMC technique; it has achieved attractive image reconstruction results. The high-resolution image is produced based on the reliable image samples generated by using the Gibbs sampler according to the distribution of the joint posterior density of the unknown high-resolution image and the prior image model's hyperparameter. Furthermore, the prior image model's hyperparameter is adapted simultaneously at each sample-generation step using the Metropolis-Hastings algorithm throughout the proposed MCMC process. Extensive experimental results presented clearly show that the proposed joint adaptation scheme yields a fairly close performance to that of the optimally-determined hyperparameter value via exhaustive search. Moreover, a bound of the length of the burn-in period is mathematically derived for accurately identifying reliable samples. In addition, the *variation-sensitive* bilateral filtering is proposed as a post-processing method to effectively reduce the image artifacts presented in the

Appendix 3.A Derivation of (3.16)

reconstructed high-resolution image, when the number of image samples generated for the MCMC process is insufficient. The proposed *variation-sensitive* bilateral filtering constantly yields superior performance to that of the conventional bilateral filtering on removing artifacts and preserving image details, as evidenced by our experimental results. On the other hand, the reconstructed high-resolution image of the MCMC SR approach are usually too smooth to lose much detail information. In view of this, we improve this aspect by developing an edge-directed MCMC SR approach. The proposed method can effectively emphasizing the high-frequency contents of the reconstructed high-resolution image to enhance its edge and detail information, as verified by the experimental results.

Appendix 3.A Derivation of (3.16)

Let X_n denote the n -th component of the lexicographic-order $M_1 M_2 \times 1$ vector $\mathbf{X}$, and $\mathbf{X}_{-n}$ denote the same vector of $\mathbf{X}$ except with the n -th component being excluded; that is, $\mathbf{X}_{-n} = \{X_1, X_2, \dots, X_{n-1}, X_{n+1}, \dots, X_{M_1 M_2}\}$. According to (3.14), the probability density function of X_n , conditional on $\mathbf{X}_{-n}$, λ , and $\mathbf{Y}$, can be expressed as

$$\begin{aligned}
 & p(X_n | \mathbf{X}_{-n}, \lambda, \mathbf{Y}) \\
 & \propto \frac{1}{Z(\lambda)} \exp \left\{ - \sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)} \mathbf{X}\|^2 - \frac{1}{2} \lambda \mathbf{X}^T \mathbf{C} \mathbf{X} \right\} \frac{1}{\lambda_{max} - \lambda_{min}} \\
 & \propto \frac{1}{Z(\lambda)} \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j=1}^{M_1 M_2} H_{i,j}^{(k)} X_j \right)^2 - \frac{1}{2} \lambda \sum_{s=1}^{M_1 M_2} \sum_{t=1}^{M_1 M_2} X_s C_{s,t} X_t \right\} \frac{1}{\lambda_{max} - \lambda_{min}}.
 \end{aligned} \tag{3.46}$$

Appendix 3.A Derivation of (3.16)

Then, (3.46) can be rewritten as

$$\begin{aligned}
 & p(X_n | \mathbf{X}_{-n}, \lambda, \mathbf{Y}) \\
 & \propto \frac{1}{Z(\lambda)} \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j=1}^{M_1 M_2} H_{i,j}^{(k)} X_j \right)^2 - \frac{1}{2} \lambda \sum_{s=1}^{M_1 M_2} \sum_{t=1}^{M_1 M_2} X_s C_{s,t} X_t \right\} \frac{1}{\lambda_{\max} - \lambda_{\min}} \\
 & = \frac{1}{Z(\lambda)} \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j=1}^{n-1} H_{i,j}^{(k)} X_j - H_{i,n}^{(k)} X_n - \sum_{j=n+1}^{M_1 M_2} H_{i,j}^{(k)} X_j \right)^2 \right\} \\
 & \quad \cdot \exp \left\{ - \frac{1}{2} \lambda \sum_{s \neq n} X_s C_{sn} X_n - \frac{1}{2} \lambda \sum_{t \neq n} X_n C_{nt} X_t - \frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t - \frac{1}{2} \lambda X_n C_{nn} X_n \right\} \frac{1}{\lambda_{\max} - \lambda_{\min}} \\
 & = \frac{1}{Z(\lambda)} \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} X_j - H_{i,n}^{(k)} X_n \right)^2 \right\} \\
 & \quad \cdot \exp \left\{ - \frac{1}{2} \lambda \left(\sum_{s \neq n} X_s C_{s,n} + \sum_{t \neq n} C_{n,t} X_t \right) X_n - \frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\} \frac{1}{\lambda_{\max} - \lambda_{\min}} \\
 & = \frac{1}{Z(\lambda)} \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) - H_{i,n}^{(k)} X_n \right)^2 \right\} \\
 & \quad \cdot \exp \left\{ - \frac{1}{2} \lambda \phi(X_n) X_n - \frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\} \frac{1}{\lambda_{\max} - \lambda_{\min}}, \tag{3.47}
 \end{aligned}$$

where

$$\psi_i^{(k)}(X_n) = y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} X_j, \tag{3.48}$$

$$\phi(X_n) = \sum_{s \neq n} X_s C_{s,n} + \sum_{t \neq n} C_{n,t} X_t. \tag{3.49}$$

Next, (3.47) can be simplified by removing the terms $\frac{1}{Z(\lambda)}$, $-\frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t$, and $\frac{1}{\lambda_{\max} - \lambda_{\min}}$ from (3.47), since these terms are independent of X_n . Therefore, (3.47) can be simplified as

$$p(X_n | \mathbf{X}_{-n}, \lambda, \mathbf{Y}) \propto \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) - H_{i,n}^{(k)} X_n \right)^2 - \frac{1}{2} \lambda \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\}, \tag{3.50}$$

where $\psi_i^{(k)}(X_n)$ and $\phi(X_n)$ are defined as that in (3.48) and (3.49), respectively.

Appendix 3.B Proof of (3.24)

Furthermore, (3.50) can be rewritten as

$$\begin{aligned}
 & p(X_n | \mathbf{X}_{-n}, \lambda, \mathbf{Y}) \\
 & \propto \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) - H_{i,n}^{(k)} X_n \right)^2 - \frac{1}{2} \lambda \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\} \\
 & = \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) \right)^2 + 2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \psi_i^{(k)}(X_n) H_{i,n}^{(k)} X_n - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(H_{i,n}^{(k)} \right)^2 X_n^2 \right. \\
 & \quad \left. - \frac{1}{2} \lambda \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\} \\
 & = \exp \left\{ - \left(\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(H_{i,n}^{(k)} \right)^2 + \frac{1}{2} \lambda C_{n,n} \right) X_n^2 + \left(\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{\sigma_k^2} \psi_i^{(k)}(X_n) H_{i,n}^{(k)} - \frac{1}{2} \lambda \phi(X_n) \right) X_n \right. \\
 & \quad \left. - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) \right)^2 \right\} \\
 & \propto \exp \left\{ - \frac{1}{2\tau_{X_n}} (X_n - \mu_{X_n})^2 \right\}, \tag{3.51}
 \end{aligned}$$

where

$$\mu_{X_n} = \frac{\tau_{X_n}}{2} \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(X_n) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(X_n) \right), \tag{3.52}$$

$$\tau_{X_n} = \frac{1}{\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\left(H_{i,n}^{(k)} \right)^2}{\sigma_k^2} + \lambda C_{n,n}}. \tag{3.53}$$

Appendix 3.B Proof of (3.24)

The objective is to prove that the burn-in parameter T computed via (3.24) ensures that the distance between any chains of image samples is no larger than the user-defined tolerance ϵ after T sample-generation steps; that is,

$$d(\mathbf{f}^{(T+1)}, \mathbf{g}^{(T+1)}) \leq \epsilon. \tag{3.54}$$

Firstly, considering that the distance between the above-mentioned chains of image

Appendix 3.C Derivation of (3.26)

samples decreases c times after each sample-generation step (see (3.22)), we have

$$d(\mathbf{f}^{(T+1)}, \mathbf{g}^{(T+1)}) \leq c \cdot d(\mathbf{f}^{(T)}, \mathbf{g}^{(T)}) \leq c \cdot c \cdot d(\mathbf{f}^{(T-1)}, \mathbf{g}^{(T-1)}) \leq \dots \leq c^T \cdot d(\mathbf{f}^{(1)}, \mathbf{g}^{(1)}). \quad (3.55)$$

Note that $D(\mathcal{X})$ is the largest distance between any two image samples in the space $\mathcal{X}$; that is, $d(\mathbf{f}^{(1)}, \mathbf{g}^{(1)}) \leq D(\mathcal{X})$. Next, substituting $d(\mathbf{f}^{(1)}, \mathbf{g}^{(1)}) \leq D(\mathcal{X})$ into (3.55), we have

$$\begin{aligned} d(\mathbf{f}^{(T+1)}, \mathbf{g}^{(T+1)}) &\leq c^T \cdot d(\mathbf{f}^{(1)}, \mathbf{g}^{(1)}) \\ &\leq c^T \cdot D(\mathcal{X}). \end{aligned} \quad (3.56)$$

Finally, substituting (3.24) into (3.56), we have

$$\begin{aligned} d(\mathbf{f}^{(T+1)}, \mathbf{g}^{(T+1)}) &\leq c^T \cdot D(\mathcal{X}) = c^{\frac{\ln(\epsilon/D(\mathcal{X}))}{\ln c}} \cdot D(\mathcal{X}) \\ &= c^{\log_c(\epsilon/D(\mathcal{X}))} \cdot D(\mathcal{X}) \\ &= \frac{\epsilon}{D(\mathcal{X})} \cdot D(\mathcal{X}) \\ &= \epsilon. \end{aligned} \quad (3.57)$$

Appendix 3.C Derivation of (3.26)

Suppose two chains of image samples (denoted as $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \dots, \mathbf{f}^{(t)}, \dots$ and $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(t)}, \dots$, respectively) are simultaneously run for the proposed MCMC SR process. The objective is to find the parameter $c \in (0, 1)$, such that the following condition satisfies for any sample-generation step t ,

$$E[d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \leq c \cdot d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)}), \quad (3.58)$$

where $E[\cdot | \cdot]$ is the conditional expectation function and $d(\cdot)$ is the pixel-wise distance measurement between two image samples and defined as

$$d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)}) = \sum_{n=1}^{M_1 M_2} |f_n^{(t)} - g_n^{(t)}|, \quad (3.59)$$

Appendix 3.C Derivation of (3.26)

in which $M_1 \times M_2$ is the dimension of the vectors $\mathbf{f}^{(t)}$ and $\mathbf{g}^{(t)}$, $f_n^{(t)}$ and $g_n^{(t)}$ are the n -th components of the vectors $\mathbf{f}^{(t)}$ and $\mathbf{g}^{(t)}$, respectively.

Firstly, according to (3.59), we have

$$E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] = \sum_{n=1}^{M_1 M_2} E [|f_n^{(t+1)} - g_n^{(t+1)}| | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}]. \quad (3.60)$$

Now, we analyze (3.60) by considering the following two cases: (i) the components of $\mathbf{f}^{(t)}$ are greater than that of $\mathbf{g}^{(t)}$; and (ii) the components of $\mathbf{f}^{(t)}$ are not greater than that of $\mathbf{g}^{(t)}$. We provide our derivations of (3.26) for the first case as follows (the second case can be similarly derived). For the first case, we have

$$E [|f_n^{(t+1)} - g_n^{(t+1)}| | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] = E [(f_n^{(t+1)} - g_n^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}]. \quad (3.61)$$

Then, substituting (3.61) into (3.60), we get

$$E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] = \sum_{n=1}^{M_1 M_2} E [(f_n^{(t+1)} - g_n^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}]. \quad (3.62)$$

Since the proposed MCMC algorithm randomly selects only a single pixel position of the image sample, and updates its intensity value at each MCMC sample-generation step (see Section 3.2.3), (3.62) can be rewritten as

$$\begin{aligned} & E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \\ &= \sum_{n=1}^{M_1 M_2} E [(f_n^{(t+1)} - g_n^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \\ &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{1}{M_1 M_2} E (f_n^{(t+1)} - g_n^{(t+1)}) \right\} \\ &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{1}{M_1 M_2} (\mu_{f_n^{(t+1)}} - \mu_{g_n^{(t+1)}}) \right\}, \end{aligned} \quad (3.63)$$

where $\mu_{f_n^{(t+1)}}$ and $\mu_{g_n^{(t+1)}}$ are defined as that in (3.42).

Now, the objective is to simplify (3.63) and compare it with (3.58) to find the parameter c . Firstly, substituting (3.42) into (3.63) and considering that $\tau_{f_n^{(t+1)}} =$

Appendix 3.C Derivation of (3.26)

$\tau_{g_n^{(t+1)}}$ (see (3.18)), we have

$$\begin{aligned}
 & E \left[d \left(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)} \right) \mid \mathbf{f}^{(t)}, \mathbf{g}^{(t)} \right] \\
 &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{1}{M_1 M_2} \left(\frac{\tau_{f_n^{(t+1)}}}{2} \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(f_n^{(t+1)}) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(f_n^{(t+1)}) \right) \right. \right. \\
 &\quad \left. \left. - \frac{\tau_{g_n^{(t+1)}}}{2} \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(g_n^{(t+1)}) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(g_n^{(t+1)}) \right) \right) \right\} \\
 &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{\tau_{f_n^{(t+1)}}}{2 M_1 M_2} \left(\left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(f_n^{(t+1)}) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(f_n^{(t+1)}) \right) \right. \right. \\
 &\quad \left. \left. - \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(g_n^{(t+1)}) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(g_n^{(t+1)}) \right) \right) \right\} \\
 &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{\tau_{f_n^{(t+1)}}}{M_1 M_2} \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{H_{i,n}^{(k)} (\psi_i^{(k)}(f_n^{(t+1)}) - \psi_i^{(k)}(g_n^{(t+1)}))}{\sigma_k^2} \right. \\
 &\quad \left. - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} (\phi(f_n^{(t+1)}) - \phi(g_n^{(t+1)})) \right\}, \tag{3.64}
 \end{aligned}$$

where $\tau_{f_n^{(t+1)}}$ is defined in (3.18), $\psi_i^{(k)}(f_n^{(t+1)})$ and $\psi_i^{(k)}(g_n^{(t+1)})$ are defined in (3.19), $\phi(f_n^{(t+1)})$ and $\phi(g_n^{(t+1)})$ are defined in (3.20).

Appendix 3.C Derivation of (3.26)

Then, substituting (3.19) into (3.64), we get

$$\begin{aligned}
 & E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \\
 &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) \right. \\
 &\quad + \frac{\tau_{f_n^{(t+1)}}}{M_1 M_2} \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{H_{i,n}^{(k)} \left((y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} f_n^{(t)}) - (y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} g_n^{(t)}) \right)}{\sigma_k^2} \\
 &\quad \left. - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} (\phi(f_n^{(t+1)}) - \phi(g_n^{(t+1)})) \right\} \\
 &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{\tau_{f_n^{(t+1)}}}{M_1 M_2} \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{H_{i,n}^{(k)} \sum_{j \neq n} H_{i,j}^{(k)} (-f_n^{(t)} + g_n^{(t)})}{\sigma_k^2} \right. \\
 &\quad \left. - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} (\phi(f_n^{(t+1)}) - \phi(g_n^{(t+1)})) \right\} \\
 &= \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) - \frac{\tau_{f_n^{(t+1)}}}{M_1 M_2} \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{H_{i,n}^{(k)} \sum_{j \neq n} H_{i,j}^{(k)} (f_n^{(t)} - g_n^{(t)})}{\sigma_k^2} \right. \\
 &\quad \left. - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} (\phi(f_n^{(t+1)}) - \phi(g_n^{(t+1)})) \right\}. \tag{3.65}
 \end{aligned}$$

Note that the term $\frac{\tau_{f_n^{(t+1)}}}{M_1 M_2} \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{H_{i,n}^{(k)} \sum_{j \neq n} H_{i,j}^{(k)} (f_n^{(t)} - g_n^{(t)})}{\sigma_k^2}$ in (3.65) is non-negative, since the entries of the matrix $\mathbf{H}$ are non-negative and $\tau_{f_n^{(t+1)}}$ is positive (see (3.18)).

Omitting the above-mentioned non-negative term from (3.65), it can be simplified as

$$\begin{aligned}
 & E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \\
 &\leq \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} (\phi(f_n^{(t+1)}) - \phi(g_n^{(t+1)})) \right\}. \tag{3.66}
 \end{aligned}$$

Appendix 3.C Derivation of (3.26)

Next, substituting (3.20) into (3.66) and considering that the matrix $\mathbf{C}$ is symmetrical (see (3.12)), which means that $C_{s,n} = C_{n,s}$, we have

$$\begin{aligned}
 & E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \\
 & \leq \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} \left(\sum_{s \neq n} f_n^{(t)} C_{s,n} + \sum_{t \neq n} C_{n,t} f_n^{(t)} - \sum_{s \neq n} g_n^{(t)} C_{s,n} - \sum_{t \neq n} C_{n,t} g_n^{(t)} \right) \right\} \\
 & \leq \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) - \frac{\lambda \tau_{f_n^{(t+1)}}}{2 M_1 M_2} \left(2 \sum_{s \neq n} C_{s,n} f_n^{(t)} - 2 \sum_{s \neq n} C_{s,n} g_n^{(t)} \right) \right\} \\
 & = \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) - \frac{\lambda \tau_{f_n^{(t+1)}}}{M_1 M_2} \sum_{s \neq n} C_{s,n} (f_n^{(t)} - g_n^{(t)}) \right\}. \tag{3.67}
 \end{aligned}$$

Note that all the off-diagonal entries of the matrix $\mathbf{C}$ are negative and for all n , we have $\sum_{s \neq n} C_{s,n} \geq -\gamma$ (see (3.12)), which means that $-\sum_{s \neq n} C_{s,n} \leq \gamma$. Hence, substituting $-\sum_{s \neq n} C_{s,n} \leq \gamma$ into (3.67), we can obtain the following from (3.67)

$$\begin{aligned}
 E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] & \leq \sum_{n=1}^{M_1 M_2} \left\{ \frac{M_1 M_2 - 1}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) + \frac{\lambda \gamma \tau_{f_n^{(t+1)}}}{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) \right\}, \\
 & = \sum_{n=1}^{M_1 M_2} \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{\lambda \gamma \tau_{f_n^{(t+1)}}}{M_1 M_2} \right) (f_n^{(t)} - g_n^{(t)}). \tag{3.68}
 \end{aligned}$$

Then, substituting (3.18) into (3.68), we get

$$E [d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \leq \sum_{n=1}^{M_1 M_2} \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda \gamma}{\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{(H_{i,n}^{(k)})^2}{\sigma_k^2} + \lambda C_{n,n}} \right) (f_n^{(t)} - g_n^{(t)}). \tag{3.69}$$

Since the dynamic range of λ is $[\lambda_{\min}, \lambda_{\max}]$, we have $\lambda \leq \lambda_{\max}$. Denote $H_{\min}$ as the minimum absolute value of all non-zero entries of the matrices $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \dots, \mathbf{H}^{(\rho)}$, which are defined in (3.1), respectively. Further denote σ_v^2 as the largest value among different σ_k^2 . Therefore, we have $\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{(H_{i,n}^{(k)})^2}{\sigma_k^2} \geq \frac{\rho H_{\min}^2}{\sigma_v^2}$. Substituting $\lambda \leq \lambda_{\max}$ and $C_{n,n} = \gamma$ (see (3.12)), $\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{(H_{i,n}^{(k)})^2}{\sigma_k^2} \geq \frac{\rho H_{\min}^2}{\sigma_v^2}$ into (3.69), we

Appendix 3.D Derivation of (3.41)

have

$$\begin{aligned} E[d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] &\leq \sum_{n=1}^{M_1 M_2} \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda_{\max} \gamma}{\frac{\rho H_{\min}^2}{\sigma_v^2} + \lambda_{\max} \gamma} \right) (f_n^{(t)} - g_n^{(t)}) \\ &= \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda_{\max} \gamma}{\frac{\rho H_{\min}^2}{\sigma_v^2} + \lambda_{\max} \gamma} \right) \sum_{n=1}^{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) . \end{aligned} \quad (3.70)$$

According to (3.59), we have $\sum_{n=1}^{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) = d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)})$. Substituting $\sum_{n=1}^{M_1 M_2} (f_n^{(t)} - g_n^{(t)}) = d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)})$ into (3.70), we can obtain

$$E[d(\mathbf{f}^{(t+1)}, \mathbf{g}^{(t+1)}) | \mathbf{f}^{(t)}, \mathbf{g}^{(t)}] \leq \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda_{\max} \gamma}{\frac{\rho H_{\min}^2}{\sigma_v^2} + \lambda_{\max} \gamma} \right) \cdot d(\mathbf{f}^{(t)}, \mathbf{g}^{(t)}) . \quad (3.71)$$

Finally, comparing (3.58) and (3.71), the parameter c can be obtained as

$$c = \left(\frac{M_1 M_2 - 1}{M_1 M_2} + \frac{1}{M_1 M_2} \frac{\lambda_{\max} \gamma}{\frac{\rho H_{\min}^2}{\sigma_v^2} + \lambda_{\max} \gamma} \right) . \quad (3.72)$$

Appendix 3.D Derivation of (3.41)

Let X_n denote the n -th component of the lexicographic-order $M_1 M_2 \times 1$ vector $\mathbf{X}$, and $\mathbf{X}_{-n}$ denote the same vector of $\mathbf{X}$ except with the n -th component being excluded; that is, $\mathbf{X}_{-n} = \{X_1, X_2, \dots, X_{n-1}, X_{n+1}, \dots, X_{M_1 M_2}\}$. According to (3.38), the probability density function of X_n , conditional on $\mathbf{X}_{-n}$ and $\mathbf{Y}$, can be expressed as

$$\begin{aligned} p(X_n | \mathbf{X}_{-n}, \mathbf{Y}) &\propto \exp \left\{ - \sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|\mathbf{y}^{(k)} - \mathbf{H}^{(k)} \mathbf{X}\|^2 - \frac{1}{2} (\lambda \mathbf{X}^T \mathbf{C} \mathbf{X} + \beta \|\mathbf{G}(\mathbf{E} * \mathbf{X})\|_1) \right\} \\ &\propto \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j=1}^{M_1 M_2} H_{i,j}^{(k)} X_j \right)^2 \right. \\ &\quad \left. - \frac{1}{2} \lambda \sum_{s=1}^{M_1 M_2} \sum_{t=1}^{M_1 M_2} X_s C_{s,t} X_t - \frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} \sum_{m=1}^{M_1 M_2} G_{l,m} E_m X_m \right\} . \end{aligned} \quad (3.73)$$

Appendix 3.D Derivation of (3.41)

Then, (3.73) can be rewritten as

$$\begin{aligned}
& p(X_n | \mathbf{X}_{-n}, \mathbf{Y}) \\
& \propto \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j=1}^{M_1 M_2} H_{i,j}^{(k)} X_j \right)^2 \right. \\
& \quad \left. - \frac{1}{2} \lambda \sum_{s=1}^{M_1 M_2} \sum_{t=1}^{M_1 M_2} X_s C_{s,t} X_t - \frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} \sum_{m=1}^{M_1 M_2} G_{l,m} E_m X_m \right\} \\
& = \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j=1}^{n-1} H_{i,j}^{(k)} X_j - H_{i,n}^{(k)} X_n - \sum_{j=n+1}^{M_1 M_2} H_{i,j}^{(k)} X_j \right)^2 \right. \\
& \quad - \frac{1}{2} \lambda \sum_{s \neq n} X_s C_{sn} X_n - \frac{1}{2} \lambda \sum_{t \neq n} X_n C_{nt} X_t - \frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t - \frac{1}{2} \lambda X_n C_{nn} X_n \\
& \quad \left. - \frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} \sum_{m \neq n} G_{l,m} E_m X_m - \frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} G_{l,n} E_n X_n \right\} \\
& = \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} X_j - H_{i,n}^{(k)} X_n \right)^2 \right. \\
& \quad - \frac{1}{2} \left(\lambda \sum_{s \neq n} X_s C_{s,n} + \lambda \sum_{t \neq n} C_{n,t} X_t + \beta \sum_{l=1}^{M_1 M_2} G_{l,n} E_n \right) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \\
& \quad \left. - \frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t - \frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} \sum_{m \neq n} G_{l,m} E_m X_m \right\} \\
& = \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) - H_{i,n}^{(k)} X_n \right)^2 - \frac{1}{2} \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right. \\
& \quad \left. - \frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t - \frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} \sum_{m \neq n} G_{l,m} E_m X_m \right\}, \tag{3.74}
\end{aligned}$$

where

$$\psi_i^{(k)}(X_n) = y_i^{(k)} - \sum_{j \neq n} H_{i,j}^{(k)} X_j, \tag{3.75}$$

$$\phi(X_n) = \lambda \sum_{s \neq n} X_s C_{s,n} + \lambda \sum_{t \neq n} C_{n,t} X_t + \beta \sum_{l=1}^{M_1 M_2} G_{l,n} E_n. \tag{3.76}$$

Next, (3.74) can be simplified by removing the terms $-\frac{1}{2} \lambda \sum_{s \neq n} \sum_{t \neq n} X_s C_{st} X_t$ and $-\frac{1}{2} \beta \sum_{l=1}^{M_1 M_2} \sum_{m \neq n} G_{l,m} E_m X_m$ from (3.74), since these terms are independent

Appendix 3.D Derivation of (3.41)

with X_n . Therefore, (3.74) can be simplified as

$$p(X_n | \mathbf{X}_{-n}, \mathbf{Y}) \propto \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) - H_{i,n}^{(k)} X_n \right)^2 - \frac{1}{2} \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\}, \quad (3.77)$$

where $\psi_i^{(k)}(X_n)$ and $\phi(X_n)$ are defined as that in (3.75) and (3.76), respectively.

Furthermore, (3.77) can be rewritten as

$$\begin{aligned} p(X_n | \mathbf{X}_{-n}, \mathbf{Y}) &\propto \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) - H_{i,n}^{(k)} X_n \right)^2 - \frac{1}{2} \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\} \\ &= \exp \left\{ - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) \right)^2 + 2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \psi_i^{(k)}(X_n) H_{i,n}^{(k)} X_n \right. \\ &\quad \left. - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(H_{i,n}^{(k)} \right)^2 X_n^2 - \frac{1}{2} \lambda \phi(X_n) X_n - \frac{1}{2} \lambda C_{n,n} X_n^2 \right\} \\ &= \exp \left\{ - \left(\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(H_{i,n}^{(k)} \right)^2 + \frac{1}{2} \lambda C_{n,n} \right) X_n^2 \right. \\ &\quad \left. + \left(\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{\sigma_k^2} \psi_i^{(k)}(X_n) H_{i,n}^{(k)} - \frac{1}{2} \lambda \phi(X_n) \right) X_n - \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{1}{2\sigma_k^2} \left(\psi_i^{(k)}(X_n) \right)^2 \right\} \\ &\propto \exp \left\{ - \frac{1}{2\tau_{X_n}} (X_n - \mu_{X_n})^2 \right\}, \end{aligned} \quad (3.78)$$

where

$$\mu_{X_n} = \frac{\tau_{X_n}}{2} \left(2 \sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\psi_i^{(k)}(X_n) H_{i,n}^{(k)}}{\sigma_k^2} - \lambda \phi(X_n) \right), \quad (3.79)$$

$$\tau_{X_n} = \frac{1}{\sum_{k=1}^{\rho} \sum_{i=1}^{L_1 L_2} \frac{\left(H_{i,n}^{(k)} \right)^2}{\sigma_k^2} + \lambda C_{n,n}}. \quad (3.80)$$

Chapter 4

A State-space Super-resolution Video Approach

4.1 Introduction

In the SR still image case, a high-resolution still image is reconstructed based on a set of low-resolution images acquired. Likewise, in the SR video case, an image sequence with a higher resolution is produced from a group of adjacent low-resolution uncompressed image frames [122, 125, 126] or compressed image frames [159]. A direct way to apply the SR image techniques to conduct SR video reconstruction, as what have been done by most of the existing SR approaches for video signals, is based on the *sliding-window*-based Bayesian approach [11], as illustrated in Figure 4.1 (a). The sliding window selects a set of consecutive low-resolution frames for producing one high-resolution image frame; that is, the window is moved across the input frames to produce successive high-resolution frames sequentially. The major drawback of this approach is that the temporal correlations among the consecutively-reconstructed high-resolution images are not considered.

4.1 Introduction

Figure 4.1. Two conventional super-resolution video reconstruction methodologies: (a) the sliding-window-based super-resolution video scheme; and (b) the simultaneous multiple high-resolution frames reconstruction scheme.

To overcome this drawback, many works [122,125,126] have been done to simultaneously reconstruct multiple high-resolution images. Borman and Stevenson [122] proposed an algorithm to reconstruct multiple high-resolution images simultaneously by imposing temporal smoothness constraint on the prior image model, while Zibetti and Mayer [125] reduced its computational complexity by only considering the temporally-smooth low-resolution frames in the observation model. Alvarez *et al.* [126] proposed a multi-channel SR approach for the compressed image sequence, which applies the sliding window to select p low-resolution frames to reconstruct K high-resolution frames simultaneously. However, all these three algorithms require large memory storage, since these multiple high-resolution and low-resolution images are required available at the same time during the reconstruction process. Moreover, how many high-resolution frames should be exploited for simultaneous reconstructions is another important issue, which has not been addressed.

The major challenge in the SR video reconstruction is how to exploit the temporally-correlated information provided by the high-resolution images and the temporally-correlated low-resolution images respectively to improve the quality of the desired

4.1 Introduction

high-resolution images. To tackle this challenge, in this chapter, an extra observation equation is incorporated into the framework of the conventional two-equation-based Kalman filtering to build up a three-equation-based state-space model, followed by a full mathematical derivation to arrive at its closed-form solution. The proposed approach exploits the information from the previously-reconstructed high-resolution frame, the currently-observed low-resolution frame as well as the previously-observed low-resolution frame for sequentially producing the high-resolution frames. This is in contrast to the conventional two-equation-based Kalman filtering, in which only the previously-reconstructed high-resolution frame and the currently-observed low-resolution frame are exploited. Extensive simulations are conducted to demonstrate that the superior performance of the proposed state-space approach is consistently superior to that of the conventional two-equation-based Kalman filtering. Moreover, a unified theoretical analysis on the relationships of the proposed approach relative to the sliding-window-based Bayesian estimation approach and the conventional two-equation-based Kalman filtering, respectively, is provided as useful insights.

The chapter is organized as follows. In Section 4.2, a mathematical model of the SR imaging is presented first to establish the foundation for the remaining development of the chapter. A three-equation-based state-space model is proposed and the closed-form is mathematically derived in Section 4.3. Moreover, the computational complexity of the proposed algorithm and its respective relationship to other two existing SR approaches—the sliding-window-based Bayesian estimation approach and the conventional two-equation-based Kalman filtering, are discussed in Section 4.3. Extensive experiments are conducted using both the artificially-degraded videos and the real-life videos in Section 4.4. Finally, Section 4.5 concludes this chapter.

4.2 Observation Model

Figure 4.2. The relationship between the original high-resolution image $\mathbf{X}_n$ and the observed low-resolution image $\mathbf{y}_n$. The observed low-resolution image $\mathbf{y}_n$ is a warped, blurred and down-sampled version of the original high-resolution image $\mathbf{X}_n$, with the additive white Gaussian noise.

4.2 Observation Model

To address the SR problem, we begin by analyzing the image acquisition process, which is usually modeled by the following four stages: i) geometric transformation, ii) blurring, iii) down-sampling, and iv) noise adding. The geometric transformation includes translation, rotation and scaling. Various blurrings are usually modeled by convolving the image with a low-pass filter, which is modeled by a *point spread function* (PSF). The noise introduced in the image acquisition process is usually assumed to be an additive, zero-mean white Gaussian noise. This image acquisition process is illustrated in Figure 4.2, where the original high-resolution image (with a size of $M_1 \times M_2$) is warped, blurred, down-sampled and added with white Gaussian noise to generate the observed low-resolution image (with a size of $L_1 \times L_2$, where $L_1 = M_1/q_1$ and $L_2 = M_2/q_2$). Hence, the observation model can be mathematically formulated as

$$\mathbf{y}_n = \mathbf{D}_n \mathbf{P}_n \mathbf{R}_n \mathbf{X}_n + \mathbf{V}_n, \quad (4.1)$$

$$= \mathbf{H}_n \mathbf{X}_n + \mathbf{V}_n, \quad (4.2)$$

where $\mathbf{y}_n$ and $\mathbf{X}_n$ represent the low-resolution image and the high-resolution image in the temporal instance n , respectively. Both of these two images are represented in the lexicographic-ordered vector form, with a size of $L_1 L_2 \times 1$ and $M_1 M_2 \times 1$, respectively. $\mathbf{H}_n = \mathbf{D}_n \mathbf{P}_n \mathbf{R}_n$, where $\mathbf{D}_n$ is the decimation matrix with a size of $L_1 L_2 \times M_1 M_2$, $\mathbf{P}_n$ is the blurring matrix of size $M_1 M_2 \times M_1 M_2$, and $\mathbf{R}_n$ is the

4.3 Proposed State-space Super-resolution Imaging Framework

warping matrix of size $M_1 M_2 \times M_1 M_2$. Additionally, $\mathbf{V}_n$ is modeled as a zero-mean white Gaussian noise vector (with a size of $L_1 L_2 \times 1$), encountered during the image acquisition process. Furthermore, it is assumed to be independent of $\mathbf{X}_n$.

In summary, the SR problem is an *inverse problem*, and its goal is to make use of a set of low-resolution images $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ with a dimension of $L_1 \times L_2$ each to compute the high-resolution ones $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$ with a dimension of $M_1 \times M_2$ each (i.e., the enlarged resolution ratio is $q_1 \times q_2$).

4.3 Proposed State-space Super-resolution Imaging Framework

4.3.1 State-space Representations

The state-space models can be established to capture the relationship among the high-resolution images (i.e., the hidden states $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$) and the low-resolution images (i.e., the observations $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$). In this section, two state-space models are studied: i) the conventional two-equation-based Kalman filtering model, and ii) our proposed three-equation-based state-space model.

The conventional Kalman filter uses the following two equations to model the relationship among the states $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ themselves and their individual relationship with respect to corresponding observations $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, respectively (as illustrated in Fig. 4.3 (a)); that is,

$$\mathbf{X}_n = \mathbf{A}_n \mathbf{X}_{n-1} + \mathbf{W}_n, \quad (\text{the state equation}) \quad (4.3)$$

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{X}_n + \mathbf{V}_n. \quad (\text{the observation equation}) \quad (4.4)$$

The state equation (4.3) represents the relationship between two successive high-resolution images. The matrix $\mathbf{A}_n$ is an $M_1 M_2 \times M_1 M_2$ transition matrix

4.3 Proposed State-space Super-resolution Imaging Framework

between the current high-resolution image $\mathbf{X}_n$ and the previous high-resolution image $\mathbf{X}_{n-1}$. $\mathbf{W}_n$ is modeled as a zero-mean Gaussian noise with a covariance matrix $\mathbf{\Gamma}_{\mathbf{W}_n} = E\{\mathbf{W}_n \mathbf{W}_n^T\}$, for representing the estimation error, due to the inaccuracies incurred in estimating $\mathbf{A}_n$. $\mathbf{W}_n$ is further lexicographically ordered as an $M_1 M_2 \times 1$ vector. Secondly, at the discrete-time instance n , the relationship between the high-resolution image $\mathbf{X}_n$ and the low-resolution image $\mathbf{y}_n$ is formulated as the observation equation (4.4). $\mathbf{H}_n$ is as described in (4.1). $\mathbf{V}_n$ is modeled as a zero-mean Gaussian noise with a covariance matrix $\mathbf{\Gamma}_{\mathbf{V}_n} = E\{\mathbf{V}_n \mathbf{V}_n^T\}$, for representing both the estimation error, due to the inaccuracies incurred in estimating $\mathbf{H}_n$, and the additive noise encountered during the image acquisition process. $\mathbf{V}_n$ is further lexicographically ordered to an $L_1 L_2 \times 1$ vector. Furthermore, $\mathbf{V}_n$ is assumed to be independent of $\mathbf{W}_n$.

Given (4.3) and (4.4), the conventional two-equation-based Kalman filtering [160] or the adaptive filtering developed in [161–163] can be used to compute the least-squares estimation of each high-resolution image. To further exploit additional information provided by the previously-observed low-resolution image to improve the quality of the reconstructed high-resolution images, an extra *observation* equation is introduced in this chapter to represent the correlation between the successive low-resolution observations as follows,

$$\mathbf{y}_n = \mathbf{B}_n \mathbf{y}_{n-1} + \mathbf{U}_n, \quad (4.5)$$

where the matrix $\mathbf{B}_n$ is an $L_1 L_2 \times L_1 L_2$ transition matrix between the current low-resolution image $\mathbf{y}_n$ and the previous low-resolution image $\mathbf{y}_{n-1}$. $\mathbf{U}_n$ is modeled as a zero-mean Gaussian noise with a covariance matrix $\mathbf{\Gamma}_{\mathbf{U}_n} = E\{\mathbf{U}_n \mathbf{U}_n^T\}$, for representing the estimation error, due to the inaccuracies incurred in estimating $\mathbf{B}_n$. $\mathbf{U}_n$ is further lexicographically ordered as an $L_1 L_2 \times 1$ vector. Furthermore, $\mathbf{U}_n$ is assumed to be independent of $\mathbf{V}_n$.

4.3 Proposed State-space Super-resolution Imaging Framework

Figure 4.3. The relationship among the low-resolution frames $y_1, y_2, \dots, y_n$ and the high-resolution frames $X_1, X_2, \dots, X_n$ using: (a) the conventional two-equation-based Kalman filtering; (b) the proposed three-equation-based state-space filtering.

4.3 Proposed State-space Super-resolution Imaging Framework

Note that the proposed equation (4.5) is *not* obtained from any combination of (4.3) and (4.4), which can be justified from the following three aspects: i) mathematically, ii) physically, and iii) experimentally.

- i). *Mathematically*—The matrix $\mathbf{B}_n$ in (4.5) is computed by applying motion estimation process directly on the input data (i.e., the actual video frames), instead of using any derivations based on (4.3) and (4.4). Therefore, the proposed (4.5) does provide a chance to extract and exploit new information, if exists; otherwise, the performance would be degenerated to the Kalman filtering case.
- ii). *Physically*—The purpose of the proposed extra equation (4.5) is to exploit more information from the input low-resolution data. This point has been illustrated in Figure 4.3 (b), where the proposed approach exploits the information from the previously-reconstructed high-resolution frame, the currently-observed low-resolution frame as well as the previously-observed low-resolution frame for sequentially producing the high-resolution frames. This is in contrast to the conventional two-equation-based Kalman filtering, in which only the previously-reconstructed high-resolution frame and the currently-observed low-resolution frame are used (see Figure 4.3 (a)).

Further note that although both $\mathbf{U}_n$ (in (4.5)) and $\mathbf{V}_n$ (in (4.4)) are the so-called *noise*, they are totally different from the viewpoint of their actual physical meanings. To be more specific, $\mathbf{U}_n$ represents the estimation error, due to the inaccuracies incurred in estimating $\mathbf{B}_n$ in (4.5), while $\mathbf{V}_n$ represents both the estimation error (due to the inaccuracies incurred in estimating $\mathbf{H}_n$) and the additive noise encountered during the image acquisition process.

4.3 Proposed State-space Super-resolution Imaging Framework

iii). *Experimentally*—If the proposed third equation (4.5) had already been covered by the conventional two-equation-based Kalman filtering state-space framework (4.3) and (4.4), then the three-equation-based simulation results on SR imaging will be identical to that of the two-equation-based case. However, this is obviously not the case as the extensive experimental results as documented in Section 4.4. In fact, the SR results show that the proposed three-equation-based state-space framework is always superior to the conventional two-equation-based Kalman filtering. This clearly reflects the fact that the proposed (4.5) can exploit new and useful information from the input data, which are not available in either (4.3) or (4.4), or both.

4.3.2 Conventional Two-equation-based Kalman Filtering for SR Video Reconstruction

The conventional Kalman filter formulates the relationship among the high-resolution images $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ and the low-resolution images $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, using the two-equation-based state-space model (i.e., (4.3)-(4.4)). More specifically, the above-mentioned relationship can be mathematically expressed as

$$\begin{aligned}
 &P(\mathbf{X}_1, \dots, \mathbf{X}_n, \mathbf{y}_1, \dots, \mathbf{y}_n) \\
 &= P(\mathbf{X}_n, \mathbf{y}_n | \mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1), \\
 &= P(\mathbf{X}_n, \mathbf{y}_n | \mathbf{X}_{n-1}) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1), \\
 &= P(\mathbf{y}_n | \mathbf{X}_n, \mathbf{X}_{n-1}) P(\mathbf{X}_n | \mathbf{X}_{n-1}) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1), \\
 &= P(\mathbf{y}_n | \mathbf{X}_n) P(\mathbf{X}_n | \mathbf{X}_{n-1}) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1). \tag{4.6}
 \end{aligned}$$

Based on (4.3)-(4.4), the conventional two-equation-based Kalman filter [160] estimates each high-resolution frame $\mathbf{X}_n$ (denoted as $\hat{\mathbf{X}}_n$) via [164]

$$\hat{\mathbf{X}}_n = \mathbf{A}_n \hat{\mathbf{X}}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{H}_n \mathbf{A}_n \hat{\mathbf{X}}_{n-1}), \tag{4.7}$$

4.3 Proposed State-space Super-resolution Imaging Framework

where $\mathbf{K}_n$ is [164]

$$\mathbf{K}_n = \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n} \right) \mathbf{H}_n^T \left[\mathbf{H}_n \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n} \right) \mathbf{H}_n^T + \Gamma_{\mathbf{v}_n} \right]^{-1}. \quad (4.8)$$

According to (4.7), the conventional two-equation-based Kalman filter computes the least-squares estimation of the high-resolution image, $\hat{\mathbf{X}}_n$, using the prediction result based on the previous high-resolution image, $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$, plus an updating term, which consists of an adaptive weighting factor $\mathbf{K}_n$ times the difference between the current observation $\mathbf{y}_n$ and the prediction result $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$ weighted by $\mathbf{H}_n$.

4.3.3 Proposed Three-equation-based State-space SR Approach

In contrast to the above-mentioned conventional two-equation-based Kalman filtering, a new filtering is mathematically derived from the above-mentioned three-equation-based state-space model (i.e., (4.3)-(4.5)) in this section, as follows.

According to (4.3)-(4.5), the relationship among the low-resolution images and the high-resolution images can be mathematically expressed as

$$\begin{aligned} & P(\mathbf{X}_1, \dots, \mathbf{X}_n, \mathbf{y}_1, \dots, \mathbf{y}_n) \\ &= P(\mathbf{X}_n, \mathbf{y}_n | \mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1), \\ &= P(\mathbf{X}_n, \mathbf{y}_n | \mathbf{X}_{n-1}, \mathbf{y}_{n-1}) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1), \\ &= P(\mathbf{y}_n | \mathbf{X}_n, \mathbf{X}_{n-1}, \mathbf{y}_{n-1}) P(\mathbf{X}_n | \mathbf{X}_{n-1}, \mathbf{y}_{n-1}) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1), \\ &= P(\mathbf{y}_n | \mathbf{X}_n, \mathbf{y}_{n-1}) P(\mathbf{X}_n | \mathbf{X}_{n-1}) P(\mathbf{X}_{n-1}, \dots, \mathbf{X}_1, \mathbf{y}_{n-1}, \dots, \mathbf{y}_1). \end{aligned} \quad (4.9)$$

Comparing (4.6) and (4.9), one can clearly see the significant difference between the conventional two-equation-based Kalman filtering and the proposed three-equation-based state-space filtering, where the previous observation $\mathbf{y}_{n-1}$ is explicitly brought into the proposed framework via (4.9).

4.3 Proposed State-space Super-resolution Imaging Framework

The challenge of assessing the above distribution (4.9) is to compute the conditional probability $P\{\mathbf{y}_n|\mathbf{X}_n, \mathbf{y}_{n-1}\}$, which needs to be computed based on the conditional probability $P\{\mathbf{y}_n|\mathbf{X}_n\}$ and $P\{\mathbf{y}_n|\mathbf{y}_{n-1}\}$. For that, various algorithms can be found in the field of statistics to solve this conditional probability density function combining problem [165], among which the *logarithmic opinion pool* [165] is exploited and summarized as follows.

Let $P(\mathbf{C}|\mathbf{A}) \sim N(\mu_{\mathbf{A}}, \Gamma_{\mathbf{A}})$ and $P(\mathbf{C}|\mathbf{B}) \sim N(\mu_{\mathbf{B}}, \Gamma_{\mathbf{B}})$, then we have

$$P(\mathbf{C}|\mathbf{A}, \mathbf{B}) = \frac{(P(\mathbf{C}|\mathbf{A}))^{\lambda_{\mathbf{A}}}(P(\mathbf{C}|\mathbf{B}))^{\lambda_{\mathbf{B}}}}{\sum_{\mathbf{C}}(P(\mathbf{C}|\mathbf{A}))^{\lambda_{\mathbf{A}}}(P(\mathbf{C}|\mathbf{B}))^{\lambda_{\mathbf{B}}}} \sim N(\mu_{\mathbf{C}}, \Gamma_{\mathbf{C}}), \quad (4.10)$$

where

$$\begin{aligned} \Gamma_{\mathbf{C}}^{-1} &= \lambda_{\mathbf{A}}\Gamma_{\mathbf{A}}^{-1} + \lambda_{\mathbf{B}}\Gamma_{\mathbf{B}}^{-1}, \\ \mu_{\mathbf{C}} &= \Gamma_{\mathbf{C}} (\lambda_{\mathbf{A}}\Gamma_{\mathbf{A}}^{-1}\mu_{\mathbf{A}} + \lambda_{\mathbf{B}}\Gamma_{\mathbf{B}}^{-1}\mu_{\mathbf{B}}), \end{aligned} \quad (4.11)$$

$\lambda_{\mathbf{A}}$ and $\lambda_{\mathbf{B}}$ are the pooling weights such that $\lambda_{\mathbf{A}}, \lambda_{\mathbf{B}} \geq 0$ and $\lambda_{\mathbf{A}} + \lambda_{\mathbf{B}} = 1$.

Suppose both the high-resolution image sequence $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ and the low-resolution observation sequence $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ are Gaussian processes. Then (4.3)-(4.5), can be rewritten as

$$P(\mathbf{X}_n|\mathbf{X}_{n-1}) \sim N(\mathbf{A}_n\mathbf{X}_{n-1}, \Gamma_{\mathbf{W}_n}), \quad (4.12)$$

$$P(\mathbf{y}_n|\mathbf{X}_n) \sim N(\mathbf{H}_n\mathbf{X}_n, \Gamma_{\mathbf{V}_n}), \quad (4.13)$$

$$P(\mathbf{y}_n|\mathbf{y}_{n-1}) \sim N(\mathbf{B}_n\mathbf{y}_{n-1}, \Gamma_{\mathbf{U}_n}). \quad (4.14)$$

By applying (4.10) on (4.13) and (4.14), $P(\mathbf{y}_n|\mathbf{y}_{n-1}, \mathbf{X}_n)$ can be derived as follows,

$$P(\mathbf{y}_n|\mathbf{y}_{n-1}, \mathbf{X}_n) \sim N(\mu_{\mathbf{y}_n}, \Gamma_{\mathbf{y}_n}), \quad (4.15)$$

where

$$\begin{aligned} \Gamma_{\mathbf{y}_n}^{-1} &= \lambda_{\mathbf{X}}\Gamma_{\mathbf{V}_n}^{-1} + \lambda_{\mathbf{y}}\Gamma_{\mathbf{U}_n}^{-1}, \\ \mu_{\mathbf{y}_n} &= \Gamma_{\mathbf{y}_n} (\lambda_{\mathbf{X}}\Gamma_{\mathbf{V}_n}^{-1}\mathbf{H}_n\mathbf{X}_n + \lambda_{\mathbf{y}}\Gamma_{\mathbf{U}_n}^{-1}\mathbf{B}_n\mathbf{y}_{n-1}) \\ &= \Gamma_{\mathbf{y}_n} (\lambda_{\mathbf{X}}\Gamma_{\mathbf{V}_n}^{-1}\mathbf{H}_n\mathbf{X}_n + \lambda_{\mathbf{y}}\Gamma_{\mathbf{U}_n}^{-1}\mathbf{B}_n\mathbf{H}_n\mathbf{X}_{n-1}), \end{aligned} \quad (4.16)$$

4.3 Proposed State-space Super-resolution Imaging Framework

in which $\lambda_{\mathbf{X}}$ and $\lambda_{\mathbf{y}}$ are the pooling weights such that $\lambda_{\mathbf{X}}, \lambda_{\mathbf{y}} \geq 0$ and $\lambda_{\mathbf{X}} + \lambda_{\mathbf{y}} = 1$. Denote the matrix $\mathbf{T}_{n-1,n}$ as the backward transition matrix between two successive frames $\mathbf{X}_n$ and $\mathbf{X}_{n-1}$; that is, $\mathbf{X}_{n-1} = \mathbf{T}_{n-1,n}\mathbf{X}_n$. By substituting $\mathbf{X}_{n-1} = \mathbf{T}_{n-1,n}\mathbf{X}_n$ into (4.16), we can rewrite (4.16) as

$$\mu_{\mathbf{y}_n} = \Gamma_{\mathbf{y}_n} (\lambda_{\mathbf{X}} \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n \mathbf{X}_n + \lambda_{\mathbf{y}} \Gamma_{\mathbf{U}_n}^{-1} \mathbf{B}_n \mathbf{H}_n \mathbf{T}_{n-1,n} \mathbf{X}_n) = \mathbf{P}_n \mathbf{X}_n, \quad (4.17)$$

where

$$\mathbf{P}_n = \Gamma_{\mathbf{y}_n} (\lambda_{\mathbf{X}} \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n + \lambda_{\mathbf{y}} \Gamma_{\mathbf{U}_n}^{-1} \mathbf{B}_n \mathbf{H}_n \mathbf{T}_{n-1,n}). \quad (4.18)$$

The derivation of the closed-form solution of (4.3)-(4.5) for estimating the high-resolution frame $\mathbf{X}_n$ (denoted as $\hat{\mathbf{X}}_n$) can be arrived as

$$\begin{aligned} \hat{\mathbf{X}}_n &= \mathbf{A}_n \hat{\mathbf{X}}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \hat{\mathbf{X}}_{n-1}) \\ &\quad + \mathbf{G}_n (\mathbf{y}_{n-1} - \mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{A}_n \hat{\mathbf{X}}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \hat{\mathbf{X}}_{n-1}))), \end{aligned} \quad (4.19)$$

$$\Gamma_{\hat{\mathbf{X}}_n} = (\mathbf{I} - \mathbf{G}_n \mathbf{P}_{n-1} \mathbf{F}_n) (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}), \quad (4.20)$$

where

$$\begin{aligned} \mathbf{F}_n &= \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T (\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n})^{-1}, \\ \mathbf{K}_n &= (\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T [\mathbf{P}_n (\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T + \Gamma_{\mathbf{y}_n}]^{-1}, \\ \mathbf{G}_n &= (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\ &\quad \cdot [\mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\ &\quad + \mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{y}_{n-1}}]^{-1}. \end{aligned} \quad (4.21)$$

The detailed derivation steps of (4.19) are documented in Section 4.5.

Let $\hat{\mathbf{X}}_n^-$ be the *priori* estimate of $\mathbf{X}_n$ based on the previously-reconstructed high-resolution image $\mathbf{X}_{n-1}$, $\hat{\mathbf{X}}_n^-$ be the *priori* estimate of $\mathbf{X}_n$ by further incorporating the currently-observed measurement $\mathbf{y}_n$, and $\hat{\mathbf{X}}_n$ be the *posterior* estimate of $\mathbf{X}_n$ by further incorporating the previously-observed measurement $\mathbf{y}_{n-1}$. According

4.3 Proposed State-space Super-resolution Imaging Framework

to (4.19), the estimated high-resolution image at the discrete-time instance n , $\hat{\mathbf{X}}_n$, can be reconstructed by the following three steps:

- Step 1: *Predict from the previously-reconstructed high-resolution frame $\hat{\mathbf{X}}_{n-1}$,*

$$\hat{\mathbf{X}}_n^{--} = \mathbf{A}_n \hat{\mathbf{X}}_{n-1}, \quad (4.22)$$

$$\Gamma_{\hat{\mathbf{X}}_n^{--}} = \mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n}. \quad (4.23)$$

That is, the priori estimate $\hat{\mathbf{X}}_n^{--}$ is obtained by projecting the previously-reconstructed high-resolution image $\hat{\mathbf{X}}_{n-1}$ onto the next discrete-time instance via $\mathbf{A}_n$.

- Step 2: *Update $\hat{\mathbf{X}}_n^{--}$ according to the current observation $\mathbf{y}_n$,*

$$\hat{\mathbf{X}}_n^{-} = \hat{\mathbf{X}}_n^{--} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \hat{\mathbf{X}}_n^{--}), \quad (4.24)$$

$$\Gamma_{\hat{\mathbf{X}}_n^{-}} = (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) \Gamma_{\hat{\mathbf{X}}_n^{--}}. \quad (4.25)$$

That is, the priori estimate $\hat{\mathbf{X}}_n^{--}$ can be updated by incorporating the currently-observed low-resolution image $\mathbf{y}_n$.

- Step 3: *Update $\hat{\mathbf{X}}_n^{-}$ according to the previous observation $\mathbf{y}_{n-1}$,*

$$\hat{\mathbf{X}}_n = \hat{\mathbf{X}}_n^{-} + \mathbf{G}_n (\mathbf{y}_{n-1} - \mathbf{P}_{n-1} \mathbf{F}_n \hat{\mathbf{X}}_n^{-}), \quad (4.26)$$

$$\Gamma_{\hat{\mathbf{X}}_n} = (\mathbf{I} - \mathbf{G}_n \mathbf{P}_{n-1} \mathbf{F}_n) \Gamma_{\hat{\mathbf{X}}_n^{-}}. \quad (4.27)$$

That is, the posterior estimate $\hat{\mathbf{X}}_n$ can be obtained by incorporating the previously-observed low-resolution image $\mathbf{y}_{n-1}$. These three stages altogether take advantage of previously-reconstructed high-resolution frame (see (4.22)), the currently-observed low-resolution image (see (4.24)) and the previously-observed low-resolution image (see (4.26)) as well to produce each high-resolution frame.

4.3 Proposed State-space Super-resolution Imaging Framework

4.3.4 Relationships to Other SR Approaches

In this section, a theoretical analysis is provided to present the relationship between the proposed framework and the other two existing SR approaches: *the sliding-window-based Bayesian estimation approach* and *the conventional two-equation-based Kalman filtering*, respectively.

Relationship to the Sliding-window-based Bayesian Estimation Approach

The sliding-window-based Bayesian estimation approach estimates one high-resolution image based on a few adjacent low-resolution frames through some criteria, such as the *maximum likelihood* (ML) and the *maximum a posterior probability* (MAP). The successive high-resolution images are then estimated likewise based on a set of low-resolution frames, which are selected by the sliding window. In this section, an analysis is provided to show that the proposed framework, in effect, produces individual high-resolution image based on two adjacent low-resolution observations each time, if the temporal correlations among high-resolution images are omitted. Incorporating the above constraints into the proposed scheme, (4.15) becomes

$$P\{\mathbf{y}_n|\mathbf{y}_{n-1}, \mathbf{X}_n\} = P\{\mathbf{y}_n|\mathbf{X}_n\} \sim N(\mathbf{H}_n\mathbf{X}_n, \mathbf{\Gamma}_{\mathbf{v}_n}) = N(\mu_{\mathbf{y}_n}, \mathbf{\Gamma}_{\mathbf{y}_n}), \quad (4.28)$$

which indicates

$$\mu_{\mathbf{y}_n} = \mathbf{H}_n\mathbf{X}_n, \quad \mathbf{\Gamma}_{\mathbf{y}_n} = \mathbf{\Gamma}_{\mathbf{v}_n}. \quad (4.29)$$

Then, comparing (4.17) with (4.28), we have

$$\mathbf{P}_n = \mathbf{H}_n. \quad (4.30)$$

Furthermore, the uncertainty in the forward estimated motion information matrix $\mathbf{A}_n$ could be omitted, which means that

$$\mathbf{\Gamma}_{\mathbf{w}_n} = \mathbf{0}. \quad (4.31)$$

4.3 Proposed State-space Super-resolution Imaging Framework

Now, substituting (4.29)-(4.31) into (4.22)-(4.27), and applying a lemma of the inverse matrix [166], $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$, the proposed scheme (4.22)-(4.27) becomes

- Step 1:

$$\hat{\mathbf{X}}_n^{--} = \mathbf{A}_n \hat{\mathbf{X}}_{n-1}, \quad (4.32)$$

$$\Gamma_{\hat{\mathbf{X}}_n^{--}} = \mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T. \quad (4.33)$$

- Step 2:

$$\hat{\mathbf{X}}_n^- = \left(\mathbf{H}_n^T \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n + \Gamma_{\hat{\mathbf{X}}_n^{--}}^{-1} \right)^{-1} \cdot \left(\Gamma_{\hat{\mathbf{X}}_n^{--}}^{-1} \hat{\mathbf{X}}_n^{--} + \mathbf{H}_n^T \Gamma_{\mathbf{V}_n}^{-1} \mathbf{y}_n \right), \quad (4.34)$$

$$\Gamma_{\hat{\mathbf{X}}_n^-} = (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T. \quad (4.35)$$

- Step 3:

$$\begin{aligned} \hat{\mathbf{X}}_n = & \left(\left(\mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \right)^T \Gamma_{\mathbf{V}_{n-1}}^{-1} \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right)^{-1} \\ & \cdot \left(\Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \hat{\mathbf{X}}_n^- + \left(\mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \right)^T \Gamma_{\mathbf{V}_{n-1}}^{-1} \mathbf{y}_{n-1} \right), \end{aligned} \quad (4.36)$$

$$\Gamma_{\hat{\mathbf{X}}_n} = \left(\mathbf{I} - \mathbf{G}_n \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \right) (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T. \quad (4.37)$$

The complete derivations of (4.34) and (4.36) are presented in Appendix 4.5.

The following discussion will show that (4.32)-(4.37), in fact, sequentially compute the MAP estimation of the high-resolution image $\mathbf{X}_n$, based on two observations $\mathbf{y}_n$ and $\mathbf{y}_{n-1}$. Let $\hat{\mathbf{X}}_n^{Init}$ be the initial estimation of $\mathbf{X}_n$, $\hat{\mathbf{X}}_n^{MAP-}$ be the MAP estimation of $\mathbf{X}_n$ by incorporating the observation $\mathbf{y}_n$, and $\hat{\mathbf{X}}_n^{MAP}$ be the MAP estimation of $\mathbf{X}_n$ by further incorporating the observation $\mathbf{y}_{n-1}$ based on $\hat{\mathbf{X}}_n^{MAP-}$. Thus, the priori probability distribution of $\mathbf{X}_n$ can be expressed as

$$P(\mathbf{X}_n) \sim N \left(\hat{\mathbf{X}}_n^{Init}, \Gamma_{\hat{\mathbf{X}}_n^{Init}} \right). \quad (4.38)$$

Furthermore, $\hat{\mathbf{X}}_n^{--}$ is used as the initial estimation $\hat{\mathbf{X}}_n^{Init}$, therefore, $\Gamma_{\hat{\mathbf{X}}_n^{Init}} = \mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T$ (see (4.33)).

4.3 Proposed State-space Super-resolution Imaging Framework

- Find the MAP estimation of the high-resolution image $\mathbf{X}_n$, $\hat{\mathbf{X}}_n^{MAP-}$, by incorporating the observation $\mathbf{y}_n$.

To compute $\hat{\mathbf{X}}_n^{MAP-}$, which maximizes the posterior probability $P(\mathbf{X}_n|\mathbf{y}_n)$, the Bayes rule is firstly applied to express $P(\mathbf{X}_n|\mathbf{y}_n)$ as

$$P(\mathbf{X}_n|\mathbf{y}_n) = \frac{P(\mathbf{y}_n|\mathbf{X}_n)P(\mathbf{X}_n)}{P(\mathbf{y}_n)} \propto P(\mathbf{y}_n|\mathbf{X}_n)P(\mathbf{X}_n). \quad (4.39)$$

According to (4.13), the conditional probability distribution $P(\mathbf{y}_n|\mathbf{X}_n)$ can be rewritten as

$$P(\mathbf{y}_n|\mathbf{X}_n) \propto \exp\left(-\frac{1}{2}(\mathbf{y}_n - \mathbf{H}_n\mathbf{X}_n)^T \Gamma_{\mathbf{V}_n}^{-1}(\mathbf{y}_n - \mathbf{H}_n\mathbf{X}_n)\right). \quad (4.40)$$

Meanwhile, according to (4.38), $P(\mathbf{X}_n)$ can be rewritten as

$$P(\mathbf{X}_n) \propto \exp\left(-\frac{1}{2}\left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{Init}\right)^T \Gamma_{\hat{\mathbf{X}}_n^{Init}}^{-1}\left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{Init}\right)\right). \quad (4.41)$$

Then, substituting (4.40) and (4.41) into (4.39), $P(\mathbf{X}_n|\mathbf{y}_n)$ can be obtained as

$$P(\mathbf{X}_n|\mathbf{y}_n) \propto \exp\left(-\frac{1}{2}(\mathbf{y}_n - \mathbf{H}_n\mathbf{X}_n)^T \Gamma_{\mathbf{V}_n}^{-1}(\mathbf{y}_n - \mathbf{H}_n\mathbf{X}_n) - \frac{1}{2}\left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{Init}\right)^T \Gamma_{\hat{\mathbf{X}}_n^{Init}}^{-1}\left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{Init}\right)\right). \quad (4.42)$$

Finally, since $\hat{\mathbf{X}}_n^{MAP-}$ satisfies that

$$\left.\frac{\partial \log P(\mathbf{X}_n|\mathbf{y}_n)}{\partial \mathbf{X}_n}\right|_{\mathbf{X}_n=\hat{\mathbf{X}}_n^{MAP-}} = 0, \quad (4.43)$$

substituting (4.42) into (4.43), $\hat{\mathbf{X}}_n^{MAP-}$ can be obtained as

$$\hat{\mathbf{X}}_n^{MAP-} = \left(\mathbf{H}_n^T \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n + \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T\right)^{-1}\right)^{-1} \left(\left(\mathbf{A}_n \Gamma_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T\right)^{-1} \hat{\mathbf{X}}_n^{--} + \mathbf{H}_n^T \Gamma_{\mathbf{V}_n}^{-1} \mathbf{y}_n\right). \quad (4.44)$$

Comparing (4.34) with (4.44), the result of the second step of the proposed scheme is, in effect, the MAP estimation that maximizes the posterior probability $P(\mathbf{X}_n|\mathbf{y}_n)$.

4.3 Proposed State-space Super-resolution Imaging Framework

- Find the MAP estimation of the image $\mathbf{X}_n$, $\hat{\mathbf{X}}_n^{MAP}$, by further incorporating the observation $\mathbf{y}_{n-1}$.

To compute $\hat{\mathbf{X}}_n^{MAP}$, which maximizes the posterior probability $P(\mathbf{X}_n|\mathbf{y}_{n-1})$, the Bayes rule is firstly applied to express $P(\mathbf{X}_n|\mathbf{y}_{n-1})$ as

$$P(\mathbf{X}_n|\mathbf{y}_{n-1}) = \frac{P(\mathbf{y}_{n-1}|\mathbf{X}_n)P(\mathbf{X}_n)}{P(\mathbf{y}_{n-1})} \propto P(\mathbf{y}_{n-1}|\mathbf{X}_n)P(\mathbf{X}_n). \quad (4.45)$$

Referring to (4.3) and (4.4), the relationship between $\mathbf{y}_{n-1}$ and $\mathbf{X}_n$ can be expressed as

$$\mathbf{y}_{n-1} = \mathbf{H}_n \mathbf{X}_{n-1} + \mathbf{V}_{n-1} = \mathbf{H}_n \mathbf{A}_n^\dagger \mathbf{X}_n + \mathbf{V}_{n-1}, \quad (4.46)$$

where $\mathbf{A}_n^\dagger$ is the pseudoinverse of the matrix $\mathbf{A}_n$ [167]; that is $\mathbf{A}_n^\dagger = (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T$.

Hence, the conditional probability distribution $P(\mathbf{y}_{n-1}|\mathbf{X}_n)$ can be written as

$$P(\mathbf{y}_{n-1}|\mathbf{X}_n) \propto \exp \left(-\frac{1}{2} \left(\mathbf{y}_{n-1} - \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \mathbf{X}_n \right)^T \boldsymbol{\Gamma}_{\mathbf{V}_{n-1}}^{-1} \cdot \left(\mathbf{y}_{n-1} - \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \mathbf{X}_n \right) \right). \quad (4.47)$$

Meanwhile, according to (4.44), $P(\mathbf{X}_n)$ is

$$P(\mathbf{X}_n) \propto \exp \left(-\frac{1}{2} \left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{MAP-} \right)^T \boldsymbol{\Gamma}_{\hat{\mathbf{X}}_n^{MAP-}}^{-1} \left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{MAP-} \right) \right), \quad (4.48)$$

where $\boldsymbol{\Gamma}_{\hat{\mathbf{X}}_n^{MAP-}} = (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \mathbf{A}_n \boldsymbol{\Gamma}_{\hat{\mathbf{X}}_{n-1}} \mathbf{A}_n^T$. Then, substituting (4.47) and (4.48) into (4.45), we can get

$$\begin{aligned} & P(\mathbf{X}_n|\mathbf{y}_{n-1}) \\ & \propto \exp \left(-\frac{1}{2} \left(\mathbf{y}_{n-1} - \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \mathbf{X}_n \right)^T \boldsymbol{\Gamma}_{\mathbf{V}_{n-1}}^{-1} \left(\mathbf{y}_{n-1} - \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \mathbf{X}_n \right) \right. \\ & \quad \left. -\frac{1}{2} \left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{MAP-} \right)^T \boldsymbol{\Gamma}_{\hat{\mathbf{X}}_n^{MAP-}}^{-1} \left(\mathbf{X}_n - \hat{\mathbf{X}}_n^{MAP-} \right) \right). \end{aligned} \quad (4.49)$$

Finally, since $\hat{\mathbf{X}}_n^{MAP}$ satisfies that

$$\left. \frac{\partial \log P(\mathbf{X}_n|\mathbf{y}_{n-1})}{\partial \mathbf{X}_n} \right|_{\mathbf{X}_n = \hat{\mathbf{X}}_n^{MAP}} = 0, \quad (4.50)$$

4.3 Proposed State-space Super-resolution Imaging Framework

substituting (4.49) into (4.50), $\hat{\mathbf{X}}_n^{MAP}$ can be obtained as

$$\begin{aligned} \hat{\mathbf{X}}_n^{MAP} = & \left(\left(\mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \right)^T \Gamma_{\mathbf{V}_{n-1}}^{-1} \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T + \Gamma_{\hat{\mathbf{X}}_n}^{-1} \right)^{-1} \\ & \cdot \left(\Gamma_{\hat{\mathbf{X}}_n}^{-1} \hat{\mathbf{X}}_n^{--} + \left(\mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \right)^T \Gamma_{\mathbf{V}_{n-1}}^{-1} \mathbf{y}_{n-1} \right). \end{aligned} \quad (4.51)$$

Comparing (4.36) with (4.51), the result of the third step of the proposed scheme is, in effect, the MAP estimation that maximizes the posterior probability $P(\mathbf{X}_n | \mathbf{y}_n)$.

According to the above discussions, the proposed scheme boils down to computing the MAP estimation of the high-resolution image $\hat{\mathbf{X}}_n$ simply based on two adjacent low-resolution images, $\mathbf{y}_n$ and $\mathbf{y}_{n-1}$, if the temporal correlations among the high-resolution images are neglected.

Relationship to the Conventional Two-equation-based Kalman Filtering Approach

The conventional two-equation-based Kalman filter generates the high-resolution sequence with the consideration of the correlation between adjacent high-resolution frames, as well as the correlation between the current low-resolution observation and the current high-resolution frame, by using only (4.3) and (4.4) to represent the relationship among the low-resolution images and high-resolution images.

Based on (4.3) and (4.4), the conventional two-equation-based Kalman filter [160] computes the least-squares estimation of the high-resolution image, $\hat{\mathbf{X}}_n$, using the prediction result (via the motion estimation matrix $\mathbf{A}_n$) based on the previous high-resolution image, $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$, plus an updating term, which consists of an adaptive weighting factor $\mathbf{K}_n$ times the difference between the current observation $\mathbf{y}_n$ and the prediction result $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$ weighted by $\mathbf{H}_n$; that is [164]

$$\hat{\mathbf{X}}_n = \mathbf{A}_n \hat{\mathbf{X}}_{n-1} + \mathbf{K}_n \left(\mathbf{y}_n - \mathbf{H}_n \mathbf{A}_n \hat{\mathbf{X}}_{n-1} \right), \quad (4.52)$$

4.3 Proposed State-space Super-resolution Imaging Framework

where $\mathbf{K}_n$ is [164]

$$\mathbf{K}_n = \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n} \right) \mathbf{H}_n^T \left[\mathbf{H}_n \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n} \right) \mathbf{H}_n^T + \Gamma_{\mathbf{v}_n} \right]^{-1}. \quad (4.53)$$

4.3.5 Computational Complexity

In this section, the computational complexity of the proposed framework with that of the conventional two-equation-based Kalman filtering will be compared stage by stage. Throughout the following analysis, it is assumed that the high-resolution image is with a size of $M_1 \times M_2$, $L_1 \times L_2$ for the low-resolution image, and $p \times p$ for the kernel of PSF. Hence, $\mathbf{H}_n$ is a sparse matrix, in which each column contains only $p^2 L_1 L_2 / M_1 M_2$ non-zero entries. A study of the computational complexity of these two methods is made in terms of the three stages as follows:

- Stage 1: Both algorithms require $M_1^2 M_2^2$ multiplications for the computation of $\mathbf{A}_n \hat{\mathbf{x}}_{n-1}$.
- Stage 2: The conventional two-equation-based Kalman filter (see (4.52)) needs to compute $\mathbf{K}_n$ first, which requires $2M_1^3 M_2^3$ multiplications for computing $\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T$ in (4.53), and requires $p^2 L_1^2 L_2^2 + p^2 L_1^3 L_2^3 / M_1 M_2$ multiplications for multiplying $\left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n} \right)$ with $\mathbf{H}_n$ and $\mathbf{H}_n^T$ in (4.53). Furthermore, the multiplication of $\mathbf{H}_n$ with $\mathbf{A}_n \hat{\mathbf{x}}_{n-1}$ in (4.52) requires $p^2 L_1^2 L_2^2 / M_1 M_2$ multiplications, and the multiplication of $\mathbf{K}_n$ with the term $\left(\mathbf{y}_n - \mathbf{H}_n \mathbf{A}_n \hat{\mathbf{x}}_{n-1} \right)$ in (4.52) requires $L_1 L_2 M_1 M_2$ multiplications.

On the other hand, the proposed framework (see (4.24)) needs to compute $\mathbf{K}_n$ first, which requires $2M_1^3 M_2^3$ multiplications for computing $\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T$ in (4.21), and requires $L_1 L_2 M_1^2 M_2^2 + L_1^2 L_2^2 M_1 M_2$ multiplications for multiplying $\left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n} \right)$ with $\mathbf{P}_n$ and $\mathbf{P}_n^T$ in (4.21). Next, the computation

4.3 Proposed State-space Super-resolution Imaging Framework

of $\mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \hat{\mathbf{X}}_n^-)$ in (4.24) requires $2L_1 L_2 M_1 M_2$ multiplications. Hence, at this stage, compared with that of the conventional two-equation-based Kalman filtering, the additional computational cost of the proposed algorithm is $L_1 L_2 M_1^2 M_2^2 + L_1^2 L_2^2 M_1 M_2 + L_1 L_2 M_1 M_2 - p^2 L_1^2 L_2^2 - p^2 L_1^3 L_2^3 / M_1 M_2 - p^2 L_1^2 L_2^2 / M_1 M_2$ multiplications.

- Stage 3: The proposed framework (see (4.26)) needs to calculate $\mathbf{G}_n$ in (4.21) first, which requires $2L_1 L_2 M_1^2 M_2^2 + L_1^2 L_2^2 M_1 M_2$ multiplications. Then, the computation of $\mathbf{P}_{n-1} \mathbf{F}_n \hat{\mathbf{X}}_n^-$ in (4.26) requires $L_1 L_2 M_1 M_2 + M_1^2 M_2^2$ multiplications, and the multiplication of the term $\mathbf{y}_{n-1} - \mathbf{P}_{n-1} \mathbf{F}_n \hat{\mathbf{X}}_n^-$ with $\mathbf{G}_n$ in (4.26) requires $L_1 L_2 M_1 M_2$ multiplications. Hence, at this stage, compared with that of the conventional two-equation-based Kalman filtering, the additional computational cost of the proposed algorithm is $2L_1 L_2 M_1^2 M_2^2 + L_1^2 L_2^2 M_1 M_2 + 2L_1 L_2 M_1 M_2 + M_1^2 M_2^2$ multiplications.

To appreciate the above-mentioned computational complexity through a simple numerical example, let $M_1 = 176$, $M_2 = 144$, and $p = 3$. Note that the frame size of the test sequence *Foreman* used in this chapter is 176×144 . Compared with the conventional two-equation-based Kalman filtering, the proposed algorithm requires additional 43.75% and 9.77% matrix multiplications, for the cases when the enlargement ratios are 2×2 and 4×4 , respectively.

4.4 Experimental results

4.4.1 Implementation

The proposed SR algorithm produces the high-resolution sequence as follows. The first observed low-resolution frame $\mathbf{y}_1$ is interpolated (say, using the bi-cubic spline interpolation) to produce the first high-resolution frame $\hat{\mathbf{X}}_1$. Meanwhile, the state uncertainty is set to zero; that is, $\mathbf{\Gamma}_{\hat{\mathbf{X}}_1} = \mathbf{0}$, since our experiments show that the reconstructed results are independent of $\mathbf{\Gamma}_{\hat{\mathbf{X}}_1}$ after finite frames (around 10 frames). Next, the 4×4 full-search block-matching algorithm (BMA) is used to estimate the half-pixel accuracy motion information matrix $\mathbf{B}_2$ between $\mathbf{y}_1$ and $\mathbf{y}_2$, since BMA achieves reasonably accurate registration performance that is fairly sufficient for the SR problem [168]. Then the motion information matrix $\mathbf{A}_2$ is estimated by up-sampling the low-resolution frame $\mathbf{y}_2$ (say, using the bi-cubic spline interpolation), followed by applying BMA on the previously-reconstructed high-resolution frame $\hat{\mathbf{X}}_1$ and the above-mentioned upsampled frame. Then $\hat{\mathbf{X}}_1$ is projected via $\mathbf{A}_2$ to obtain the initial estimate $\hat{\mathbf{X}}_2^-$ for the next discrete-time instance (see (4.22)). The pooling weights $\lambda_{\mathbf{x}}$ and $\lambda_{\mathbf{y}}$ are both set to be 0.5, which means that the two input probability density functions are viewed equally in generating the pooled probability density function [165]. Then, the observation $\mathbf{y}_2$ is incorporated to compute $\hat{\mathbf{X}}_2^-$ by (4.24). Finally, $\hat{\mathbf{X}}_2^-$ is updated by further incorporating the previous low-resolution frame $\mathbf{y}_1$ to get the final result $\hat{\mathbf{X}}_2$ (see (4.26)). The above process is repeated to compute the successive high-resolution frames.

On the other hand, the implementation of the conventional two-equation-based Kalman filtering is presented as follows. At each time instance n , it has two steps to reconstruct the high-resolution frame $\mathbf{X}_n$: i) predict $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$ according to the previously-reconstructed high-resolution frame $\hat{\mathbf{X}}_{n-1}$ (see the term $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$ in (4.52)); and ii) update $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$ by further incorporating the current low-resolution

4.4 Experimental results

frame $\mathbf{y}_n$ (see the term $\mathbf{K}_n(\mathbf{y}_n - \mathbf{H}_n \mathbf{A}_n \hat{\mathbf{X}}_{n-1})$ in (4.52).

During the above procedure, there are three motion information matrices, $\mathbf{A}_n$, $\mathbf{B}_n$ and $\mathbf{T}_{n-1,n}$, need to be computed. Suppose that the high-resolution image is with a size of $M_1 \times M_2$, $L_1 \times L_2$ for the low-resolution image. The 4×4 full-search BMA can be applied on two successive low-resolution images to estimate the half-pixel accuracy motion vectors for computing the matrix $\mathbf{B}_n$.

On the other hand, since there is no original high-resolution frames available for computing the motion information matrix $\mathbf{A}_n$ directly, it is necessary to estimate $\mathbf{A}_n$ by upsampling the current low-resolution frame, followed by applying BMA on the previously-reconstructed high-resolution and the above-mentioned resulting upsampled frame. Let $\mathbf{u}$ and $\mathbf{v}$ be the lexicographically ordered $M_1 M_2 \times 1$ vectors, representing the horizontal and vertical components of the motion vectors between two successive high-resolution images $\mathbf{X}_{n-1}$ and $\mathbf{X}_n$, respectively. Then the entry at the location (i, j) of the matrix $\mathbf{A}_n$ (denoted as $A_n(i, j)$) can be obtained via

$$A_n(i, j) = \begin{cases} 1, & \text{if } j = i - M_1 \times v(i) - u(i), i = 1, 2, \dots, M_1 \times M_2; \\ 0, & \text{otherwise.} \end{cases} \quad (4.54)$$

According to (4.54), all entries at the i -th row of the matrix $\mathbf{A}_n$ are zeros, except that the $(i - M_1 \times v(i) - u(i))$ -th entry equals to one.

Moreover, the matrix $\mathbf{T}_{n-1,n}$, which represents the backward motion information between two successive high-resolution images, is theoretically the inverse of the matrix $\mathbf{A}_n$, provided that $\mathbf{A}_n$ is non-singular. However, in general, the matrix $\mathbf{A}_n$ is singular [169]. In this chapter, the matrix $\mathbf{T}_{n-1,n}$ is computed via the pseudoinverse matrix of $\mathbf{A}_n$, that is,

$$\mathbf{T}_{n-1,n} = (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T. \quad (4.55)$$

Before getting into the details of computing $\mathbf{T}_{n-1,n}$, we firstly show that the matrix $\mathbf{A}_n^T \mathbf{A}_n$ in (4.55) is a diagonal matrix with all off-diagonal entries being zeros. The

4.4 Experimental results

entry at the location (i, j) of the matrix $\mathbf{A}_n^T \mathbf{A}_n$, denoted as $(\mathbf{A}_n^T \mathbf{A}_n)(i, j)$, can be obtained via

$$(\mathbf{A}_n^T \mathbf{A}_n)(i, j) = \sum_k A_n^T(i, k) A_n(k, j) = \sum_k A_n(k, i) A_n(k, j). \quad (4.56)$$

According to (4.54), if $i \neq j$, we have $\sum_k A_n(k, i) A_n(k, j) = 0$, since each row of the matrix $\mathbf{A}_n$ yields a unique non-zero entry (i.e., for any k , there is only a unique non-zero entry between $A_n(k, i)$ and $A_n(k, j)$). Hence, the matrix $\mathbf{A}_n^T \mathbf{A}_n$ is a diagonal matrix with all off-diagonal entries being zeros. Referring to (4.56), we have

$$\mathbf{A}_n^T \mathbf{A}_n = \text{diag} \left[\sum_k A_n(k, 1) A_n(k, 1), \dots, \sum_k A_n(k, M_1 M_2) A_n(k, M_1 M_2) \right]. \quad (4.57)$$

Then, substituting (4.57) into (4.55), $\mathbf{T}_{n-1,n}$ can be obtained as

$$\mathbf{T}_{n-1,n} = \text{diag} \left[\sum_k A_n(k, 1) A_n(k, 1), \dots, \sum_k A_n(k, M_1 M_2) A_n(k, M_1 M_2) \right]^{-1} \mathbf{A}_n^T. \quad (4.58)$$

Furthermore, the entries of the matrix $\mathbf{T}_{n-1,n}$ are set to be zeros, under the cases when $\sum_k A_n(k, i) A_n(k, i) = 0$, which means that the projected images via these zero elements have no contribution to producing the high-resolution images. Thus, according to (4.58), the entry at the location (i, j) of the matrix $T_{n-1,n}$, denoted as $T_{n-1,n}(i, j)$, can be obtained via

$$T_{n-1,n}(i, j) \approx \begin{cases} \frac{1}{\sum_k A_n(k, i) A_n(k, i)} A_n^T(i, j), & \text{if } \sum_k A_n(k, i) A_n(k, i) \neq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (4.59)$$

Besides the above three motion information matrices, there are three covariance matrices, $\mathbf{\Gamma}_{\mathbf{W}_n}$, $\mathbf{\Gamma}_{\mathbf{U}_n}$ and $\mathbf{\Gamma}_{\mathbf{V}_n}$, need to be estimated during the above-mentioned implementation procedure. To compute $\mathbf{\Gamma}_{\mathbf{W}_n}$, it is necessary to compute $\mathbf{W}_n$ first by computing the difference between the frame $\hat{\mathbf{X}}_n$ and $\mathbf{A}_n \hat{\mathbf{X}}_{n-1}$. Then the matrix $\mathbf{W}_n$ is divided into 4×4 non-overlapping blocks, which are assumed to be independent

4.4 Experimental results

with each other. Next the variance of each block is computed to compute the matrix $\Gamma_{\mathbf{W}_n}$. On the other hand, it is necessary to compute $\mathbf{U}_n$ first by computing the difference between the frame $\mathbf{y}_n$ and $\mathbf{B}_n \mathbf{y}_{n-1}$, then compute the variance of $\mathbf{U}_n$ for estimating $\Gamma_{\mathbf{U}_n}$. Similarly, $\mathbf{V}_n$ can be computed via the difference between the frame $\mathbf{y}_n$ and $\mathbf{H}_n \hat{\mathbf{X}}_n$, then its variance needs to be computed for estimating $\Gamma_{\mathbf{V}_n}$.

On the other hand, the direct computations of $\hat{\mathbf{X}}_n^-$ in (4.24) and $\hat{\mathbf{X}}_n$ in (4.26) need the matrix inversions in computing $\mathbf{K}_n$ and $\mathbf{G}_n$ (see (4.21)). However it is computationally intractable in practice due to the large dimensions of the matrices. To avoid the matrix inversions in the direct computations of $\hat{\mathbf{X}}_n^-$ and $\hat{\mathbf{X}}_n$, the *steepest descent* (SD) algorithm [167] is applied on (4.24) to compute $\hat{\mathbf{X}}_n^-$ iteratively, as follows

$$\hat{\mathbf{X}}_n^{-(m+1)} = \hat{\mathbf{X}}_n^{-(m)} + \alpha^{(m)} \mathbf{Q}_{\hat{\mathbf{X}}_n^{-(m)}}^T \mathbf{Z}_{\hat{\mathbf{X}}_n^{-(m)}}, \quad (4.60)$$

where

$$\begin{aligned} \mathbf{Q}_{\hat{\mathbf{X}}_n^{-(m)}} &= \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1}, \\ \mathbf{Z}_{\hat{\mathbf{X}}_n^{-(m)}} &= \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{y}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \hat{\mathbf{X}}_n^- - \mathbf{Q}_{\hat{\mathbf{X}}_n^{-(m)}} \hat{\mathbf{X}}_n^{-(m)}, \end{aligned} \quad (4.61)$$

and the parameter $\alpha^{(m)}$ represents the step size at the m -th iteration,

$$\alpha^{(m)} = \frac{\left\| \mathbf{Q}_{\hat{\mathbf{X}}_n^{-(m)}}^T \mathbf{Z}_{\hat{\mathbf{X}}_n^{-(m)}} \right\|^2}{\left\| \mathbf{Q}_{\hat{\mathbf{X}}_n^{-(m)}} \mathbf{Q}_{\hat{\mathbf{X}}_n^{-(m)}}^T \mathbf{Z}_{\hat{\mathbf{X}}_n^{-(m)}} \right\|^2}. \quad (4.62)$$

And the iteration is terminated when $\frac{\left\| \hat{\mathbf{X}}_n^{-(m+1)} - \hat{\mathbf{X}}_n^{-(m)} \right\|^2}{\left\| \hat{\mathbf{X}}_n^{-(m)} \right\|^2} < 10^{-6}$. Likewise, by applying the SD algorithm [167] on (4.26), $\hat{\mathbf{X}}_n$ can be computed iteratively, as follows

$$\hat{\mathbf{X}}_n^{(m+1)} = \hat{\mathbf{X}}_n^{(m)} + \alpha^{(m)} \mathbf{Q}_{\hat{\mathbf{X}}_n^{(m)}}^T \mathbf{Z}_{\hat{\mathbf{X}}_n^{(m)}}, \quad (4.63)$$

4.4 Experimental results

where

$$\begin{aligned} \mathbf{Q}_{\hat{\mathbf{x}}_n^{(m)}} &= \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T [\mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{y}_{n-1}}]^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{x}}_n}^{-1} \right], \\ \mathbf{Z}_{\hat{\mathbf{x}}_n^{(m)}} &= (\mathbf{P}_{n-1} \mathbf{F}_n)^T [\mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{y}_{n-1}}]^{-1} \mathbf{y}_{n-1} + \Gamma_{\hat{\mathbf{x}}_n}^{-1} \hat{\mathbf{x}}_n - \mathbf{Q}_{\hat{\mathbf{x}}_n^{(m)}} \hat{\mathbf{x}}_n^{(m)}, \end{aligned} \quad (4.64)$$

and the parameter $\alpha^{(m)}$ represents the step size at the m -th iteration,

$$\alpha^{(m)} = \frac{\left\| \mathbf{Q}_{\hat{\mathbf{x}}_n^{(m)}}^T \mathbf{Z}_{\hat{\mathbf{x}}_n^{(m)}} \right\|^2}{\left\| \mathbf{Q}_{\hat{\mathbf{x}}_n^{(m)}} \mathbf{Q}_{\hat{\mathbf{x}}_n^{(m)}}^T \mathbf{Z}_{\hat{\mathbf{x}}_n^{(m)}} \right\|^2}. \quad (4.65)$$

And the iteration is terminated when $\frac{\left\| \hat{\mathbf{x}}_n^{(m+1)} - \hat{\mathbf{x}}_n^{(m)} \right\|^2}{\left\| \hat{\mathbf{x}}_n^{(m)} \right\|^2} < 10^{-6}$.

4.4.2 Simulation Results

Extensive simulation experiments are conducted to demonstrate the performance of the proposed three-equation state-space filtering method and compare it with that of the *bi-cubic spline interpolation*, the SR algorithm developed in [20], and the conventional two-equation-based Kalman filtering. In [20], a sliding window moving across three consecutive low-resolution frames for producing one high-resolution image frame is implemented. Furthermore, the method [20] is applied to each block (with a same size as that used in BMA-based motion estimation).

Two kinds of experiments are conducted and presented as follows: i) the SR computation, and ii) the deblurring computation (i.e., the ratio of enlargement in the spatial domain is 1×1), each of which is conducted using both the artificially-degraded image sequence and the real-life image sequence captured by a built-in camera of a handphone.

4.4 Experimental results

The SR Computation

The first simulation experiment is conducting the SR computation using the artificially-degraded image sequence. In this case, the original image sequence is served as the high-resolution ground truth and compared with the reconstructed SR sequence for performance evaluation. Two test sequences (*Foreman* and *Bus* with a frame size of 144×176) are used in our experiments. In our simulation, the low-resolution sequence is generated from the original sequence by applying a series of operations as follows. Each frame of the original sequence is blurred by a 3×3 uniform averaging filter, followed by down-sampling in both the horizontal and vertical directions and adding with a zero-mean white Gaussian noise to yield a 25 dB signal-to-noise ratio (SNR) low-resolution image. Here the SNR is defined as $10 \times \log_{10}(\sigma_s^2/\sigma_n^2)$, where σ_s^2 and σ_n^2 are the variances of the noise-free image and the additive white noise, respectively. Figure 4.4 shows four versions of the 2×2 down-sampled 84-th frame of the *Foreman* image sequence, and likewise, for the 2×2 down-sampled 94-th frame of the *Bus* image sequence in Figure 4.5.

The above-mentioned four SR algorithms are applied individually to produce the higher-resolution sequences (100 frames each) with the enlargement ratio 2×2 and 4×4 , respectively. Note that the bi-cubic spline interpolation algorithm is applied on a single low-resolution frame only to produce one high-resolution frame. To provide an objective performance evaluation, two image quality metrics are exploited in this chapter; that is, the *PSNR* and the *structural similarity* (SSIM) [25]. The SSIM measurement is fairly consistent with the perceptual assessment, and it has been proved to yield a fairly reliable performance for evaluating the SR algorithms [170]. Note that the permissible range of the SSIM value is $[0, 1]$; a larger SSIM value indicates a better image quality. Table 4.2 and Table 4.3 compare the PSNR and SSIM performances (averaged over all 100 frames) of the above-mentioned four SR

algorithms, respectively. Figure 4.6 and Figure 4.7 show the 2×2 enlarged results of the 84-th frame of the *Foreman* image sequence and the 94-th frame of the *Bus* image sequence, respectively. As seen from the above Figures and Tables, the proposed algorithm always has the best objective and subjective performances among these four algorithms.

Note that the proposed method does *not* claim that the *estimation performance* of the conventional two-equation-based Kalman filtering has been improved. The performance improvement comes from the fact that the proposed approach has an opportunity to exploit valid information from the input data through the third equation inserted in the state-space framework.

In the second simulation experiment, the original image sequence is directly used, without going through those early-mentioned degradation processes (blurring, adding noise, and so on), as the input of the SR algorithms for producing the sequence with a higher resolution. *Foreman* and *Bus* image sequences are used in our simulation. The above-mentioned four SR algorithms are applied individually to produce the high-resolution sequence (100 frames) with the enlargement ratio 2×2 and 4×4 , respectively. Since there is no ground truth for performance evaluation in this experiment, two perceptual image quality measurement metrics (without the need for the ground truth) are exploited to provide the objective performance evaluation; that is, the *variance*-based scheme [26] and the *no-reference quality assessment* (NRQA) [27], which has been proved to yield a fairly reliable performance for evaluating the SR algorithms [170]. Note that a larger variance value or a larger NRQA value represents a better image quality. Table 4.4 and Table 4.5 compares the variance and NRQA performances (averaged over all 100 frames) of the above-mentioned four SR algorithms, respectively. Figure 4.8 and Figure 4.9 show the 2×2 enlarged results of the 84-th frame of the *Foreman* image sequence and the 94-th frame of the *Bus* image sequence, respectively. As seen

4.4 Experimental results

from the above Figures and Tables, the proposed algorithm also achieves the best performance among these four algorithms in all experimental cases.

The third SR performance evaluations are conducted for two real-life image sequences, which are captured by a built-in camera of a handphone (LGTM; model KU970) using its functionalities “Multishot” and “Black & White”. The above two sequences are denoted as *License-Plate* and *EEE-Board*, and each sequence consists of nine images, with a frame size of 80×80 and 100×160 , respectively. The above-mentioned four SR algorithms are applied individually to each sequence and produce the higher-resolution sequence (also 9 frames each) with the enlargement ratio 2×2 and 4×4 , respectively. In our simulations, the camera’s PSF is assumed to be a 3×3 Gaussian low-pass filter with a standard deviation of 1 for each frame of the captured sequences.

Since there is no ground truth available for conducting objective performance evaluation in these experiments, two perceptual image quality measurement metrics (without the need for the ground truth) are exploited in this chapter. They are the *variance*-based scheme [26] and the *no-reference quality assessment* (NRQA) [27]. Table 4.6 and Table 4.7 compare the variance and NRQA performances (averaged over all 9 frames) of the above-mentioned four SR algorithms, respectively. As observed from these Tables, our derived three-equation-based state-space filtering consistently delivers better objective performance than the conventional two-equation-based Kalman filtering. On the other hand, to demonstrate the superior subjective performance of the proposed method, Figure 4.10 shows the 2×2 enlarged results of the 4-th frame of the *License-Plate* image sequence. As seen from Figure 4.10, the proposed algorithm consistently achieves better visual quality than the conventional two-equation-based Kalman filtering. More specifically, the last “6” of the plate number “8667” can not be confidently recognized in the conventional Kalman filtering’s result (see Figure 4.10 (e)), while it is much more recognizable in our

result (see Figure 4.10 (f)). Such level of differentiation might be extremely crucial to forensic application. Likewise, the above-concluded comments are also applicable to the results of the 6-th frame of the *EEE-Board* image sequence, as shown in Figure 4.11, from which one can clearly see the difference of reconstructed “a” of “Electrical” by using two methods.

The Deblurring Computation

Experiments are conducted to show that the proposed three-equation-based state-space filtering also addresses the deblurring problem (i.e., the ratio of enlargement in the spatial domain is 1×1 , the downsampling matrix $\mathbf{D}_n$ becomes an identity matrix) with attractive performance. Two kinds of experiments are conducted using: i) the artificially-blurred image sequence, and ii) the real-life image sequence captured by a handphone’s built-in camera as described previously.

In the first simulation experiment, the original image sequences *Foreman* and *Bus* are served as the high-resolution ground truth and compared with the deblurred results produced by two SR algorithms for performance evaluation (with respect to the ground truth). The blurred sequence is generated by convolving the original sequence with a 3×3 uniform averaging filter; such operation is applied to each frame of the original sequence independently. The above-mentioned blurring function is assumed to be known in advance for two SR algorithms. Tables 4.8 and 4.9 compare their resulted PSNR and SSIM [25] performances (averaged over all 100 frames), respectively. Figures 4.12 and 4.13 show the deblurred results of the 84-th frame of the *Foreman* image sequence and the 94-th frame of the *Bus* image sequence, respectively.

In the second simulation experiment, two real-life image sequences, *License-Plate* and *EEE-Board*, are used as the test data. The camera’s PSF is assumed

4.4 Experimental results

to be a 3×3 Gaussian low-pass filter with a standard deviation of 1; such assumption is applied for each frame of the captured sequences. Tables 4.10 and 4.11 compare their resulted average variance [26] and NRQA [27] performances (averaged over all 9 frames). Figures 4.14 and 4.15 show the deblurred results of the 7-th frame of the *License-Plate* image sequence and the 3-rd frame of the *EEE-Board* image sequence, respectively. As seen from the above two experiments, the proposed approach consistently outperforms the conventional two-equation-based Kalman filtering on dealing with the blurring problem.

Justification of The Independence Between $\mathbf{V}_n$ (in (4.4)) and $\mathbf{U}_n$ (in (4.5))

Experiments are conducted to justify that $\mathbf{V}_n$ (in (4.4)) is independent of $\mathbf{U}_n$ (in (4.5)), at each discrete-time instance n . For that, the aim is to justify that their joint *probability density function* (pdf) (i.e., $p(\mathbf{V}_n, \mathbf{U}_n)$) is equal to the product of their respective pdfs (i.e., $p(\mathbf{V}_n)$ and $p(\mathbf{U}_n)$); that is

$$p(\mathbf{V}_n, \mathbf{U}_n) = p(\mathbf{V}_n) \times p(\mathbf{U}_n). \quad (4.66)$$

The above three pdfs $p(\mathbf{V}_n, \mathbf{U}_n)$, $p(\mathbf{V}_n)$ and $p(\mathbf{U}_n)$ in (4.66) can be computed by fitting both the original image sequence (i.e., $\mathbf{X}_1, \mathbf{X}_2, \dots$) and the simulated low-resolution image sequence (i.e., $\mathbf{y}_1, \mathbf{y}_2, \dots$) to the model defined in (4.4) and (4.5). Denote d_n as the absolute difference between $p(\mathbf{V}_n, \mathbf{U}_n)$ and $p(\mathbf{V}_n) \times p(\mathbf{U}_n)$ as

$$d_n = \sum_{\mathbf{V}_n} \sum_{\mathbf{U}_n} |p(\mathbf{V}_n, \mathbf{U}_n) - p(\mathbf{V}_n) \times p(\mathbf{U}_n)|. \quad (4.67)$$

Note that the permissible range of d_n is $[0, 1]$. Furthermore, note that d_n defined in (4.67) can serve as a criterion to measure how independent between $\mathbf{V}_n$ and $\mathbf{U}_n$; a smaller d_n value represents a better independence between $\mathbf{V}_n$ and $\mathbf{U}_n$ [171].

To determine whether $\mathbf{V}_n$ and $\mathbf{U}_n$ are independent from each other, we explore the absolute difference d_n between $p(\mathbf{V}_n, \mathbf{U}_n)$ and $p(\mathbf{V}_n) \times p(\mathbf{U}_n)$ using the *Foreman*

4.4 Experimental results

Table 4.1
THE ABSOLUTE DIFFERENCE d_n (SEE (4.67)) BETWEEN $p(\mathbf{V}_n, \mathbf{U}_n)$ AND
 $p(\mathbf{V}_n) \times p(\mathbf{U}_n)$.

Test sequence	2 × 2 SR computation		4 × 4 SR computation	
	<i>Foreman</i>	<i>Bus</i>	<i>Foreman</i>	<i>Bus</i>
d_n defined in (4.67)	0.15	0.05	0.03	0.04

and *Bus* sequences used in our first SR computation experiment. The experimental results are presented in Table 4.1, where one can see d_n is fairly small. Therefore, $\mathbf{V}_n$ and $\mathbf{U}_n$ can be considered as independent from each other.

Computational Time

All of four SR approaches mentioned above are implemented using the *Matlab* programming language and run on a PC with a Pentium 2.6 GHz CPU and a 4096 MB RAM. One hundred experiments are conducted for each of the above-mentioned four SR approaches, and their respective average run times are compared in Tables 4.12, 4.13 and 4.14 for above-mentioned three SR computation simulations, respectively.

Memory Requirement

In this section, the memory requirement of the proposed framework and that of the above-mentioned three SR approaches are compared. Throughout the following analysis, it is assumed that a high-resolution image sequence (each frame has a size of $M_1 \times M_2$) needs to be reconstructed by performing 2 × 2 SR computation using a low-resolution image sequence (each frame has a size of $L_1 \times L_2$, where $L_1 = M_1/2$ and $L_2 = M_2/2$).

A study of the memory requirements of these four methods is provided as follows:

- The *bi-cubic interpolation* approach utilizes a sliding window moving across

4.4 Experimental results

one low-resolution frame for producing one high-resolution image frame. Therefore, it requires storing one low-resolution frame (i.e., one $L_1 \times L_2$ matrix).

- The SR approach [20] utilizes a sliding window moving across three consecutive low-resolution frames for producing one high-resolution image frame. Therefore, it requires storing three low-resolution images (i.e., three $L_1 \times L_2$ matrices), plus two motion information matrices, each of which represents the motion information between each two consecutive low-resolution frames.
- The conventional two-equation-based Kalman filtering approach reconstructs one high-resolution frame by using the previous low-resolution frame (i.e., one $L_1 \times L_2$ matrix), the current low-resolution frame (i.e., one $L_1 \times L_2$ matrix) and the previously-reconstructed high-resolution frame (i.e., one $M_1 \times M_2$ matrix). Furthermore, it requires storing one motion information matrix $\mathbf{A}_n$, and one motion information matrix $\mathbf{B}_n$, plus four variance matrices $\Gamma_{\hat{\mathbf{x}}_{n-1}}$, $\Gamma_{\hat{\mathbf{y}}_n}$, $\Gamma_{\mathbf{v}_n}$, $\Gamma_{\mathbf{w}_n}$, each of which is a sparse matrix, plus one matrix $\mathbf{H}_n$, which is a sparse matrix and each column contains only $p^2 L_1 L_2 / M_1 M_2 = p^2 / 4$ non-zero entries.
- The proposed three-equation-based state-space filtering approach reconstructs one high-resolution frame by using the current low-resolution frame (i.e., one $L_1 \times L_2$ matrix) and the previously-reconstructed high-resolution frame (i.e., one $M_1 \times M_2$ matrix). Furthermore, it requires storing one motion information matrix $\mathbf{A}_n$, and four variance matrices $\Gamma_{\hat{\mathbf{x}}_{n-1}}$, $\Gamma_{\hat{\mathbf{y}}_n}$, $\Gamma_{\mathbf{v}_n}$, $\Gamma_{\mathbf{w}_n}$, each of which is a sparse matrix, plus one matrix $\mathbf{H}_n$, which is a sparse matrix and each column contains only $p^2 L_1 L_2 / M_1 M_2 = p^2 / 4$ non-zero entries.

4.4 Experimental results

Table 4.2
THE AVERAGE PSNR PERFORMANCE COMPARISON ON THE SR COMPUTATION USING
THE ARTIFICIALLY-DEGRADED IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	2×2	27.52 dB	27.98 dB	29.24 dB	29.76 dB
	4×4	21.07 dB	22.94 dB	23.51 dB	23.92 dB
<i>Bus</i>	2×2	22.94 dB	21.85 dB	23.81 dB	24.18 dB
	4×4	18.02 dB	17.08 dB	19.42 dB	19.71 dB

Table 4.3
THE AVERAGE SSIM [25] PERFORMANCE COMPARISON ON THE SR COMPUTATION
USING THE ARTIFICIALLY-DEGRADED IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	2×2	0.81	0.85	0.87	0.88
	4×4	0.69	0.74	0.77	0.78
<i>Bus</i>	2×2	0.72	0.66	0.77	0.80
	4×4	0.49	0.44	0.50	0.52

Table 4.4
THE AVERAGE VARIANCE [26] PERFORMANCE COMPARISON ON THE SR COMPUTATION
USING THE ORIGINAL IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	2×2	2666.06	2674.98	2692.28	2840.73
	4×4	2652.31	2660.99	2660.96	2683.83
<i>Bus</i>	2×2	2372.46	2348.45	2449.21	2821.80
	4×4	2369.12	2348.45	2395.69	2447.74

4.4 Experimental results

Table 4.5

THE AVERAGE NRQA [27] PERFORMANCE COMPARISON ON THE SR COMPUTATION USING THE ORIGINAL IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	2×2	43.97	44.96	45.06	48.83
	4×4	32.58	32.16	33.17	33.66
<i>Bus</i>	2×2	40.13	38.49	43.07	50.13
	4×4	32.67	30.49	34.10	34.65

Table 4.6

THE AVERAGE VARIANCE [26] PERFORMANCE COMPARISON ON THE SR COMPUTATION USING THE REAL-LIFE IMAGE SEQUENCE CAPTURED BY A BUILT-IN CAMERA OF A HANDPHONE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>License-Plate</i>	2×2	2297.22	2313.81	2329.35	2572.16
	4×4	2298.86	2183.26	2387.76	2423.07
<i>EEE-Board</i>	2×2	1257.81	1263.83	1279.67	1376.60
	4×4	1258.57	1263.98	1354.61	1362.13

Table 4.7

THE AVERAGE NRQA [27] PERFORMANCE COMPARISON ON THE SR COMPUTATION USING THE REAL-LIFE IMAGE SEQUENCE CAPTURED BY A BUILT-IN CAMERA OF A HANDPHONE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>License-Plate</i>	2×2	44.24	46.36	53.36	56.31
	4×4	38.29	38.19	48.08	50.54
<i>EEE-Board</i>	2×2	53.93	55.83	65.67	68.02
	4×4	38.93	43.72	60.78	63.29

4.4 Experimental results

Table 4.8

THE AVERAGE PSNR PERFORMANCE COMPARISON ON THE DEBLURRING COMPUTATION USING THE ARTIFICIALLY DEGRADED IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	1×1	30.64 dB	31.21 dB
<i>Bus</i>	1×1	25.36 dB	26.35 dB

Table 4.9

THE AVERAGE SSIM [25] PERFORMANCE COMPARISON ON THE DEBLURRING COMPUTATION USING THE ARTIFICIALLY DEGRADED IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	1×1	0.89	0.90
<i>Bus</i>	1×1	0.82	0.85

Table 4.10

THE AVERAGE VARIANCE [26] PERFORMANCE COMPARISON ON THE DEBLURRING COMPUTATION USING THE REAL-LIFE VIDEO SEQUENCE CAPTURED BY A BUILT-IN CAMERA OF A HANDPHONE.

Test sequence	Enlargement ratio	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>License-Plate</i>	1×1	2520.53	2758.87
<i>EEE-Board</i>	1×1	1335.84	1458.11

Table 4.11

THE AVERAGE NRQA [27] PERFORMANCE COMPARISON ON THE DEBLURRING COMPUTATION USING THE REAL-LIFE VIDEO SEQUENCE CAPTURED BY A BUILT-IN CAMERA OF A HANDPHONE.

Test sequence	Enlargement ratio	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>License-Plate</i>	1×1	77.40	78.88
<i>EEE-Board</i>	1×1	59.08	60.91

4.4 Experimental results

Figure 4.4. Four versions of the 84-th frame of the *Foreman* image sequence: (a) the original image X (144×176); (b) a blurred image X_B , after convolving X with a 3×3 uniform averaging filter; (c) the down-sampled image X_D , after down-sampling X_B by a factor of two in both the horizontal and the vertical directions; and (d) the simulated low-resolution image, after adding the zero-mean white Gaussian noise to X_D . (For demonstration, images (c) and (d) are scaled up to the same size as that of X through pixel duplication.)

4.4 Experimental results

Figure 4.5. Four versions of the 94-th frame of the *Bus* image sequence: (a) the original image X (144×176); (b) a blurred image X_B , after convolving X with a 3×3 uniform averaging filter; (c) the down-sampled image X_D , after down-sampling X_B by a factor of two in both the horizontal and the vertical directions; and (d) the simulated low-resolution image, after adding the zero-mean white Gaussian noise to X_D . (For demonstration, images (c) and (d) are scaled up to the same size as that of X through pixel duplication.)

4.4 Experimental results

Figure 4.6. Various 2×2 enlarged results of the 84-th frame of the *Foreman* image sequence using the artificially-degraded image sequence: (a) the bi-cubic spline interpolation approach (PSNR = 27.49 dB); (b) Vandewalle *et al.*'s algorithm [20] (PSNR = 28.19 dB); (c) the conventional two-equation-based Kalman filtering (PSNR = 29.20 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 29.76 dB).

4.4 Experimental results

Figure 4.7. Various 2×2 enlarged results of the 94-th frame of the *Bus* image sequence using the artificially-degraded image sequence: (a) the bi-cubic spline interpolation approach (PSNR = 21.64 dB); (b) Vandewalle *et al.*'s algorithm [20] (PSNR = 21.38 dB); (c) the conventional two-equation-based Kalman filtering (PSNR = 23.86 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 24.39 dB).

4.4 Experimental results

Figure 4.8. Various 2×2 enlarged results (in CIF) of the 84-th frame of the *Foreman* image sequence using the original image sequence (QCIF frame size): (a) the bi-cubic spline interpolation approach; (b) Vandewalle *et al.*'s algorithm [20]; (c) the conventional two-equation-based Kalman filtering; and (d) the proposed three-equation-based state-space filtering.

4.4 Experimental results

Figure 4.9. Various 2×2 enlarged results (in CIF) of the 94-th frame of the *Bus* image sequence using the original image sequence (QCIF frame size): (a) the bi-cubic spline interpolation approach; (b) Vandewalle *et al.*'s algorithm [20]; (c) the conventional two-equation-based Kalman filtering; and (d) the proposed three-equation-based state-space filtering.

4.4 Experimental results

Figure 4.10. Various 2×2 enlarged results of the 4-th frame of the *License-Plate* image sequence: (a) original image (80×80); (b) the conventional two-equation-based Kalman filtering; (c) the proposed three-equation-based state-space filtering; image (d)-(f) are the zoomed parts of images (a)-(c) using pixel duplication, respectively. (For demonstration, images (a) and (d) are scaled up through pixel duplication.)

4.4 Experimental results

Figure 4.11. Various 2×2 enlarged results of the 6-th frame of the *EEE-Board* image sequence: (a) original image (100×160); (b) the conventional two-equation-based Kalman filtering; (c) the proposed three-equation-based state-space filtering; image (d)-(f) are the zoomed parts of images (a)-(c) using pixel duplication, respectively. (For demonstration, images (a) and (d) are scaled up through pixel duplication.)

4.4 Experimental results

Figure 4.12. Various deblurred results of the 84-th frame of the *Foreman* image sequence: (a) the original image; (b) the simulated blurring image; (c) the conventional two-equation-based Kalman filtering (PSNR = 30.63 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 31.24 dB).

4.4 Experimental results

Figure 4.13. Various deblurred results of the 94-th frame of the *Bus* image sequence: (a) the original image; (b) the simulated blurring image; (c) the conventional two-equation-based Kalman filtering (PSNR = 25.32 dB); and (d) the proposed three-equation-based state-space filtering (PSNR = 26.59 dB).

4.4 Experimental results

Figure 4.14. Various deblurred results of the 7-th frame of the *License-Plate* image sequence: (a) original image; (b) the conventional two-equation-based Kalman filtering; and (c) the proposed three-equation-based state-space filtering.

Figure 4.15. Various deblurred results of the 3-rd frame of the *EEE-Board* image sequence: (a) original image; (b) the conventional two-equation-based Kalman filtering; and (c) the proposed three-equation-based state-space filtering.

4.4 Experimental results

Table 4.12
THE RUN TIME (AVERAGED OVER ALL 100 FRAMES) COMPARISON ON THE SR
COMPUTATION USING THE ARTIFICIALLY-DEGRADED IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	2×2	0.04 seconds	2.29 seconds	1.05 seconds	1.32 seconds
	4×4	0.08 seconds	3.07 seconds	1.19 seconds	1.35 seconds
<i>Bus</i>	2×2	0.04 seconds	2.24 seconds	1.09 seconds	1.38 seconds
	4×4	0.08 seconds	3.03 seconds	1.69 seconds	1.87 seconds

Table 4.13
THE RUN TIME (AVERAGED OVER ALL 100 FRAMES) COMPARISON ON THE SR
COMPUTATION ON THE SR COMPUTATION USING THE ORIGINAL IMAGE SEQUENCE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>Foreman</i>	2×2	0.17 seconds	4.18 seconds	1.97 seconds	3.01 seconds
	4×4	1.38 seconds	7.52 seconds	5.11 seconds	6.76 seconds
<i>Bus</i>	2×2	0.18 seconds	4.26 seconds	2.71 seconds	5.39 seconds
	4×4	1.42 seconds	7.28 seconds	10.46 seconds	11.42 seconds

Table 4.14
THE RUN TIME (AVERAGED OVER ALL 9 FRAMES) COMPARISON ON THE SR
COMPUTATION USING THE REAL-LIFE IMAGE SEQUENCE CAPTURED BY A BUILT-IN
CAMERA OF A HANDPHONE.

Test sequence	Enlargement ratio	Bi-cubic spline interpolation	Vandewalle <i>et al.</i> 's method [20]	Conventional two-equation-based Kalman filtering	Proposed three-equation-based state-space filtering
<i>License-Plate</i>	2×2	0.10 seconds	2.65 seconds	0.99 seconds	1.34 seconds
	4×4	0.89 seconds	4.67 seconds	2.04 seconds	2.24 seconds
<i>EEE-Board</i>	2×2	0.04 seconds	1.06 seconds	0.51 seconds	0.65 seconds
	4×4	0.32 seconds	1.84 seconds	2.24 seconds	2.74 seconds

4.5 Conclusions

In this chapter, a *three*-equation-based state-space filtering is mathematically derived to perform SR video reconstruction. The proposed approach inserts an extra observation equation into the conventional *two*-equation-based Kalman framework to further exploit the information provided by the previously-observed low-resolution frame. Due to more observed information being jointly considered, the performance of the proposed method is always superior to that of the conventional two-equation-based Kalman filtering, provided that the extra information can be properly incorporated into the estimation. Furthermore, a unified analysis is provided to show that the proposed framework can be boiled down to sequentially computing the MAP estimation of individual higher-resolution frame, based on two adjacent low-resolution frames each time, if the temporal correlations among the high-resolution frames are omitted. On the other hand, the conventional two-equation-based Kalman filtering becomes a special case of the proposed scheme, if the previously-observed low-resolution frame is neglected during the estimation of each unknown high-resolution frame.

Appendix 4.A Derivation of (4.19)

Figure 4.16. The relationship between the states $X_1, X_2, \dots, X_n$ and the observations $y_1, y_2, \dots, y_n$: (a) at the discrete-time step n ; (b) at the discrete-time step $n - 1$.

Appendix 4.A Derivation of (4.19)

A detailed proof of (4.19) is provided in this section.

The proposed algorithm specifies a joint probability distribution over the states $X_1, X_2, \dots, X_n, \dots$ and the observations $y_1, y_2, \dots, y_n, \dots$. This relationship is illustrated in Figure 4.16 (a). Let the set $\{X_n, y_n\}$ separate the set of the past states and observations, P_n , which is defined as $P_n = \{X_1, X_2, \dots, X_{n-1}, y_1, y_2, \dots, y_{n-1}\}$, and the set of the future states and observations, F_n , which is defined as $F_n = \{X_{n+1}, \dots, y_{n+1}, \dots\}$.

Consequently,

$$P(P_n, y_n, X_n, F_n) = P(P_n, X_n) \cdot P(y_n | P_n, X_n) \cdot P(F_n | P_n, X_n, y_n), \quad (4.68)$$

which is the foundation for the development of the following proofs. The objective is to incorporate the state X_{n-1} and the observation y_n into (4.68) to obtain the update equation (4.19).

According to (4.3)-(4.5) and the first-order Markov assumption, which means that $P(y_n | P_n, X_n) = P(y_n | y_{n-1}, X_n)$ and $P(y_{n-1} | P_{n-1}, X_{n-1}) = P(y_{n-1} | y_{n-2}, X_{n-1})$,

Appendix 4.A Derivation of (4.19)

the following equation can be extracted from (4.68)

$$\begin{aligned}
 & P(P_n, \mathbf{y}_n, \mathbf{X}_n, F_n) \\
 &= \int_{\mathbf{X}_{n-1}} P(P_{n-1}, \mathbf{X}_{n-1} = \mathbf{s}, \mathbf{y}_{n-1}, \mathbf{X}_n) d\mathbf{X}_{n-1} \cdot P(\mathbf{y}_n | P_n, \mathbf{X}_n) \cdot P(F_n | P_n, \mathbf{X}_n, \mathbf{y}_n) \\
 &= \int_{\mathbf{X}_{n-1}} P(\mathbf{X}_n | \mathbf{X}_{n-1} = \mathbf{s}) \cdot P(\mathbf{y}_{n-1} | \mathbf{y}_{n-2}, \mathbf{X}_{n-1} = \mathbf{s}) \cdot P(\mathbf{X}_{n-1} = \mathbf{s}, P_{n-1}) d\mathbf{X}_{n-1} \\
 &\quad \cdot P(\mathbf{y}_n | \mathbf{y}_{n-1}, \mathbf{X}_n) \cdot P(F_n | P_n, \mathbf{X}_n, \mathbf{y}_n) \\
 &= T(\mathbf{X}_n) \cdot P(F_n | P_n, \mathbf{X}_n, \mathbf{y}_n),
 \end{aligned}$$

where

$$T(\mathbf{X}_n) = \int_{\mathbf{X}_{n-1}} P(\mathbf{X}_n | \mathbf{X}_{n-1} = \mathbf{s}) \cdot P(\mathbf{y}_{n-1} | \mathbf{y}_{n-2}, \mathbf{X}_{n-1} = \mathbf{s}) \cdot P(\mathbf{X}_{n-1} = \mathbf{s}, P_{n-1}) d\mathbf{X}_{n-1} \cdot P(\mathbf{y}_n | \mathbf{y}_{n-1}, \mathbf{X}_n). \quad (4.69)$$

Now applying the *logarithmic opinion pool* [165] on $P(\mathbf{y}_n | \mathbf{X}_n)$ and $P(\mathbf{y}_n | \mathbf{X}_n)$, $P(\mathbf{y}_n | \mathbf{y}_{n-1}, \mathbf{X}_n)$ can be computed as

$$P(\mathbf{y}_n | \mathbf{y}_{n-1}, \mathbf{X}_n) \sim N(\mu_{\mathbf{y}_n} \mathbf{X}_n, \Gamma_{\mathbf{y}_n}), \quad (4.70)$$

where

$$\begin{aligned}
 \Gamma_{\mathbf{y}_n}^{-1} &= \lambda_x \Gamma_{\mathbf{V}_n}^{-1} + \lambda_y \Gamma_{\mathbf{U}_n}^{-1}, \\
 \mu_{\mathbf{y}_n} &= \Gamma_{\mathbf{y}_n} (\lambda_x \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n \mathbf{X}_n + \lambda_y \Gamma_{\mathbf{U}_n}^{-1} \mathbf{B}_n \mathbf{y}_{n-1}) = \Gamma_{\mathbf{y}_n} (\lambda_x \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n \mathbf{X}_n + \lambda_y \Gamma_{\mathbf{U}_n}^{-1} \mathbf{B}_n \mathbf{H}_n \mathbf{X}_{n-1}),
 \end{aligned} \quad (4.71)$$

in which λ_x and λ_y are the pooling weights such that $\lambda_x, \lambda_y \geq 0$ and $\lambda_x + \lambda_y = 1$.

Denote the matrix $\mathbf{T}_{n-1,n}$ as the backward motion information matrix between $\mathbf{X}_n$ and $\mathbf{X}_{n-1}$; that is, $\mathbf{X}_{n-1} = \mathbf{T}_{n-1,n} \mathbf{X}_n$. Hence, (4.71) can be rewritten as

$$\mu_{\mathbf{y}_n} = \mathbf{P}_n \mathbf{X}_n, \quad (4.72)$$

where

$$\mathbf{P}_n = \Gamma_{\mathbf{y}_n} (\lambda_x \Gamma_{\mathbf{V}_n}^{-1} \mathbf{H}_n + \lambda_y \Gamma_{\mathbf{U}_n}^{-1} \mathbf{B}_n \mathbf{H}_n \mathbf{T}_{n-1,n}). \quad (4.73)$$

Appendix 4.A Derivation of (4.19)

Moreover, we have the following lemma of conditional Gaussian probability density function [171]. Let $X \sim N(X; \mu, \sigma^2)$ denote the Gaussian process X with mean μ and variance σ^2 . Suppose we have two jointly Gaussian vectors: $\mathbf{X} \sim N(\mathbf{X}; \mu_X, \mathbf{V}_X)$ and $\mathbf{y} = \mathbf{A}\mathbf{X} + \mathbf{n}$, where $\mathbf{n} \sim N(\mathbf{n}; \mathbf{0}, \mathbf{V}_y)$, then

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{y} \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{X} \\ \mathbf{y} \end{bmatrix}; \begin{bmatrix} \mu_X \\ \mathbf{A}\mu_X \end{bmatrix}, \begin{bmatrix} \mathbf{V}_X & \mathbf{V}_X \mathbf{A}^T \\ \mathbf{A}\mathbf{V}_X & \mathbf{A}\mathbf{V}_X \mathbf{A}^T + \mathbf{V}_y \end{bmatrix} \right).$$

Since $P(\mathbf{y}|\mathbf{X})P(\mathbf{X}) = P(\mathbf{X}|\mathbf{y})P(\mathbf{y})$, we get

$$\begin{aligned} & N(\mathbf{y}; \mathbf{A}\mathbf{X}, \mathbf{V}_y)N(\mathbf{X}; \mu_X, \mathbf{V}_X) \\ &= N(\mathbf{X}; \mu_X + \mathbf{K}(\mathbf{y} - \mathbf{A}\mu_X), (\mathbf{I} - \mathbf{K}\mathbf{A})\mathbf{V}_X)N(\mathbf{y}; \mathbf{A}\mu_X, \mathbf{A}\mathbf{V}_X \mathbf{A}^T + \mathbf{V}_y), \end{aligned} \quad (4.74)$$

where $\mathbf{K} = \mathbf{V}_X \mathbf{A}^T (\mathbf{A}\mathbf{V}_X \mathbf{A}^T + \mathbf{V}_y)^{-1}$.

Now, applying (4.74) to (4.69) twice, (4.69) becomes

$$\begin{aligned} T(\mathbf{X}_n) &= N(\mathbf{y}_n; f(\mathbf{H}_n \mathbf{X}_n, \mathbf{B}_n \mathbf{y}_{n-1}), g(\Gamma_{\mathbf{V}_n}, \Gamma_{\mathbf{U}_n})) \\ &\quad \cdot \int_{\mathbf{X}_{n-1}} (N(\mathbf{X}_n; \mathbf{A}_n \mathbf{X}_{n-1}, \Gamma_{\mathbf{W}_n}) \cdot N(\mathbf{y}_{n-1}; f(\mathbf{H}_n \mathbf{X}_{n-1}, \mathbf{B}_{n-1} \mathbf{y}_{n-2}), g(\Gamma_{\mathbf{V}_{n-1}}, \Gamma_{\mathbf{U}_{n-1}})) \\ &\quad \cdot N(\mathbf{X}_{n-1}; \mu_{\mathbf{X}_{n-1}}, \Gamma_{\mathbf{X}_{n-1}})) d\mathbf{X}_{n-1} \\ &= N(\mathbf{y}_n; f(\mathbf{H}_n \mathbf{X}_n, \mathbf{B}_n \mathbf{y}_{n-1}), g(\Gamma_{\mathbf{V}_n}, \Gamma_{\mathbf{U}_n})) \\ &\quad \cdot \int_{\mathbf{X}_{n-1}} (N(\mathbf{X}_n; \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}, \mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \\ &\quad \cdot N(\mathbf{y}_{n-1}; f(\mathbf{H}_n \mathbf{X}_{n-1}, \mathbf{B}_{n-1} \mathbf{y}_{n-2}), g(\Gamma_{\mathbf{V}_{n-1}}, \Gamma_{\mathbf{U}_{n-1}})) \\ &\quad \cdot N(\mathbf{X}_{n-1}; \mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}), (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}})) d\mathbf{X}_{n-1} \\ &= N(\mathbf{y}_n; f(\mathbf{H}_n \mathbf{X}_n, \mathbf{B}_n \mathbf{y}_{n-1}), g(\Gamma_{\mathbf{V}_n}, \Gamma_{\mathbf{U}_n})) \cdot N(\mathbf{X}_n; \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}, \mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \\ &\quad \cdot \int_{\mathbf{X}_{n-1}} (N(\mathbf{y}_{n-1}; f(\mathbf{H}_n \mathbf{X}_{n-1}, \mathbf{B}_{n-1} \mathbf{y}_{n-2}), g(\Gamma_{\mathbf{V}_{n-1}}, \Gamma_{\mathbf{U}_{n-1}})) \\ &\quad \cdot N(\mathbf{X}_{n-1}; \mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}), (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}})) d\mathbf{X}_{n-1}, \end{aligned} \quad (4.75)$$

where $\mathbf{F}_n = \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n})^{-1}$, f and g are two functions in *logarithmic opinion pool*. Suppose we have two Gaussian distributions $N(\mu_A, \Gamma_A)$ and

Appendix 4.A Derivation of (4.19)

$N(\mu_B, \Gamma_B)$, then

$$\begin{aligned} g(\Gamma_A, \Gamma_B) &= (\lambda_A \Gamma_A^{-1} + \lambda_B \Gamma_B^{-1})^{-1}, \\ f(\mu_A, \mu_B) &= g(\Gamma_A, \Gamma_B)(\lambda_A \Gamma_A \mu_A + \lambda_B \Gamma_B \mu_B), \end{aligned} \quad (4.76)$$

in which λ_A and λ_B are the pooling weights such that $\lambda_A, \lambda_B \geq 0$ and $\lambda_A + \lambda_B = 1$.

The integral in (4.75) becomes

$$\begin{aligned} & \int_{\mathbf{X}_{n-1}} (N(\mathbf{y}_{n-1}; f(\mathbf{H}_n \mathbf{X}_{n-1}, \mathbf{B}_{n-1} \mathbf{y}_{n-2}), g(\Gamma_{\mathbf{V}_n}, \Gamma_{\mathbf{U}_n})) \\ & \cdot N(\mathbf{X}_{n-1}; \mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}), (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}})) d\mathbf{X}_{n-1} \\ &= \int_{\mathbf{X}_{n-1}} (N(\mathbf{y}_{n-1}; \mathbf{P}_{n-1}(\mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}})), \mu_{y_{n-1}}(\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mu_{y_{n-1}}^T + \Gamma_{y_{n-1}})) \\ & \cdot N(\mathbf{X}_{n-1}; \mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}) + \mathbf{G}_{\mathbf{A}_n}(\mathbf{y}_{n-1} - \mathbf{P}_{n-1}(\mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}))) \\ & (\mathbf{I} - \mathbf{G}_{\mathbf{A}_n} \mathbf{P}_n)(\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}})) d\mathbf{X}_{n-1}, \end{aligned} \quad (4.77)$$

where $\mathbf{G}_{\mathbf{A}_n} = (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mathbf{P}_{n-1}^T [\mathbf{P}_{n-1}(\mathbf{I} - \mathbf{F}_n) \mathbf{P}_{n-1}^T + \Gamma_{y_{n-1}}]^{-1}$.

Then substituting (4.77) into (4.75), (4.75) becomes

$$\begin{aligned} T(\mathbf{X}_n) &= N(\mathbf{y}_n; \mathbf{P}_n \mathbf{X}_n, \Gamma_{y_n}) \cdot N(\mathbf{X}_n; \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}, \mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \\ & \cdot N(\mathbf{y}_{n-1}; \mathbf{P}_{n-1}(\mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n(\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}})), \mathbf{P}_{n-1}(\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{y_{n-1}}). \end{aligned} \quad (4.78)$$

Appendix 4.A Derivation of (4.19)

Applying (4.74) to (4.78) twice, (4.78) can be rewritten as

$$\begin{aligned}
T(\mathbf{X}_n) &= N(\mathbf{y}_n; \mathbf{P}_n \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}, \mathbf{P}_n (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T + \Gamma_{\mathbf{y}_n}) \\
&\quad \cdot N(\mathbf{X}_n; \mathbf{A}_n \mu_{\mathbf{X}_{n-1}} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}), (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n})) \\
&\quad \cdot N(\mathbf{y}_{n-1}; \mathbf{P}_{n-1} (\mu_{\mathbf{X}_{n-1}} + \mathbf{F}_n (\mathbf{X}_n - \mathbf{A}_n \mu_{\mathbf{X}_{n-1}})), \mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{y}_{n-1}}) \\
&= N(\mathbf{y}_n; \mathbf{P}_n \mathbf{A}_n \mu_{\mathbf{X}_{n-1}}, \mathbf{P}_n (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T + \Gamma_{\mathbf{y}_n}) \\
&\quad \cdot N(\mathbf{y}_{n-1}; \mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{A}_n \mu_{\mathbf{X}_{n-1}} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \mu_{\mathbf{X}_{n-1}})), \\
&\quad \mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{y}_{n-1}}) \\
&\quad \cdot N(\mathbf{X}_n; \mathbf{A}_n \mu_{\mathbf{X}_{n-1}} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \mu_{\mathbf{X}_{n-1}})) \\
&\quad + \mathbf{G}_n (\mathbf{y}_{n-1} - \mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{A}_n \mathbf{X}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \mathbf{X}_{n-1}))), \\
&\quad (\mathbf{I} - \mathbf{G}_n \mathbf{P}_{n-1} \mathbf{F}_n) (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n})), \tag{4.79}
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{K}_n &= (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T [\mathbf{P}_n (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T + \Gamma_{\mathbf{y}_n}]^{-1}, \\
\mathbf{G}_n &= (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\
&\quad \cdot [\mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\
&\quad + \mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{y}_{n-1}}]^{-1}.
\end{aligned}$$

Now it is easy to get the following update equations from the third term of (4.79),

$$\begin{aligned}
\mu_{\mathbf{X}_{n-1}} &= \\
&\mathbf{A}_n \mathbf{X}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \mathbf{X}_{n-1}) + \mathbf{G}_n (\mathbf{y}_{n-1} - \mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{A}_n \mathbf{X}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{P}_n \mathbf{A}_n \mathbf{X}_{n-1}))), \\
\Gamma_{\mathbf{X}_n} &= (\mathbf{I} - \mathbf{G}_n \mathbf{P}_{n-1} \mathbf{F}_n) (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}),
\end{aligned}$$

Appendix 4.B Derivations of (4.34) and (4.36)

where

$$\begin{aligned} \mathbf{F}_n &= \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n})^{-1}, \\ \mathbf{K}_n &= (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T [\mathbf{P}_n (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) \mathbf{P}_n^T + \Gamma_{\mathbf{Y}_n}]^{-1}, \\ \mathbf{G}_n &= (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\ &\quad \cdot [\mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\mathbf{X}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{W}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\ &\quad + \mathbf{P}_{n-1} (\mathbf{I} - \mathbf{F}_n \mathbf{A}_n) \Gamma_{\mathbf{X}_{n-1}} \mathbf{P}_{n-1}^T + \Gamma_{\mathbf{Y}_{n-1}}]^{-1}. \end{aligned}$$

Appendix 4.B Derivations of (4.34) and (4.36)

In this section, the detailed derivation of (4.34) and (4.36), from (4.24) and (4.26), respectively, will be presented. In the development of the following proofs, we will use (4.29)-(4.31) and a lemma of the inverse matrix [166],

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1}. \quad (4.80)$$

We firstly present the derivation of (4.34) from (4.24). Applying (4.80) to (4.24), we can get

$$\begin{aligned} \hat{\mathbf{X}}_n^- &= \hat{\mathbf{X}}_n^{--} + \Gamma_{\hat{\mathbf{X}}_n^{--}} \mathbf{P}_n^T (\Gamma_{\mathbf{Y}_n} + \mathbf{P}_n \Gamma_{\hat{\mathbf{X}}_n^{--}} \mathbf{P}_n^T)^{-1} (\mathbf{y}_n - \mathbf{P}_n \hat{\mathbf{X}}_n^{--}) \\ &= \hat{\mathbf{X}}_n^{--} + \Gamma_{\hat{\mathbf{X}}_n^{--}} \mathbf{P}_n^T \left(\Gamma_{\mathbf{Y}_n}^{-1} - \Gamma_{\mathbf{Y}_n}^{-1} \mathbf{P}_n (\mathbf{P}_n^T \Gamma_{\mathbf{Y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{X}}_n^{--}}^{-1})^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{Y}_n}^{-1} \right) (\mathbf{y}_n - \mathbf{P}_n \hat{\mathbf{X}}_n^{--}) \\ &= T_1 \hat{\mathbf{X}}_n^{--} + T_2 \mathbf{y}_n, \end{aligned} \quad (4.81)$$

where

$$T_1 = \mathbf{I} - \Gamma_{\hat{\mathbf{X}}_n^{--}} \mathbf{P}_n^T \left(\Gamma_{\mathbf{Y}_n}^{-1} - \Gamma_{\mathbf{Y}_n}^{-1} \mathbf{P}_n (\mathbf{P}_n^T \Gamma_{\mathbf{Y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{X}}_n^{--}}^{-1})^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{Y}_n}^{-1} \right) \mathbf{P}_n, \quad (4.82)$$

$$T_2 = \Gamma_{\hat{\mathbf{X}}_n^{--}} \mathbf{P}_n^T \left(\Gamma_{\mathbf{Y}_n}^{-1} - \Gamma_{\mathbf{Y}_n}^{-1} \mathbf{P}_n (\mathbf{P}_n^T \Gamma_{\mathbf{Y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{X}}_n^{--}}^{-1})^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{Y}_n}^{-1} \right). \quad (4.83)$$

Appendix 4.B Derivations of (4.34) and (4.36)

Next, (4.82) and (4.83) can be further simplified as

$$\begin{aligned}
 T_1 &= \mathbf{I} - \left(\Gamma_{\hat{\mathbf{x}}_n^-} \left(\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right) - \Gamma_{\hat{\mathbf{x}}_n^-} \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n \right) \left(\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right)^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n \\
 &= \mathbf{I} - \left[\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right]^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n \\
 &= \left[\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right]^{-1} \left(\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} - \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n \right) \\
 &= \left[\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right]^{-1} \Gamma_{\hat{\mathbf{x}}_n^-}^{-1}, \tag{4.84}
 \end{aligned}$$

$$\begin{aligned}
 T_2 &= \left(\Gamma_{\hat{\mathbf{x}}_n^-} \left(\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right) - \Gamma_{\hat{\mathbf{x}}_n^-} \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n \right) \left(\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right)^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \\
 &= \left[\mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1} \mathbf{P}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right]^{-1} \mathbf{P}_n^T \Gamma_{\mathbf{y}_n}^{-1}. \tag{4.85}
 \end{aligned}$$

Finally, incorporating (4.29) and (4.30) into (4.84) and (4.85), (4.81) can be simplified as

$$\hat{\mathbf{x}}_n^- = \left(\mathbf{H}_n^T \Gamma_{\mathbf{v}_n}^{-1} \mathbf{H}_n + \Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \right)^{-1} \cdot \left(\Gamma_{\hat{\mathbf{x}}_n^-}^{-1} \hat{\mathbf{x}}_n^- + \mathbf{H}_n^T \Gamma_{\mathbf{v}_n}^{-1} \mathbf{y}_n \right). \tag{4.86}$$

Now, we move to the derivation of (4.36) from (4.26). In order to simplify (4.26), it is necessary to simplify $\mathbf{F}_n$ and $\mathbf{G}_n$ first. By substituting (4.31) into (4.21), $\mathbf{F}_n$ can be simplified as

$$\mathbf{F}_n = \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T \right)^{-1}. \tag{4.87}$$

According to (4.87), we have

$$\begin{aligned}
 \mathbf{I} - \mathbf{F}_n \mathbf{A}_n &= \mathbf{I} - \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T \right)^{-1} \mathbf{A}_n \\
 &= \mathbf{I} - \left(\mathbf{A}_n^T \mathbf{A}_n \right)^{-1} \mathbf{A}_n^T \mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T \left(\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T \right)^{-1} \mathbf{A}_n \\
 &= \mathbf{I} - \left(\mathbf{A}_n^T \mathbf{A}_n \right)^{-1} \mathbf{A}_n^T \mathbf{A}_n = \mathbf{0}. \tag{4.88}
 \end{aligned}$$

Then substituting (4.88) into (4.21), $\mathbf{G}_n$ can be simplified as

$$\begin{aligned}
 \mathbf{G}_n &= (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T \\
 &\quad [\mathbf{P}_{n-1} \mathbf{F}_n (\mathbf{I} - \mathbf{K}_n \mathbf{P}_n) (\mathbf{A}_n \Gamma_{\hat{\mathbf{x}}_{n-1}} \mathbf{A}_n^T + \Gamma_{\mathbf{w}_n}) (\mathbf{P}_{n-1} \mathbf{F}_n)^T + \Gamma_{\mathbf{y}_{n-1}}]^{-1}. \tag{4.89}
 \end{aligned}$$

Appendix 4.B Derivations of (4.34) and (4.36)

Now substituting (4.89) into (4.26) and applying (4.80), we have

$$\begin{aligned}\hat{\mathbf{X}}_n &= \hat{\mathbf{X}}_n^- + \Gamma_{\hat{\mathbf{X}}_n^-} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \left(\Gamma_{\mathbf{y}_{n-1}}^{-1} - \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} \right. \\ &\quad \left. \cdot (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \right) (\mathbf{y}_{n-1} - \mathbf{P}_{n-1} \mathbf{F}_n \hat{\mathbf{X}}_n^-) \\ &= T_3 \hat{\mathbf{X}}_n^- + T_4 \mathbf{y}_{n-1},\end{aligned}\tag{4.90}$$

where

$$\begin{aligned}T_3 &= \mathbf{I} - \Gamma_{\hat{\mathbf{X}}_n^-} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \left(\Gamma_{\mathbf{y}_{n-1}}^{-1} - \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} \right. \\ &\quad \left. (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \right) \mathbf{P}_{n-1} \mathbf{F}_n,\end{aligned}\tag{4.91}$$

$$\begin{aligned}T_4 &= \Gamma_{\hat{\mathbf{X}}_n^-} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \left(\Gamma_{\mathbf{y}_{n-1}}^{-1} - \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} \right. \\ &\quad \left. (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \right).\end{aligned}\tag{4.92}$$

Next, (4.91) and (4.92) can be further simplified as

$$\begin{aligned}T_3 &= \mathbf{I} - \left(\Gamma_{\hat{\mathbf{X}}_n^-} \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right] - \Gamma_{\hat{\mathbf{X}}_n^-} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \right) \\ &\quad \cdot \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \\ &= \mathbf{I} - \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \\ &= \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} \\ &\quad \cdot \left((\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} - (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \right) \\ &= \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} \Gamma_{\hat{\mathbf{X}}_n^-}^{-1},\end{aligned}\tag{4.93}$$

$$\begin{aligned}T_4 &= \left(\Gamma_{\hat{\mathbf{X}}_n^-} \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right] - \Gamma_{\hat{\mathbf{X}}_n^-} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n \right) \\ &\quad \cdot \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \\ &= \left[(\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1} \mathbf{P}_{n-1} \mathbf{F}_n + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right]^{-1} (\mathbf{P}_{n-1} \mathbf{F}_n)^T \Gamma_{\mathbf{y}_{n-1}}^{-1}.\end{aligned}\tag{4.94}$$

Finally, incorporating (4.29)-(4.31) into (4.93) and (4.94), (4.90) can be simplified as

$$\begin{aligned}\hat{\mathbf{X}}_n &= \left(\left(\mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \right)^T \Gamma_{\mathbf{v}_{n-1}}^{-1} \mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T + \Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \right)^{-1} \\ &\quad \cdot \left(\Gamma_{\hat{\mathbf{X}}_n^-}^{-1} \hat{\mathbf{X}}_n^- + (\mathbf{H}_n (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T)^T \Gamma_{\mathbf{v}_{n-1}}^{-1} \mathbf{y}_{n-1} \right).\end{aligned}\tag{4.95}$$

Chapter 5

Conclusions and Future Works

In this chapter, we summarize the contributions of this thesis and present some directions for future research.

5.1 Conclusions

SR imaging has been one of the most spotlighted research areas in image processing, since it can overcome the inherent hardware limitation of the imaging system and improve the performance of several digital image processing applications. It has been proved to be effective in a wide variety of application areas, such as face recognition [172–174], remote sensing [175–177], security and surveillance [178,179], medical imaging [180–183]. In view of this, a comprehensive review on the existing SR imaging research works is provided in Chapter 2.

Motivated by the fact that the SR image computation is, in essence, an ill-posed inverse problem, a Bayesian inference framework is mathematically formulated in thesis based on both the observed low-resolution images and the prior model of unknown high-resolution image. The major challenge of the above Bayesian-based approach is that the formulated probability distribution of unknown high-resolution

5.1 Conclusions

image is usually very complicate; thus, its statistics is quite difficult to compute. To tackle this challenge, the MCMC technique has been successfully introduced in Chapter 3 to conduct stochastic SR image reconstruction, and the convergence bound of the proposed MCMC SR approach has been mathematically derived. Furthermore, the proposed SR approach produces the unknown high-resolution image with the simultaneous-adapted prior image model (through the automatically-updated hyperparameter in the prior image model). Extensive experimental results are presented to show that the proposed method not only achieves better performance than the bi-cubic interpolation method, but also yields fairly close performance to the case where the optimal prior image model is exploited. In addition, we propose an outlier-sensitive bilateral filtering method as a ‘complementary’ post-processing scheme to further enhance the reconstructed high-resolution image, by removing the artifacts that are only incurred due to insufficient number of samples used in the MCMC SR algorithm. The proposed outlier-sensitive bilateral filtering method yields superior performance than the conventional bilateral filtering approach, as verified by the experimental results.

The SR video approaches aim to reconstruct an image sequence with a high-resolution from a group of adjacent low-resolution image frames. In view of that the conventional *two*-equation-based Kalman filter only considers the information from the previously-reconstructed high-resolution frame and the currently-observed low-resolution frame for producing each higher-resolution frame, a three-equation-based state-space filtering approach is mathematically derived in Chapter 4 by inserting an extra observation equation into the conventional two-equation-based Kalman framework to further exploit the information provided by the previously-observed low-resolution frames for sequentially producing the high-resolution sequence. Due to more information brought in and with the correct mathematical derivation, the

performance of the proposed method is always superior to the conventional two-equation-based Kalman filtering method. Furthermore, a unified analysis is provided to show that the proposed framework boils down to sequentially computing the MAP estimation of individual higher-resolution image, based on two adjacent low-resolution images each time, if the temporal correlations among the high-resolution images are omitted; on the other hand, the proposed scheme becomes the conventional two-equation-based Kalman filtering approach, if the previously-observed low-resolution frame is neglected during the estimation of the unknown high-resolution frame at each time step.

5.2 Future Works

Based on the research in this thesis, we provide the following promising areas that need to be studied more thoroughly in the future.

SR image reconstruction

In Chapter 3, a stochastic SR image approach is proposed by exploiting the Bayesian inference and the MCMC technique based on both the observed low-resolution images and the prior model of unknown high-resolution image. The proposed stochastic SR image approach is able to simultaneously update the hyperparameter value, along with the updating of the image sample at each sample-generation step throughout the proposed MCMC process. This is motivated by the fact that the hyperparameter serves a key role to control the degree of regularization, which is imposed by the prior image model, and controls the quality of the reconstructed high-resolution image. However, a *global* hyperparameter is used for the whole image in our developed approach. In view of this, a potential direction is to adaptively assign different hyperparameter for various image regions.

5.2 Future Works

The second issue is how to shorten the burn-in period of the developed MCMC SR process so as to further shorten the run time of the proposed MCMC approach. A good attempt has been provided in Section 3.2.4 by using a bi-cubic interpolated version of any one of the low-resolution input images as the initial image sample, rather than starting with a randomly-generated image. On the other hand, some recently-developed fast MCMC techniques, such as the hybrid Monte Carlo technique [184,185], can be considered to be incorporated into our developed framework. It could shorten the burn-in period of the MCMC process and converge to the target distribution more rapidly by exploiting the gradient information of the target distribution.

The third issue is to investigate which fundamental issues control the performance of SR image computation. The performance of SR computation is mainly relied on a number of factors [186], including interpolation error, pixel-domain quantization error, motion estimation error and optical-blur error [187–191]. Furthermore, perceptual image quality measurement has been developed for evaluating the performance of various SR algorithms [170,192]. On the other hand, A thorough study of SR performance limits would understand the fundamental limits of the SR imaging, thereby helping practitioners to find the correct balance between expensive optical imaging systems and image reconstruction algorithms. Baker and Kanade [193] showed that, for a large enough resolution enhancement factor, any smoothness prior image model will result in reconstructions with very little high-frequency content. Lin *et al.* studied a numerical perturbation model of reconstruction-based SR algorithms for the case of translational motion [194] and the learning-based SR algorithms [195]. Robinson and Milanfar [196] analyzed this issue from statistical principles using *Cramer-Rao* bounds.

The forth issue is development of new grid computing infrastructure particularly for stochastic image processing simulations, as well as design of new schedule

algorithm for executing image processing tasks over grid computing facilities. On the other hand, most SR approaches that are developed in the literature have fairly large computational loads, since they usually employ an iterative scheme. To reduce this computational cost, many attempts have been done to develop fast SR algorithms for particular application cases. Elad and Hel-Or [197] proposed a fast SR case where the motions among the low-resolution images are pure translational, and the blurring incurred in the image acquisition is space-invariant. Zhang *et al.* [198] proposed a parallel SR approach, where the multiple input low-resolution images are split into several sub-sets according to their translation relations, and then an MAP estimation algorithm is applied to each subset to reconstruct a high-resolution image, finally, the high-resolution images obtained by various sub-set of low-resolution images are fused to a single final high-resolution image result. Investigation of how to implement the above-mentioned fast SR approaches in grid computing environment is a good direction for future research.

SR video reconstruction

A three-equation-based state-space approach is mathematically developed in Chapter 4 to exploit the information from the previously-reconstructed high-resolution frame, the currently-observed low-resolution frame as well as the previously-observed low-resolution frame for sequentially producing the high-resolution frames. However, the future low-resolution frames are neglected, which can be further incorporated to improve the quality of the reconstructed high-resolution image. By this way, the high-resolution frames could be reconstructed with some frames delay, but with possible improved quality.

On the other hand, various motion estimation techniques could be considered to be incorporated into our developed state-space approach. Schultz *et al.* [168] discussed various motion estimation methods, including the parametric models,

5.2 Future Works

block matching [199] and optical flow [200–204]. In view of this, a good attempt is to exploit the optical flow technique [205], since it can provide more accurate motion estimation parameters than the block matching algorithm that is currently used in our developed approach.

There are several very important technical challenges that are closely related with the SR problem; that is, registration and blur identification. On the other hand, SR imaging can be further expanded to not only obtain an high-resolution representation of the spatial resolution of scene intensity, but also to other purposes [206], such as temporal resolution enhancement and compression artifacts suppression, respectively. All of the above-mentioned directions are presented in detail as follows.

Registration

Image registration is an essential stage for the reconstruction-based SR approaches, since the multiple low-resolution frames need to be aligned well so that the spatio-temporal information provided by these frames can be well exploited. Nevertheless, it is difficult to achieve accurate registration in the low-resolution, which will give rise to suboptimal results in the reconstructed high-resolution images [9].

This motivates the study of improving the registration accuracy. The first way is to investigate how to obtain accurate registration result for blurred and noisy image data. This is a challenging task, accurate registration in the low-resolution remains difficult to achieve, due to the presence of aliasing, blur and noise in the captured low-resolution images [207]. Several attempts have been done to employ an iterative scheme to estimate the HR image and the registration parameters alternately [50, 84, 128]. This is contrast to that conventional SR algorithms commonly perform two sequential stages: an image registration stage followed by an inverse estimation that integrates image fusion and deblurring into a single step. Ng

5.2 Future Works

et al. [38] proposed a regularized constrained total least-squares method to minimize a non-convex cost function, which is projected onto the image and registration parameter domains one at a time, and minimized iteratively. Chung *et al.* [208] proposed a nonlinear cost function, and estimated the registration parameters and the high-resolution image using the Gauss-Newton method. Robinson *et al.* [209] proposed to estimate the registration parameters by minimizing a *variable projection* function. Then, the estimates of the motion parameters are used to reconstruct the high-resolution image using Bayesian estimation. He *et al.* [210, 211] exploited a *nonlinear least squares* (NLS) method that is fairly robust to the error incurred in the initialization of registration parameters.

5.2 Future Works

The second way is to take into account the impact of unreliable registration parameters. A SR image reconstruction algorithm has been proposed by Lee and Kang [212] to address the inaccurate registration parameters, based on the assumption that the registration error can be modeled as a Gaussian noise with a standard deviation proportional to the degree of the registration error. Jiang *et al.* [72] proposed to detect and discard bad matchings in the initial alignment maps obtained using robust optical flow algorithm. Hence, the quality of super-resolved images can be improved by rejecting such errors. Eren *et al.* [213] proposed a robust objected-based SR approach by exploiting only objects (regions), where reliable motion parameters can be utilized. Zomet *et al.* [214] used the median filter in the iterative SR reconstruction to produce SR images by discarding measurements that are inconsistent with the imaging model. Ng and Bose mathematically analyzed how the registration error can result in inaccurate reconstructed high-resolution image [215]. Trimeche *et al.* [216] proposed an adaptive filtering method to reject the outlier image regions. The adaptation criterion is determined by the error between the median of the samples from the low-resolution images and the output of their proposed filter. El-Yamany and Papamichalis [217] proposed to suppresses the outliers due to violations of the assumed observation model by incorporating a robust error norm into the objective function, and adapting the estimation process to each of the low-resolution frames.

Blur Identification

The identification of the blur incurred in the image acquisition is important to the SR problem, although it is assumed to be known in most of SR algorithms developed in the literature. However, the blurs are usually unknown in practice. Therefore, this motivates the study of performing the high-resolution image reconstruction with limited knowledge or no knowledge of the blurring function [218–221].

Figure 5.1. A conceptual framework of simultaneous resolution enhancement in both the spatial domain and the temporal domain.

For that, Shekarforoush and Chellappa [222] used a data-driven regularization method, where the PSF are assumed unknown; hence, they are required to be estimated from the observed data. He *et al.* [223, 224] exploited a MAP estimation framework for tackling the SR computation with space-invariant blur, by imposing prior information on the unknown blur parameter. Then, an alternating minimization method is developed to estimate the blur parameter and high-resolution image progressively. Yang *et al.* [225] proposed to use Lanczos algorithm and Gauss quadrature theory to alternatively estimate the blur and high-resolution image. Despite the space-invariant blur addressed in the above research works [222–225], various kinds of blurs have been investigated in the literature, such as the space-variant blur [226], motion blur [227], time-variant blur [129].

Simultaneous Super-resolution in Spatial Domain and Temporal Domain

The research works presented in this thesis are focused on resolution enhancement in spatial domain. Recently, the challenge of simultaneous resolution enhancement in the time domain as well as the space domain has received growing attention [228–231]. Temporal resolution refers to the number of frames captured per second and is also commonly known as the frame rate. It is related to the amount of perceptible motion between the frames. Higher frame rates result in less smearing

5.2 Future Works

due to movements in the scene. The typical frame rate suitable for a pleasing view is about 25 frames per second or above. Robertson and Stevenson [228] considered to increase the frame rate of video compressed by inserting images between received frames of the sequence. They further exploited a Bayesian framework to prevent the spatial compression artifacts in the received frame from propagating to the adjacent reconstructed high-resolution frames. Shechtman *et al.* [229] proposed a joint space-time SR framework by using multiple image sequence which has different spatial resolution and different frame rate. Watanabe *et al.* [230] utilized the redundant discrete wavelet transform to develop a method for obtaining a high spatio-temporal resolution video from two video sequences. The first sequence has a high spatial resolution but a low frame rate, while the second sequence has a low spatial resolution but a high frame rate. Both of these two sequences are captured by a dual sensor camera that can capture these sequences with the same field of view simultaneously.

Simultaneous Super-resolution and Demosaicing

The SR approaches proposed in this thesis are focused on tackling SR computation for gray images only. Recently, some attempts have been done on performing SR computation for *color filter array* (CFA) data, where the aim is to produce higher-resolution image from its CFA counterparts, by recovering the full color information for each pixel position and increasing the spatial resolution at the same time. The most commonly used CFA pattern is the Bayer pattern, by which the captured CFA image contains only one of the three color components (i.e., red, green and blue) at each pixel position. The conventional CFA demosaicing (or spelled as demosaicking) aims to produce a color image with full color information at each pixel positions by interpolating missing color information [232, 233]. Gotoh and Okutomi [234] proposed to address the demosaicing and SR problems within a unified context

by utilizing an Tikhonov regularization framework. Farsiu *et al.* [235] developed a hybrid method of SR and demosaicing, based on a MAP estimation technique by minimizing a multiple term cost function. The L_1 norm is used for measuring the difference between the projected estimate of the high-resolution image and each low-resolution image; while a bilateral regularization is used for spatially regularizing the luminance component, which results in sharp edges and forcing interpolation along the edges and not across them, while exploiting Tikhonov regularization on the chrominance components. Shao *et al.* [236] proposed a partition-based interpolation approach for performing joint demosaicing and SR by estimating missing pixels through linear operations on neighboring pixels that exploit the captured image structure. Mega *et al.* [237] addressed a SR computation case where the blurring parameters need to be estimated from the input image. Gevrekci *et al.* [238] developed a *projections onto convex sets* (POCS) approach for Bayer-sampled image sequence [239, 240].

Simultaneous Super-resolution and Compression Artifacts Suppression

Most of the SR video research works described before are based on the assumption that the low-resolution images are available in the time domain and there is no compression involved. However, compression has become a standard component of almost every data communication application. In view of this, Bayesian-based SR video algorithms have been extended to address compressed video data [159, 241–243]. Wang *et al.* [244] improved the method developed in [241] by providing more accurate registration performance and better initial estimation of the high-resolution image, where the registration is done based on the upsampled versions of the low-resolution image, rather than directly on the low-resolution image. Park *et al.* [245] considered the spatially adaptive properties for the DCT compressed images. Gunturk and Altunbasak [246, 247] proposed a SR method in the

5.2 Future Works

discrete cosine transform (DCT) domain, and exploited the quantization information in the compressed video stream to establish the constraint sets of the desired high-resolution images [248], since the quantization error introduced during the compression process is often the dominant source of error when the compression ratio is high. Shao [249] proposed a SR method based on the classification of local image patterns using both structure information and activity measure. The structure information is represented by adaptive dynamic-range coding, while the activity measure is the combination of local entropy and dynamic range. On the other hand, a unified treatment for image restoration, image deinterlacing and the SR computation was proposed by Martins and Forchhammer [250].

On the other hand, the principle of the SR technique can be utilized for improving the image compression performance, motivated by that the downsampling operation can be applied during image encoding, while the SR techniques can be applied to upsample the decoded image during image decoding. More specifically, the original (high-resolution) image sequence can be downsampled before they are encoded, and then the decoded image sequence is imposed by SR techniques for generating an high-resolution image sequence (with the same resolution as that of the original image sequence) [251]. For that, Barreto *et al.* [252] proposed to use an image segmentation stage to ‘guide’ the above-mentioned downsampling and SR stages. To be more specific, image segmentation is performed at the encoder to classify their blocks into three different types according to their respective characteristics (e.g., motion and texture). These obtained segmentation information is used to define the downsampling process at the encoder, then it is encoded and provided to the decoder as side information for guiding the SR reconstruction process.

Appendix A Grid Computing

Enabled Super-resolution Imaging

A.1 Introduction

In Chapter 3, the *Markov chain Monte Carlo* (MCMC) technique has been successfully introduced first time for addressing the SR problem with attractive performance [24]. The developed MCMC SR algorithm provides a flexible framework for the incorporation of the prior knowledge of the unknown high-resolution image, as well as for the modeling of the dependence between the observed low-resolution images and the unknown high-resolution image. Despite its powerful performance, the MCMC SR approach requires substantial amounts of computational resources, for it needs to generate a fairly large number of image samples for producing high-resolution images. To be more specific, suppose that the MCMC SR algorithm is exploited to produce one high-resolution gray image with a dimension of 256×256 , by utilizing four low-resolution gray images with a dimension of 128×128 each. Note that the dynamic range of the image's gray level is $[0, 255]$ for the gray images. The MCMC SR approach needs to generate 8.9455×10^7 image samples with a size of 256×256 each [24]. Furthermore, since the *hyperparameter* of the prior image model used in the MCMC SR approach intimately controls the degree of the regularity and the quality of the reconstructed high-resolution image, the MCMC

Appendix A Grid Computing Enabled Super-resolution Imaging

SR approach requires exhaustive search over a hyperparameter set specified to obtain an optimal value for the prior image model. All these lend the MCMC SR approach developed in [24] with fairly large computational time complexity.

Grid computing appears as a promising technology to handle the intensive computation requirements of the above-mentioned MCMC SR problem. Grid computing is typically characterized by the cooperation among the distributed resources, such as the computational processors and the data storages, to constitute a high-performance computational environment [253]. The basic idea of grid computing is to perform many independent computations at once using multiple computation processes, rather than doing these computations sequentially using a single computation process only. Obviously, grid computing provides a tremendous computing power for enabling the computationally-complicated algorithms. Motivated by this, grid computing has gained a widespread popularity in the signal processing research community [254], especially for those requiring huge amounts of computation powers and data storages, such as medical image processing [255–258], 3-D image registration [259], remote sensing image processing [260], and evolutionary computation [261, 262].

In this appendix, grid computing is introduced to tackle the computationally-intensive MCMC SR problem. The proposed grid-based MCMC SR approach firstly divides the complicated MCMC SR task into a number of independent tasks, which are much easier to be handled individually. These independent tasks are then dispatched to many different compute nodes available in the grid computing service. Finally, the respective results obtained from different tasks are assembled to generate the final high-resolution image result for the entire MCMC SR task.

A.2 Proposed Grid-based Super-resolution Imaging

Summarizing the previously-presented key-points in Chapter 3, the MCMC SR approach requires fairly huge computational resources, due to the following two

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.1. An overview of image samples generation procedure.

reasons: i) it needs to generate a fairly huge number of image samples for estimating the unknown high-resolution image; that is, N defined in (3.27) is fairly large; ii) it needs exhaustive search over a hyperparameter set (i.e., the user-defined set $S = \{\lambda_1, \lambda_2, \dots, \lambda_L\}$) for obtaining the optimal hyperparameter value used for the prior image model. The image sample generation process is illustrated in Figure A.1. To summarize, the MCMC SR approach with exhaustive search over a range of hyperparameter values can be stated as below.

- For each hyperparameter value λ_i in the user-defined hyperparameter set $S = \{\lambda_1, \lambda_2, \dots, \lambda_L\}$,
 - Randomly generate $\mathbf{X}^{(1)}$ as initialization.
 - For $t = 1 : N - 1$,
 - * Generate a new image sample $\mathbf{X}^{(t+1)}$ using the algorithm described in Chapter 3.
 - Discard the initial T samples and compute the high-resolution image $\hat{\mathbf{X}}_{\lambda_i}$ for the current hyperparameter value λ_i .
- Compare the results obtained by different hyperparameter values (i.e., $\hat{\mathbf{X}}_{\lambda_1}, \hat{\mathbf{X}}_{\lambda_2}, \dots, \hat{\mathbf{X}}_{\lambda_L}$), and then select the best one (in terms of the PSNR performance) as the final high-resolution image result.

To tackle the above two computation challenges, grid computing offers attractive solution.

Appendix A Grid Computing Enabled Super-resolution Imaging

A.2.1 Grid Computing

Grid computing provides a high-performance computing environment through a set of distributed computational processors and data storages. Figure A.2 illustrates a typical grid computing structure, where a user can access and submit tasks to the grid computing facilities via the Internet. Typically, a grid computing system consists of multiple *clusters*. Each cluster consists of a group of loosely-coupled computers, called the *nodes*, which work together corporately so that they can be viewed as a single virtual computer. These nodes have different functionalities. The *head node* acts as an interface of the cluster, while the *compute node* is responsible for computing the distributed tasks. All of the nodes within a cluster share the same data storage and are connected to each other through fast *local area network* (LAN).

Over the above-mentioned grid computing structure, many independent computation tasks, which do not have to share data during the execution of the computation processes, can be run on different clusters at the same time. For that, many software toolkits have been developed to manage the assignment of the tasks. Among them, the *Globus* toolkit [263] is a popular software for providing a set of basic grid services, such as computational and storage resource management, security protocols, and remote data access facilities. It allows users to easily access the grid computing facilities for developing the grid-based applications.

A.2.2 Proposed Grid-based MCMC SR Framework

To tackle the intensive MCMC SR computation problem as described in Chapter 3, the idea is to split the complicated MCMC SR task (see Figure A.3 (a)) into several small and independent tasks (see Figure A.3 (b)) for parallel processing in the grid computing environment. More specifically, rather than trying out all possible integer-valued hyperparameters using a single MCMC SR task, Figure A.3 (b) demonstrates the proposed grid-based MCMC SR framework with a set of

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.2. A general framework of the grid computing structure.

MCMC SR tasks, each task is responsible for exploiting the MCMC SR algorithm to produce a high-resolution image based on one fixed, but different, hyperparameter value. Their computed results (i.e., the high-resolution images that are obtained by using different hyperparameter values) are compared to select the best one (in terms of PSNR) as the final result for the entire MCMC SR task.

An overview of the proposed grid-based MCMC SR framework is demonstrated in Figure A.4, which contains the following seven modules: *task splitting*, *resources information collection*, *tasks scheduling (meta-scheduler)*, *tasks scheduling (local scheduler)*, *tasks execution*, *tasks storing* and *tasks collection*. Their functionalities are briefly described as follows.

- *Task splitting*: splitting the original MCMC SR task into a number of independent and relatively small tasks.
- *Resources information collection*: assigning the MCMC SR tasks to the available clusters in the grid computing facilities. It firstly needs to obtain a list of the available clusters by exploiting the *Globus Monitoring and Discovery System (MDS)*.
- *Tasks scheduling (meta-scheduler)*: assigning the MCMC SR tasks to the

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.3. (a) A single MCMC SR task that needs to sequentially exploit all the hyperparameter values permissible in the hyperparameter set chosen. (b) A set of independent MCMC SR tasks that are dispatched in the grid computing environment for parallel computing, in which each task runs under one fixed, but different, hyperparameter value.

Appendix A Grid Computing Enabled Super-resolution Imaging

available clusters, based on the information collected by the MDS.

- *Tasks scheduling (local scheduler)*: dispatching the sub-sets of MCMC SR tasks across different available compute nodes.
- *Tasks execution*: exploiting the computation resources provided by the compute nodes to executing the tasks.
- *Tasks storing*: exploiting the storage resources for providing the capability of storing the program codes, data, and intermediate results (i.e., the image samples generated during the MCMC SR process).
- *Task collection*: gathering the results obtained from many tasks. In our case, many high-resolution images are obtained, each of which is under a different hyperparameter value for the prior image model. Among them, the best one (in terms of the PSNR) will be selected as the final high-resolution image result of the entire MCMC SR task.

To summarize, the proposed grid-based MCMC SR approach starts with dividing the MCMC SR task into a number of independent tasks. The meta-scheduler then dispatches them to different available clusters, according to the availability information collected by the MDS. Next, the local scheduler assigns a sub-set of the MCMC SR tasks over different available compute nodes. Finally, the results obtained from different tasks are compared and the best one (in terms of PSNR) is selected as the final high-resolution image result for the entire MCMC SR task.

A.3 Experimental Results

To study the performance of the proposed grid-based MCMC framework, four test images are used in our experiments; that is, 256×256 *Airplane*, 256×256 *Boat*, 256×256 *Bridge* and 200×200 *Text*, as shown in Figure A.5. In our simulation, four low-resolution images are independently generated from the original image by

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.4. An overview of the proposed grid-based MCMC SR framework.

applying a series of operations on each test image, respectively, as follows. We first apply a shift, which is drawn from a continuous uniform distribution over the interval $(-2, 2)$ in the units of pixels, followed by applying a rotation, which is drawn from a continuous uniform distribution over the interval $[-4, 4]$ in the units of degrees. Next, the processed images are convolved with a point spread function, which is a Gaussian low-pass filter with a size of 4×4 and the standard derivation of 2, followed by down-sampling with a factor of two in both the horizontal and the vertical directions, respectively. Lastly, the processed image is added with a zero-mean Gaussian white noise to yield a 25 dB *signal-to-noise ratio* (SNR) image. Here the SNR is defined as $10 \times \log_{10}(\sigma_s^2/\sigma_v^2)$, where σ_s^2 and σ_v^2 are the variances of the noise-free image and the additive white noise, respectively.

The proposed grid-based MCMC SR framework is implemented and tested on

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.5. Four test images used in the experiments: (a) Airplane; (b) Boat; (c) Bridge; and (d) Text.

the grid computing environment offered by *National Grid Office* (NGO), Singapore¹. The grid computing facilities we used is composed of six compute nodes, each of which comprises two Intel Xeon 2.4G Hz CPUs running *Sun Grid Engine* (SGE) 5.3 as the local scheduler for each cluster.

Figure A.6 demonstrates the work-flow of the proposed grid-based MCMC SR approach. To access the grid computing facilities, the user needs to request and create a certificate via *Grid Security Infrastructure* (GSI) first, which is an authentication process for the user to use the computing and storage resources available in the grid computing service. Upon the successful request of the certificate to access the grid computing service over the Internet, the proposed grid-based MCMC SR task is divided into fifty equal tasks, each of which is responsible for exploiting one hyperparameter value only. In each task, the MCMC SR algorithm described in Chapter 3 is then applied to produce a high-resolution image. Finally, the respective results of different tasks (i.e., the high-resolution images that are obtained by using different hyperparameter values) are compared to select the best one (in terms of the PSNR) as the final result of the entire MCMC SR task.

In our case, the permissible hyperparameter value range is $[1, 50]$, and only integer values are used. This range is experimentally selected and has been proved

¹The detailed information can be found on the website of NGO <http://www.ngpp.ngp.org.sg/ngppresources.html>.

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.6. The work-flow of the proposed grid-based MCMC SR framework.

to yield fairly good performance in our simulations. Furthermore, in each of the above-mentioned tasks, the MCMC SR algorithm described in Chapter 3 firstly randomly generates the initial image sample and then computes the convergence bound T via (3.27). In our simulation, the parameter ϵ is set to be 10, and γ is set to be 4. The MCMC technique is applied to generate N samples, where the parameter N is defined by the user and it should be large enough to meet the statistics of the target distribution. In our simulation, N is selected to be $10 \times T$. Next, the first generated T samples are removed as they are considered as unreliable samples, and the follow-up generated $9 \times T$ samples are used to produce a high-resolution image.

Experiments are conducted to compare the computational time of the proposed grid-based MCMC SR approach with that of the single-PC-based MCMC SR approach developed in [24]. The later approach is implemented on a single stand-alone PC with a Pentium 2.4 GHz CPU and a 512 MB RAM. With the benefit from grid computing, the computational time of the MCMC SR approach has been effectively

Appendix A Grid Computing Enabled Super-resolution Imaging

Figure A.7. The comparison of computational time.

reduced by 85.10%, 84.78%, 85.01% and 85.54% for the above four test images, respectively, as seen from Figure A.7.

Appendix B Markov Chain Monte Carlo

The Monte Carlo technique is placed among the 10 scientific computing algorithms that have had the greatest influence on the development and practice of science and engineering in the 20-th century [131]. It firstly became popular in physics [42], and it has spanned over numerous applications [130].

The basic idea of the Monte Carlo technique is to provide an accurate estimation through stochastic simulations—that is, by generating a fairly large number of samples (say, N samples are denoted as $x^{(1)}, x^{(2)}, \dots, x^{(N)}$) according to a target probability density $p(x)$ defined on a space χ (e.g., the set of possible configurations), and then use these samples to approximate the target probability density. The key challenge of the Monte Carlo method is how to generate samples. For that, the Markov chain Monte Carlo method generate these samples conditionally by further incorporating the mechanism of Markov chain; that is, the generation of the current sample only depends on the previous one (for the first-order Markov chain case). Numerous MCMC methods have been developed, among of which two most popular ones are the *Metropolis-Hastings* (M-H) algorithm [42] and the *Gibbs sampler* [43]. A brief overview of these two algorithms is provided as follows. More comprehensive reviews of the fundamental MCMC material can be found in [130, 132, 137, 264].

B.1 The Metropolis-Hastings Algorithm

To generate N samples (denoted as $x^{(1)}, x^{(2)}, \dots, x^{(N)}$) according to the distribution $p(x)$, the M-H algorithm [42] starts from a randomly-generated sample $x^{(1)}$, then proposes a candidate sample at each sample generation step, and then accepts or rejects this candidate based on an acceptance probability ratio, which is computed according to the target distribution and the previous sample. The sample generation process can be summarized as follows.

- Randomly initialize the first sample $x^{(1)}$.
- For $t = 1 : N - 1$,
 - Generate a candidate $x^{(t+1)*}$, which is sampled according to a distribution $q(x^{(t+1)*}|x^{(t)})$; this distribution is usually selected as a Gaussian distribution with the mean value as sample at the previous step, and the user-defined variance value σ^2 .
 - Whether the candidate $x^{(t+1)*}$ is accepted as a value for the new sample at the $(t + 1)$ -th step (i.e., $x^{(t+1)} = x^{(t+1)*}$) or not is determined by an acceptance probability ratio A , which is given by $A = \min \left\{ 1, \frac{p(x^{(t+1)*})}{p(x^{(t)})} \right\}$. If $x^{(t+1)*}$ is rejected, then $x^{(t)}$ will be remained to be used; that is $x^{(t+1)} = x^{(t)}$.

An example is provided to contribute some insightful comments for understanding the above sample-generation process of the M-H method, by using the M-H method to generate samples according to a distribution

$$p(x) = 0.3 \times \frac{1}{2\sqrt{2\pi}} e^{-\frac{x^2}{8}} + 0.7 \times \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-10)^2}{8}}. \quad (\text{B.1})$$

The distribution used for generate the candidate sample at each sample generation

Appendix B Markov chain Monte Carlo

Figure B.1. The comparison between the distribution defined in (B.1) and the histogram of the samples generated by the M-H method according to this distribution. The histogram of the samples become more consistent to their distribution when a larger number of samples are generated: (a) the distribution defined in (B.1); (b) the histogram of 50 samples generated by the M-H method; (c) the histogram of 500 samples generated by the M-H method; and (d) the histogram of 5000 samples generated by the M-H method.

step of the M-H method is experimentally selected as a Gaussian distribution with the mean as the sample at the previous step (i.e., $x^{(t)}$) and the variance 100; that is, $q(x^{(t+1)}|x^{(t)}) \sim N(x^{(t)}, 100)$. Figure B.1 demonstrates the results of running the M-H algorithm for generating 50, 500 and 5000 samples, respectively. As expected, the distribution of these samples meets the target distribution, especially when a large number of samples are used.

B.2 The Gibbs Sampler

Although the M-H method is efficient enough to handle the one-dimensional sample generation (i.e., the sample x has dimension of 1 in the example presented in the previous section), the Gibbs sampler [43] is suitable to handle the high-dimensional sample generation (particularly the image data). The key idea of the Gibbs sampler is to generate new sample by updating its only one component according to its conditional distribution given the rest components at each sample generation step, rather than update its all components according to its joint distribution at one time. This is due to that such conditional distributions are far easier to simulate than the complex joint distributions. Let x_n denote the n -th component of the $M \times 1$ vector $\mathbf{x}$. The sample generation process can be summarized as follows.

- Randomly initialize the first sample $\mathbf{x}^{(1)}$ by randomly generating all of its components $x_1^{(1)}, x_2^{(1)}, \dots, x_M^{(1)}$.
- For $t = 1 : N - 1$,
 - Randomly select the i -th component of $\mathbf{x}^{(t+1)}$, and update its value by sampling according to its conditional distribution given the rest components $p(x_i^{(t+1)} | x_1^{(t)}, \dots, x_{i-1}^{(t)}, x_{i+1}^{(t)}, \dots, x_M^{(t)})$. that is, $\mathbf{x}^{(t+1)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_{n-1}^{(t)}, x_n^{(t+1)}, x_{n+1}^{(t)}, \dots, x_M^{(t)}\}$.

To provide a better understanding of the concept of the Gibbs sampler method, an example is presented as follows. Let $x \sim N(x; \mu, \sigma^2)$ denote the Gaussian probability density function with a mean μ and a variance σ^2 . Suppose the goal is to generate samples according to the following 2-D Gaussian distribution

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim N \left(\begin{pmatrix} x \\ y \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix} \right). \quad (\text{B.2})$$

Appendix B Markov chain Monte Carlo

Figure B.2. The samples generated by the Gibbs sampler according to the distribution defined in (B.2): (a) 10 samples starting from the position $(10, 10)$; (b) 200 samples starting from the position $(10, 10)$; (c) 1000 samples starting from the position $(10, 10)$; and (d) 1000 samples starting from the position $(-10, -10)$.

Experiments are conducted using the Gibbs sampler to generate samples (illustrated in Figure B.2) according to the distribution (B.2). As seen from Figure B.2, one can draw the following four key conclusions: (i) the samples are generated in an order of a Markov chain; that is, each sample is dependent on the previous sample (see Figure B.2 (a)); (ii) the mean value computed via more number of samples (see 1000 samples in Figure B.2 (c)) is more reliable than that computed via insufficient number of samples (see 200 samples in Figure B.2 (b)); (iii) determine when the samples become reliable for computing the mean value of the samples is a key issue, since the firstly-generated samples are dependent on the randomly-generated initial sample; thus they are unreliable and should be discarded; and (iv) the mean value computed via reliable samples are independent with the starting point of the initial sample (compare Figure B.2 (c) starting from the position $(10, 10)$ and Figure B.2 (d) starting from the position $(-10, -10)$).

Appendix C Markov Random Field

The *Markov random field* (MRF) model aims to model the image data that are spatially interacted and thus form a joint probability distribution. It has been extensively used in various image processing and computer vision tasks [145, 146], including image restoration, image segmentation, missing data treatment.

The MRF model forms a probabilistic model for a set of variables (image's intensities) that interact on a lattice structure. For an $M \times M$ image, it can be represented by a two dimensional lattice that is denoted by $S = (p, q)$, with p, q being integers from the set $1, 2, \dots, M$. Let $\mathbf{x} = \{x(s), s \in S\}$ denote the image's intensity values, where $x(s) \in \Lambda$ (Λ is the set of the image's intensity values) and $x \in \chi$ (χ is the set of all possible configurations of the image intensity values). Then, an MRF model is defined as a joint probability density on the set of all possible configuration $\mathbf{x}$ satisfying the following conditions [145]

- *Positivity*: $p(x(s)) > 0$ for all s .
- *Markovianity*: $p(x(s) | \text{image's intensity values at the positions in the lattice except } s) = p(x(s) | \text{image's intensity values of the neighbors of } s)$.
- The distribution of the image's intensity value at a particular pixel position is conditioned on that of the neighborhood surrounding that pixel position.

Appendix C Markov random field

The above MRF model can be mathematically expressed as [145]

$$p(\mathbf{x}) = \frac{1}{Z(\lambda)} \exp \{-\lambda U(\mathbf{x})\}, \quad (\text{C.1})$$

where λ is the so-called *hyperparameter*, and $Z(\lambda)$ is a normalizing factor, which is called the *partition function* and defined as [145]

$$Z(\lambda) = \int_{\chi} \exp \{-\lambda U(\mathbf{x})\} d\mathbf{x}, \quad (\text{C.2})$$

where χ represents all possible configurations of the unknown high-resolution image $\mathbf{x}$, and the *energy function* $U(\mathbf{x})$ is made of [145]

$$U(\mathbf{x}) = \sum_{c \in C} V_c(\mathbf{x}), \quad (\text{C.3})$$

in which $V_c(\mathbf{x})$ is a *potential function* of a local group of pixels c (called the *clique*) and its value depends on the local configuration on the clique c . Hence, $U(\mathbf{x})$ is obtained by computing the potential function over all possible cliques (denoted as the set C).

Author's Publications

Journals

- [J1] Jing Tian and Kai-Kuang Ma, "A state-space super-resolution approach for video reconstruction," *Signal, Image and Video and Processing*, accepted.
- [J2] Jing Tian and Kai-Kuang Ma, "Stochastic image resolution enhancement," *Journal of Visual Communication and Image Representation*, under review.
- [J3] Jing Tian and Kai-Kuang Ma, "Super-resolution imaging — A survey and new results," in preparation.

Conferences

- [C1] Jing Tian and Kai-Kuang Ma, "A MCMC approach for Bayesian super-resolution image reconstruction," in *Proc. IEEE Int. Conf. on Image Processing*. Genoa, Italy, Sept. 2005, pp. 45-48.
- [C2] Jing Tian and Kai-Kuang Ma, "A new state-space approach for super-resolution image sequence reconstruction," in *Proc. IEEE Int. Conf. on Image Processing*. Genoa, Italy, Sept. 2005, pp. 881-884.
- [C3] Jing Tian and Kai-Kuang Ma, "Markov chain Monte Carlo super-resolution image reconstruction with simultaneous adaptation of the prior image model," in *Pacific-Rim Conference on Multimedia*. Nov. 2006, pp. 287-294.

-
- [C4] Jing Tian and Kai-Kuang Ma, "Markov chain Monte Carlo super-resolution image reconstruction with artifacts suppression," in *Proc. IEEE Asia-Pacific Conference on Circuits and Systems*. Singapore, Dec. 2006, pp. 940-943.
- [C5] Jing Tian and Kai-Kuang Ma, "Super-resolution imaging using grid computing," in *Proc. IEEE Int. Symp. on Cluster Computing and the Grid*. Rio de Janeiro, Brazil, May 2007, pp. 293-300.
- [C6] Jing Tian and Kai-Kuang Ma, "Grid computing for stochastic super-resolution imaging: Fundamentals and algorithms," in *Proc. Int. Workshop on Grid Computing & Applications*. Singapore, Jun. 2006, pp. 190-195.
- [C7] Jing Tian and Kai-Kuang Ma, "Edge-adaptive super-resolution image reconstruction using a Markov chain Monte Carlo approach," in *Proc. Int. Conf. on Information, Communications and Signal Processing*, Singapore, Dec. 2007.
- [C8] Kai-Kuang Ma and Jing Tian, "Super-resolution imaging: MCMC and state-space methods," in *Proc. Int. Conf. on Information, Communications and Signal Processing*, Singapore, Dec. 2007. (Invited Paper)

Bibliography

- [1] S. Chaudhuri, *Super-Resolution Imaging*. Boston: Kluwer Academic Publishers, 2001.
- [2] A. K. Katsaggelos, R. Molina, and J. Mateos, *Super Resolution of Images and Video*. Morgan & Claypool, 2007.
- [3] A. Katartzis and M. Petrou, "High resolution images from a sequence of low resolution observations," in *Image Fusion Algorithms and Applications*, T. Stathaki, Ed. Academic Press, 2008.
- [4] M. G. Kang and S. Chaudhuri, "Super-resolution image reconstruction," *IEEE Signal Processing Mag.*, vol. 20, pp. 19–20, May 2003.
- [5] N. K. Bose, R. H. Chan, and M. K. Ng, *Special Issue on High Resolution Image Reconstruction, International Journal of Imaging Systems and Technology*, vol. 14, no. 2-3, Aug. 2004.
- [6] M. K. Ng, T. Chan, M. G. Kang, and P. Milanfar, *Special Issue on Super-resolution Imaging: Analysis, Algorithms, and Applications, EURASIP Journal on Applied Signal Processing*, 2006.
- [7] M. K. Ng, E. Lam, and C. S. Tong, *Special Issue on Superresolution Imaging: Theory, Algorithms and Applications, Multidimensional Systems and Signal Processing*, vol. 18, Sep. 2007.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [8] R. C. Hardie, R. R. Schultz, and K. E. Barner, *Special Issue on Super-Resolution Enhancement of Digital Video, EURASIP Journal on Applied Signal Processing*, 2007.
- [9] A. K. Katsaggelos and R. Molina, *Special Issue on Super Resolution, The Computer Journal*, 2008.
- [10] B. R. Hunt, "Superresolution of images: Algorithms, principles, performance," *International Journal of Imaging Systems and Technology*, vol. 6, no. 4, pp. 297–304, 1995.
- [11] S. Borman and R. L. Stevenson, "Spatial resolution enhancement of low-resolution image sequences: A comprehensive review with directions for future research," Univ. of Notre Dame, Tech. Rep., 1998.
- [12] —, "Super-resolution from image sequences — a review," in *Proc. IEEE Midwest Symp. on Circuits and Systems*. Notre Dame, IN, Aug. 1998, pp. 374–378.
- [13] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Mag.*, vol. 20, pp. 21–36, May 2003.
- [14] M. K. Ng and N. K. Bose, "Mathematical analysis of super-resolution methodology," *IEEE Signal Processing Mag.*, vol. 20, pp. 62–74, May 2003.
- [15] D. Capel and A. Zisserman, "Computer vision applied to super resolution," *IEEE Signal Processing Mag.*, vol. 20, pp. 75–86, May 2003.
- [16] E. Choi, J. Choi, and M. G. Kang, "Super-resolution approach to overcome physical limitations of imaging sensors: An overview," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 36–46, Aug. 2004.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [17] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 47–57, Aug. 2004.
- [18] J. D. Ouwerkerk, "Image super-resolution survey," *Image and Vision Computing*, vol. 24, pp. 1039–1052, Oct. 2006.
- [19] G. Cristobal, E. Gila, F. Sroubek, J. Flusser, C. Miravet, and F. B. Rodriguez, "Superresolution imaging: A survey of current techniques," in *Proc. SPIE Conf. on Advanced Signal Processing Algorithms, Architectures, and Implementations*, Aug. 2008, pp. 70 740C.1–70 740C.18.
- [20] P. Vandewalle, S. Susstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolutionn," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 71459, 2006.
- [21] T. Q. Pham, L. J. van Vliet, and K. Schutte, "Robust fusion of irregularly sampled data using adaptive normalized convolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 83268, 2006.
- [22] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Trans. Image Processing*, vol. 13, pp. 1327–1344, Oct. 2004.
- [23] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. on Computer Vision*. Bombay, India, Jan. 1998, pp. 839–846.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [24] J. Tian and K.-K. Ma, "A MCMC approach for Bayesian super-resolution image reconstruction," in *Proc. IEEE Int. Conf. on Image Processing*. Genoa, Italy, Sept. 2005, pp. 45–48.
- [25] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, pp. 600–612, Apr. 2004.
- [26] Y. Qu, Z. Pu, H. Zhao, and Y. Zhao, "Comparison of different quality assessment functions in autoregulative illumination intensity algorithms," *Optical Engineering*, vol. 45, pp. 117 201.1–117 201.6, Nov. 2006.
- [27] H. R. Sheikh, A. C. Bovik, and L. K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Trans. Image Processing*, vol. 14, pp. 1918–1927, Dec. 2005.
- [28] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy and undersampled measured images," *IEEE Trans. Image Processing*, vol. 6, pp. 1646–1658, Dec. 1997.
- [29] D. Rajan and S. Chaudhuri, "Generalized interpolation and its application in super-resolution imaging," *Image and Vision Computing*, vol. 19, pp. 957–969, Nov. 2001.
- [30] —, "An MRF-based approach to generation of super-resolution images from blurred observations," *Journal of Mathematical Imaging and Vision*, vol. 16, pp. 5–15, Jan. 2002.
- [31] M. K. Ng and A. C. Yau, "Super-resolution image restoration from blurred low-resolution images," *Journal of Mathematical Imaging and Vision*, vol. 23, pp. 367–378, Nov. 2005.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [32] M. V. Joshi, S. Chaudhuri, and R. Panuganti, "Super-resolution imaging: Use of zoom as a cue," *Image and Vision Computing*, vol. 22, pp. 1185–1196, Dec. 2004.
- [33] —, "A learning-based method for image super-resolution from zoomed observations," *IEEE Trans. on Systems, Man and Cybernetics, Part B*, vol. 35, pp. 527–537, Jun. 2005.
- [34] S. Chaudhuri and M. V. Joshi, *Motion-free Super-resolution*. New York: Springer, 2005.
- [35] M. K. Ng, H.-F. Shen, S. Chaudhuri, and A. C. Yau, "Zoom based super-resolution reconstruction approach using total variation prior," *Optical Engineering*, vol. 46, pp. 127 003.1–127 003.11, Dec. 2007.
- [36] N. K. Bose and K. J. Boo, "High-resolution image reconstruction with multisensors," *International Journal of Imaging Systems and Technology*, vol. 9, no. 4, pp. 294–304, 1998.
- [37] M. K. Ng and A. M. Yip, "A fast MAP algorithm for high-resolution image reconstruction with multisensors," *Multidimensional Systems and Signal Processing*, vol. 12, pp. 143–164, Apr. 2001.
- [38] M. K. Ng, J. Koo, and N. K. Bose, "Constrained total least squares computations for high resolution image reconstruction with multisensors," *International Journal of Imaging Systems and Technology*, vol. 12, pp. 35–42, Jan. 2002.
- [39] W. Ching, M. K. Ng, K. Sze, and A. C. Yau, "Super-resolution image reconstruction from blurred observations by multisensors," *International Journal of Imaging Systems and Technology*, vol. 13, pp. 153–160, Sep. 2003.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [40] Y. Wen, M. K. Ng, and W. Ching, "High-resolution image reconstruction from rotated and translated low-resolution images with multisensors," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 75–83, Aug. 2004.
- [41] N. K. Bose, "Multispectral sensing," in *Distributed Sensor Networks*, S. S. Iyengar and R. R. Brooks, Eds. CRC Press, 2004.
- [42] N. Metropolis, M. N. Rosenbluth, A. W. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, pp. 1087–1092, Jun. 1953.
- [43] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, Nov. 1984.
- [44] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, T. S. Huang, Ed. London: JAI Press Inc., 1984.
- [45] Z.-Z. Wang and F.-H. Qi, "On ambiguities in super-resolution modeling," *IEEE Signal Processing Letters*, vol. 11, pp. 678–681, Aug. 2004.
- [46] G. Rochefort, F. Champagnat, G. L. Besnerais, and J. F. Giovannelli, "An improved observation model for super-resolution under affine motion," *IEEE Trans. on Image Processing*, vol. 15, pp. 3325–3337, Nov. 2006.
- [47] S. P. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1013–1027, Jun. 1990.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [48] S. P. Kim and W. Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Trans. Image Processing*, vol. 2, pp. 534–539, Apr. 1993.
- [49] S. Rhee and M. G. Kang, "Discrete cosine transform based regularized high-resolution image reconstruction algorithm," *Optical Engineering*, vol. 38, pp. 1348–1356, Aug. 1999.
- [50] N. A. Woods, N. P. Galatsanos, and A. K. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. Image Processing*, vol. 15, pp. 201–213, Jan. 2006.
- [51] C. S. Tong and K. T. Leung, "Super-resolution reconstruction based on linear interpolation of wavelet coefficients," *Multidimensional Systems and Signal Processing*, vol. 18, pp. 153–171, Sep. 2007.
- [52] S. E. Ei-Khamy, M. M. Hadhoud, M. I. Dessouky, B. M. Salam, and F. E. Ei-Samie, "Regularized super-resolution reconstruction of images using wavelet fusion," *Optical Engineering*, vol. 44, pp. 097 001.1–097 001.10, Sep. 2005.
- [53] —, "Wavelet fusion: A tool to break the limits on LMMSE image super-resolution," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 4, pp. 105–118, Mar. 2006.
- [54] H. Ji and C. Fermuller, "Wavelet-based super-resolution reconstruction: Theory and algorithm," in *Proc. European Conf. on Computer Vision*. Graz, Austria, May 2006, pp. 295–307.
- [55] M. B. Chappalli and N. K. Bose, "Simultaneous noise filtering and super-resolution with second-generation wavelets," *IEEE Signal Processing Letters*, vol. 12, pp. 772–775, Nov. 2005.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [56] N. Nguyen and P. Milanfar, "A wavelet-based interpolation-restoration method for superresolution," *Circuits, Systems, and Signal Processing*, vol. 19, pp. 321–338, Aug. 2000.
- [57] R. Chan, T. Chan, L. Shen, and Z. Shen, "Wavelet algorithms for high-resolution image reconstruction," *SIAM Journal on Scientific Computing*, vol. 24, pp. 1408–1432, Apr. 2003.
- [58] P. M. Shankar and M. A. Neifeld, "Sparsity constrained regularization for multiframe image restoration," *Optical Society of America, Journal, A: Optics and Image Science*, vol. 25, pp. 1199–1214, May 2008.
- [59] N. K. Bose, S. Lertrattanapanich, and M. B. Chappalli, "Superresolution with second generation wavelets," *Signal Processing: Image Communication*, vol. 19, pp. 387–391, Aug. 2004.
- [60] N. K. Bose and M. B. Chappalli, "A second generation wavelet framework for superresolution with noise filtering," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2-3, pp. 84–89, Aug. 2004.
- [61] L. Shen and Q. Sun, "Biorthogonal wavelet system for high-resolution image reconstruction," *IEEE Trans. Signal Processing*, vol. 52, pp. 1997–2011, Jul. 2004.
- [62] Y. Lu, L. Shen, and Y. Xu, "Multi-parameter regularization methods for high-resolution image reconstruction with displacement errors," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 54, pp. 1788–1799, Aug. 2007.
- [63] N. Goldberg, A. Feuer, and G. C. Goodwin, "Super-resolution reconstruction using spatio-temporal filtering," *Image and Vision Computing*, vol. 14, pp. 508–525, Dec. 2003.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [64] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 181–186, Mar. 1992.
- [65] S. Prasad, "Digital superresolution and the generalized sampling theorem," *Optical Society of America, Journal, A: Optics and Image Science*, vol. 24, pp. 311–325, Feb. 2007.
- [66] M.-C. Chiang and T. E. Boulton, "Efficient super-resolution via image warping," *Image and Vision Computing*, vol. 28, pp. 761–771, Jul. 2000.
- [67] C. T. Lin, K. W. Fan, H. C. Pu, S. M. Lu, and S. F. Liang, "An HVS-directed neural-network-based image resolution enhancement scheme for image resizing," *IEEE Trans. on Fuzzy Systems*, vol. 15, pp. 605–615, Aug. 2007.
- [68] N. K. Bose and N. A. Ahuja, "Superresolution and noise filtering using moving least squares," *IEEE Trans. on Image Processing*, vol. 15, pp. 2239–2248, Aug. 2006.
- [69] S. Lertrattanapanich and N. K. Bose, "High resolution image formation from low resolution frames using Delaunay triangulation," *IEEE Trans. on Image Processing*, vol. 11, pp. 1427–1441, Dec. 2002.
- [70] X. Li, "Super-resolution for synthetic zooming," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 58195, 2006.
- [71] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 231–239, May 1991.
- [72] Z. D. Jiang, T. T. Wong, and H. J. Bao, "Practical super-resolution from dynamic video sequences," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*. Madison, WI, Jun. 2003, pp. 549–554.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [73] C.-Y. Chen, Y.-C. Kuo, and C.-S. Fuh, "Image reconstruction with improved super-resolution algorithm," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, pp. 1513–1527, Dec. 2004.
- [74] Y. Lu and M. Inamura, "Pyramid-based super-resolution of the undersampled and subpixel shifted image sequence," *International Journal of Imaging Systems and Technology*, vol. 12, pp. 254–263, Apr. 2002.
- [75] —, "Super-resolution of undersampled and subpixel shifted image sequence by pyramid iterative backprojection," *IEICE Trans. on Information and Systems*, vol. 85, pp. 1929–1937, Dec. 2002.
- [76] —, "Super-resolution image pyramid," *IEICE Trans. on Information and Systems*, vol. 86, pp. 1436–1446, Aug. 2003.
- [77] B.-W. Jeon, R.-H. Park, and S. Yang, "Resolution enhancement by prediction of the high-frequency image based on the Laplacian pyramid," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 72520, 2006.
- [78] A. J. Patti and A. M. Tekalp, "Super resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Trans. Image Processing*, vol. 6, pp. 1446–1451, Oct. 1997.
- [79] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Philadelphia, PA: IOP Publishing Ltd, 1998.
- [80] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images," in *Proc. IEEE Int. Conf. on Image Processing*. Washington, DC, Oct. 1995, pp. 539–542.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [81] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Processing*, vol. 5, pp. 996–1011, Jun. 1996.
- [82] D. Kong, M. Han, W. Xu, H. Tao, and Y. Gong, "A conditional random field model for video super-resolution," in *Proc. IEEE Int. Conf. on Pattern Recognition*. Hong Kong, Aug. 2006, pp. 619–622.
- [83] K. V. Suresh and A. N. Rajagopalan, "Robust and computationally efficient superresolution algorithm," *Optical Society of America, Journal, A: Optics and Image Science*, vol. 24, pp. 984–992, Apr. 2007.
- [84] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach for joint motion estimation, segmentation, and super resolution," *IEEE Trans. on Image Processing*, vol. 16, pp. 479–490, Feb. 2007.
- [85] F. Humblot and A. Mohammad-Djafari, "Super-resolution using hidden Markov model and Bayesian detection estimation framework," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 36971, 2006.
- [86] A. Mohammad-Djafari, "Super-resolution: A short review, a new method based on hidden Markov modeling of hr image and future challenges," *The Computer Journal*, 2008.
- [87] D. Capel and A. Zisserman, "Super-resolution enhancement of text image sequences," in *Proc. IEEE Int. Conf. on Pattern Recognition*. Barcelona, Spain, Sep. 2000, pp. 600–605.
- [88] A. Chakrabarti, A. N. Rajagopalan, and R. Chellappa, "Super-resolution of face images using kernel PCA-based prior," *IEEE Trans. on Multimedia*, vol. 9, pp. 888–892, Jun. 2007.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [89] K. I. Kim, M. Franz, and B. Scholkopf, "Kernel hebbian algorithm for single-frame super-resolution," in *Proc. ECCV Workshop on Statistical Learning in Computer Vision*. Prague, Czech Republic, May 2004, pp. 135–149.
- [90] U. Mudenagudi, R. Singla, P. Kalra, and S. Banerjee, "Super resolution using graph-cut," in *Proc. Asian Conference on Computer Vision*. Hyderabad, India, Jan. 2006, pp. 385–394.
- [91] T. Chang, T. Liu, and J. Chuang, "Direct energy minimization for super-resolution on nonlinear manifolds," in *Proc. European Conf. on Computer Vision*. Graz, Austria, May 2006, pp. 281–294.
- [92] A. J. Storkey, "Dynamic structure super-resolution," in *Advances in Neural Information Processing Systems*, S. Thrun, L. Saul, and B. Schölkopf, Eds. Cambridge, MA: MIT Press, 2003.
- [93] Y.-W. Tai, W.-S. Tong, and C.-K. Tang, "Perceptually-inspired and edge-directed color image super-resolution," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*. New York, NY, Jun. 2006, pp. 1948–1955.
- [94] M. Elad and D. Datsenko, "Example-based regularization deployed to super-resolution reconstruction of a single image," *The Computer Journal*, 2008.
- [95] C. V. Jiji, S. Chaudhuri, and P. Chatterjee, "Single frame image super-resolution: Should we process locally or globally?" *Multidimensional Systems and Signal Processing*, vol. 18, pp. 123–152, Sep. 2007.
- [96] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image analogies," in *Proc. SIGGRAPH*. Los Angeles, Aug. 2001, pp. 327–340.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [97] F. M. Candocia and J. C. Principe, "Super-resolution of images based on local correlations," *IEEE Trans. on Neural Networks*, vol. 10, pp. 372–380, Mar. 1999.
- [98] —, "Superresolution of images with learned multiple reconstruction kernels," in *Multimedia Image and Video Processing*, L. Guan, S. Kung, and J. Larsen, Eds. Boca Raton: CRC Press, 2001.
- [99] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International Journal of Computer Vision*, vol. 40, pp. 25–47, Oct. 2000.
- [100] I. Begin and F. R. Ferrie, "Blind super-resolution using a learning-based approach," in *Proc. Int. Conf. on Pattern Recognition*. Cambridge, UK, Aug. 2004, pp. 85–89.
- [101] J. Qiao, J. Liu, and C. Zhao, "A novel SVM-based blind super-resolution algorithm," in *International Joint Conference on Neural Networks*. Vancouver, Canada, Jul. 2006, pp. 2523–2528.
- [102] L. C. Pickup, S. J. Roberts, and A. Zisserman, "A sampled texture prior for image super-resolution," in *Advances in Neural Information Processing Systems*, S. Thrun, L. Saul, and B. Scholkopf, Eds. Cambridge, MA: MIT Press, 2003.
- [103] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*. Washington, D.C., Jul. 2004, pp. 275–282.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [104] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, "Single frame image super-resolution using learnt wavelet coefficients," *International Journal on Imaging Science and Technology*, vol. 14, pp. 105–112, Sep. 2004.
- [105] C. V. Jiji and S. Chaudhuri, "Single-frame image super-resolution through contourlet learning," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 73767, 2006.
- [106] X. Li, "Image resolution enhancement via data-driven parametric models in the wavelet space," *EURASIP Journal on Image and Video Processing*, vol. 2007, Article ID 41516, 2007.
- [107] D. Datsenko and M. Elad, "Example-based single document image super-resolution: A global MAP approach with outlier rejection," *Multidimensional Systems and Signal Processing*, vol. 18, pp. 103–121, Sep. 2007.
- [108] F. Pan and L. Zhang, "New image super-resolution scheme based on residual error restoration by neural networks," *Optical Engineering*, vol. 42, pp. 3038–3046, Oct. 2003.
- [109] G. Qiu, "Interresolution look-up table for improved spatial magnification of image," *Journal of Visual Communication and Image Representation*, vol. 11, pp. 360–373, Dec. 2000.
- [110] C. Miraveta and F. B. Rodriguez, "A two-step neural-network based algorithm for fast image super-resolution," *Image and Vision Computing*, vol. 25, pp. 1449–1473, Sep. 2007.
- [111] K. S. Ni and T. Q. Nguyen, "Image superresolution using support vector regression," *IEEE Trans. on Image Processing*, vol. 16, pp. 1596–1610, Jun. 2007.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [112] D. Capel and A. Zisserman, "Super-resolution from multiple views using learnt image models," in *Proc. Int. Conf. on Computer Vision and Pattern Recognition*. Hawaii, HW, Dec. 2001, pp. 627–634.
- [113] Q. Wang, X. Tang, and H. Shum, "Patch based blind image super resolution," in *Proc. IEEE Int. Conf. on Computer Vision*. Beijing, China, Oct. 2005, pp. 709–716.
- [114] C. Liu, H.-Y. Shum, and W. T. Freeman, "Face hallucination: Theory and practice," *International Journal of Computer Vision*, vol. 75, pp. 115–134, Oct. 2007.
- [115] S. Rajaram, M. D. Gupta, N. Petrovic, and T. S. Huang, "Learning-based nonparametric image super-resolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 51306, 2006.
- [116] K. V. Suresh, G. M. Kumar, and A. N. Rajagopalan, "Superresolution of license plates in real traffic videos," *IEEE Trans. on Intelligent Transportation Systems*, vol. 8, pp. 321–331, Jun. 2007.
- [117] B. Narayanan, R. C. Hardie, K. E. Barner, and M. Shao, "A computationally efficient super-resolution algorithm for video processing using partition filters," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, pp. 621–634, May 2007.
- [118] M. K. Ng, H.-F. Shen, L.-P. Zhang, and E. Lam, "A total variation regularization based super-resolution reconstruction algorithm for digital video," *EURASIP Journal on Applied Signal Processing*, vol. 2007, Article ID 74585, 2007.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [119] V. Patanavijit and S. Jitapunkul, "A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian-Tikhonov regularization," *EURASIP Journal on Applied Signal Processing*, vol. 2007, Article ID 34821, 2007.
- [120] B. C. Tom and A. K. Katsaggelos, "Resolution enhancement of monochrome and color video using motion compensation," *IEEE Trans. Image Processing*, vol. 10, pp. 278–287, Feb. 2001.
- [121] N. R. Shah and A. Zakhor, "Resolution enhancement of color video sequences," *IEEE Trans. Image Processing*, vol. 8, pp. 879–885, Jun. 1999.
- [122] S. Borman and R. L. Stevenson, "Simultaneous multi-frame MAP super-resolution video enhancement using spatio-temporal priors," in *Proc. IEEE Int. Conf. on Image Processing*. Kobe, Japan, Oct. 1999, pp. 469–473.
- [123] M. Zibetti and J. Mayer, "Simultaneous super-resolution for video sequences," in *Proc. IEEE Int. Conf. on Image Processing*. Genoa, Italy, Sep. 2005, pp. 877–880.
- [124] —, "Outlier robust and edge-preserving simultaneous super-resolution," in *Proc. IEEE Int. Conf. on Image Processing*. Atlanta, GA, Oct. 2006, pp. 1741–1744.
- [125] —, "A robust and computationally efficient simultaneous super-resolution scheme for image sequences," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, pp. 1288–1300, Oct. 2007.
- [126] L. D. Alvarez, R. Molina, and A. K. Katsaggelos, "Multi-channel reconstruction of video sequences from low-resolution and compressed observations," in

BIBLIOGRAPHY

BIBLIOGRAPHY

- the 8th Iberoamerican Congress on Pattern Recognition.* Havana, Cuba, Nov. 2003, pp. 46–53.
- [127] C. M. Bishop, A. Blake, and B. Marthi, “Super-resolution enhancement of video,” in *Proc. Int. Conf. on Artificial Intelligence and Statistics*. Key West, FL, Jan. 2003.
- [128] 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 Trans. Image Processing*, vol. 6, pp. 1621–1633, Dec. 1997.
- [129] H. He and L. P. Kondi, “An image super-resolution algorithm for different error levels per frame,” *IEEE Trans. Image Processing*, vol. 15, pp. 592–623, Mar. 2006.
- [130] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall, 1996.
- [131] J. Dongarra and F. Sullivan, “Introduction to the top 10 algorithms,” *Computing in Science & Engineering*, vol. 2, pp. 22–23, Jan.-Feb. 2000.
- [132] W. J. Fitzgerald, “Markov chain Monte Carlo methods with applications to signal processing,” *Signal Processing*, vol. 81, pp. 3–18, Jan. 2001.
- [133] A. Jalobeanu, L. Blanc-Feraud, and J. Zerubia, “Hyperparameter estimation for satellite image restoration using a MCMC maximum-likelihood method,” *Pattern Recognition*, vol. 35, pp. 341–352, Feb. 2002.
- [134] A. C. Kokaram and S. J. Godsill, “MCMC for joint noise reduction and missing data treatment in degraded video,” *IEEE Trans. Signal Processing*, vol. 50, pp. 189–205, Feb. 2002.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [135] A. C. Kokaram, "On missing data treatment for degraded video and film archives: A survey and a new Bayesian approach for joint noise reduction and missing data treatment in degraded video," *IEEE Trans. Image Processing*, vol. 13, pp. 397–415, Mar. 2004.
- [136] —, "Practical, unified, motion and missing data treatment in degraded video," *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 163–177, January–March 2004.
- [137] A. Doucet and X. Wang, "Monte Carlo methods for signal processing," *IEEE Signal Processing Mag.*, vol. 22, pp. 152–170, Nov. 2005.
- [138] G. Papandreou, P. Maragos, and A. Kokaram, "Image inpainting with a wavelet domain hidden markov tree model," in *Proc. IEEE Int. Conf. on Acoustic, Speech and Signal Processing*. Las Vegas, NV, Mar.-Apr. 2008, pp. 773–776.
- [139] M. E. Tipping and C. M. Bishop, "Bayesian image super-resolution," in *Advances in Neural Information Processing Systems*, S. Becker, S. Thrun, and K. Obermeyer, Eds. Cambridge: MIT Press, 2002.
- [140] L. C. Pickup, D. Capel, S. J. Roberts, and A. Zisserman, "Bayesian image super-resolution, continued," in *Advances in Neural Information Processing Systems*, B. Scholkopf, J. Platt, and T. Hoffman, Eds., Cambridge, MA: MIT Press, 2006, pp. 1089–1096.
- [141] H. He and L. P. Kondi, "Resolution enhancement of video sequences with simultaneous estimation of the regularization parameter," *Journal of Electronic Imaging*, vol. 13, pp. 586–596, Jul. 2004.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [142] M. Zibetti, F. Bazan, and J. Mayer, "Determining the regularization parameters for super-resolution problems," *Signal Processing*, vol. 88, pp. 2890–2901, Dec. 2008.
- [143] N. K. Bose, S. Lertrattanapanich, and J. Koo, "Advances in superresolution using L-curve," in *Proc. IEEE Int. Symp. on Circuits and Systems*. Sydney, Australia, May 2001, pp. 433–436.
- [144] C. F. J. Wu, "On the convergence properties of the EM algorithm," *The Annals of Statistics*, vol. 11, pp. 95–103, Mar. 1983.
- [145] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. New York: Springer-Verlag, 1995.
- [146] H. Rue, *Gaussian Markov Random Fields: Theory and Applications*. Boca Raton: Chapman & Hall, 2005.
- [147] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, and A. K. Katsaggelos, "Hierarchical Bayesian image restoration from partially known blurs," *IEEE Trans. Image Processing*, vol. 9, pp. 1784–1797, Oct. 2000.
- [148] M. Figueiredo and R. Nowak, "Wavelet-based image estimation: An empirical Bayes approach using Jeffreys' noninformative prior," *IEEE Trans. Image Processing*, vol. 10, pp. 1322–1331, Sep. 2001.
- [149] J. Tian and K.-K. Ma, "Markov chain Monte Carlo super-resolution image reconstruction with simultaneous adaptation of the prior image model," in *Advances in Multimedia Information Processing - PCM 2006, Pacific Rim Conference on Multimedia, Hangzhou, China, Nov. 2-4, 2006*, ser. Lecture Notes in Computer Science, Y.-T. Zhuang, S.-Q. Yang, Y. Rui, and Q.-M. He, Eds. Springer, vol. 4261, pp. 287–294.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [150] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction*. New York: Springer, 2003.
- [151] A. L. Gibbs, "Convergence in the Wasserstein metric for Markov chain Monte Carlo algorithms with applications to image restoration." *Stochastic Models*, vol. 20, pp. 473–492, Oct. 2004.
- [152] J. Tian and K.-K. Ma, "Markov chain Monte Carlo super-resolution image reconstruction with artifacts suppression," in *Proc. IEEE Asia-Pacific Conference on Circuits and Systems*. Singapore, Dec. 2006, pp. 940–943.
- [153] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Processing*, vol. 6, pp. 298–311, Feb. 1997.
- [154] W. Vanzella, F. A. Pellegrino, and V. Torre, "Self-adaptive regularization," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26, pp. 804–809, Jun. 2004.
- [155] M. Mignotte, "A segmentation-based regularization term for image deconvolution," *IEEE Trans. Image Processing*, vol. 15, pp. 1973–1984, Jul. 2006.
- [156] B. Choi, S. D. Kim, and J. B. Ra, "Region-based super-resolution using adaptive diffusion regularization," *Optical Engineering*, vol. 47, pp. 027 006.1–027 006.11, Feb. 2008.
- [157] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Processing*, vol. 9, pp. 505–510, Mar. 2000.
- [158] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 679–698, Nov. 1986.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [159] C. A. Segall, R. Molina, and A. K. Katsaggelos, "High-resolution images from low-resolution compressed video," *IEEE Signal Processing Mag.*, vol. 20, pp. 37–48, May 2003.
- [160] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35–45, May 1960.
- [161] M. Elad and A. Feuer, "Superresolution restoration of an image sequence: Adaptive filtering approach," *IEEE Trans. Image Processing*, vol. 8, pp. 387–395, Mar. 1999.
- [162] ———, "Super-resolution reconstruction of image sequences," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 817–834, Sep. 1999.
- [163] S. Farsiu, M. Elad, and P. Milanfar, "Video-to-video dynamic superresolution for grayscale and color sequences," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 61859, 2006.
- [164] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, N.J.: Prentice Hall, 1996.
- [165] C. Genest and J. V. Zidek, "Combining probability distributions: A critique and an annotated bibliography," *Statistical Science*, vol. 1, pp. 114–148, Jan. 1986.
- [166] J. Ruanaidh and W. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing, Statistics and Computing*. New York: Springer Verlag, 1996.
- [167] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, 1996.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [168] R. R. Schultz, M. Li, and R. L. Stevenson, "Subpixel motion estimation for super-resolution image sequence enhancement," *Journal of Visual Communication and Image Representation*, vol. 9, pp. 38–50, Mar. 1998.
- [169] M. Bertero, T. Poggio, and V. Torre, "Ill-posed problems in early vision," *Proceedings of the IEEE*, vol. 76, pp. 869–889, Aug. 1988.
- [170] A. R. Reibman, R. M. Bell, and S. Gray, "Quality assessment for super-resolution image enhancement," in *Proc. IEEE Int. Conf. on Image Processing*. Atlanta, GA, Oct. 2006, pp. 2017–2020.
- [171] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, N.J.: PTR Prentice-Hall, 1993.
- [172] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau, "Eigenface-domain super-resolution for face recognition," *IEEE Trans. on Image Processing*, vol. 12, pp. 597–606, May 2003.
- [173] G. Pan, S. Han, Z. Wu, and Y. Wang, "Super-resolution of 3D face," in *Proc. European Conf. on Computer Vision*. Graz, Austria, May 2006, pp. 389–401.
- [174] K. Jia and S. Gong, "Generalized face super-resolution," *IEEE Trans. on Image Processing*, vol. 17, pp. 873–886, Jun. 2008.
- [175] T. Akgun, Y. Altunbasak, and R. M. Mersereau, "Super-resolution reconstruction of hyperspectral images," *IEEE Trans. on Image Processing*, vol. 14, pp. 1860–1875, Nov. 2005.
- [176] M. T. Merino and J. Nunez, "Super-resolution of remotely sensed images with variable-pixel linear reconstruction," in *IEEE Trans. on Geoscience and Remote Sensing*, vol. 45, May 2007, pp. 1446–1457.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [177] H.-F. Shen, M. K. Ng, P.-X. Li, and L.-P. Zhang, "Super-resolution reconstruction algorithm to MODIS remote sensing images," *The Computer Journal*, 2007.
- [178] H. Nagahara, Y. Yagi, and M. Yachida, "Superresolution modeling using an omnidirectional image sensor," *IEEE Trans. on Systems, Man and Cybernetics, Part B*, vol. 33, pp. 607–615, Aug. 2003.
- [179] S. Chaudhuri and D. R. Taur, "High-resolution slow-motion sequencing: How to generate a slow-motion sequence from a bit stream," *IEEE Signal Processing Magazine*, vol. 22, pp. 16–24, Mar. 2005.
- [180] X. Pan, L. Yu, and C.-M. Kao, "Spatial-resolution enhancement in computed tomography," *IEEE Trans. on Medical Imaging*, vol. 24, pp. 246–253, Feb. 2005.
- [181] J. A. Kennedy, O. Israel, A. Frenkel, R. Bar-Shalom, and H. Azhari, "Super-resolution in PET imaging," *IEEE Trans. on Medical Imaging*, vol. 25, pp. 137–147, Feb. 2006.
- [182] —, "Improved image fusion in PET/CT using hybrid image reconstruction and super-resolution," *EURASIP Journal on Applied Signal Processing*, vol. 2007, Article ID 46846, 2007.
- [183] H. Greenspan, "Super-resolution in medical imaging," *The Computer Journal*, 2008.
- [184] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid Monte Carlo," *Physics Letters B*, vol. 195, pp. 216–222, Sep. 1987.
- [185] R. M. Neal, *Bayesian learning for neural networks*, ser. Lecture Notes in Statistics. New York: Springer-Verlag, 1996.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [186] G. Caner, A. M. Tekalp, and W. Heinzelman, "Performance evaluation of super-resolution reconstruction from video," in *Proc. IS&T/SPIE's Annual Symposium on Electronic Imaging*. Santa Clara, CA, Jan. 2003, pp. 177–185.
- [187] M. Tanaka and M. Okutomi, "Theoretical analysis on reconstruction-based super-resolution for an arbitrary PSF," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*. San Diego, CA, Jun. 2005, pp. 947–954.
- [188] A. N. Rajagopalan and V. P. Kiran, "Motion-free superresolution and the role of relative blur," *Optical Society of America, Journal, A: Optics and Image Science*, vol. 20, pp. 2022–2032, Nov. 2003.
- [189] T. Q. Pham, L. J. van Vliet, and K. Schutte, "Influence of signal-to-noise ratio and point spread function on limits of superresolution," in *Proc. SPIE Conf. on Image Processing: Algorithms and Systems IV*. San Jose, CA, Jan. 2005, pp. 169–180.
- [190] Z.-Z. Wang and F.-H. Qi, "Analysis of multiframe super-resolution reconstruction for image anti-aliasing and deblurring," *Image and Vision Computing*, vol. 23, pp. 393–404, Apr. 2005.
- [191] A. W. M. Eekeren, K. Schutte, O. R. Oudegeest, and L. J. Vliet, "Performance evaluation of super-resolution reconstruction methods on real-world data," *EURASIP Journal on Applied Signal Processing*, vol. 2007, Article ID 43953, 2007.
- [192] A. R. Reibman and T. Schaper, "Subjective performance evaluation for super-resolution image enhancement," in *Proc. Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics*. Scottsdale, AZ, Jan. 2006.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [193] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, pp. 1167–1183, Sep. 2002.
- [194] Z. Lin and H.-Y. Shum, "Fundamental limits of reconstruction-based super-resolution algorithms under local translation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26, pp. 83–97, Jan. 2004.
- [195] Z.-C. Lin, J.-F. He, X.-O. Tang, and C.-K. Tang, "Limits of learning-based superresolution algorithms," in *Proc. IEEE Int. Conf. on Computer Vision*. Rio de Janeiro, Brazil, Oct. 2007.
- [196] D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," *IEEE Trans. on Image Processing*, vol. 15, pp. 1413–1428, Jun. 2006.
- [197] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Trans. on Image Processing*, vol. 10, pp. 1187–1193, Aug. 2001.
- [198] D. Zhang, H. Li, and M.-H. Du, "Fast MAP-based multiframe super-resolution image reconstruction," *Image and Vision Computing*, vol. 23, pp. 671–679, Jul. 2005.
- [199] D. Barreto, L. D. Alvarez, and J. Abad, "Motion estimation techniques in super-resolution image reconstruction: A performance evaluation," in *Virtual Observatory. Plate Content Digitalization, Archive Mining and Image Sequence Processing*, M. Tsvetkov, V. Golev, F. Murtagh, and R. Molina, Eds. Sofia, Bulgaria, Apr. 2005, pp. 254–268.
- [200] S. Baker and T. Kanade, "Super resolution optical flow," The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep., Oct. 1999.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [201] W. Zhao and H. Sawhney, "Is super-resolution with optical flow feasible?" in *Proc. European Conf. on Computer Vision*. Copenhagen, Denmark, Jun. 2002, pp. 599–613.
- [202] F. Lin, C. Fookes, V. Chandran, and S. Sridharan, "Investigation into optical flow super-resolution for surveillance applications," *Proc. APRS Workshop on Digital Image Computing*, pp. 73–78, 2005.
- [203] H. V. Le and G. Seetharaman, "A super-resolution imaging method based on dense subpixel-accurate motion fields," *Journal of VLSI Signal Processing*, vol. 42, pp. 79–89, Jan. 2006.
- [204] R. Fransens, C. Strecha, and L. V. Gool, "Optical flow based super-resolution: A probabilistic approach," *Computer Vision and Image Understanding*, vol. 106, pp. 106–115, Apr. 2007.
- [205] A. M. Tekalp, *Digital Video Processing*. Prentice Hall, 1995.
- [206] D. Rajan, S. Chaudhuri, and M. V. Joshi, "Multi-objective super resolution: Concepts and examples," *IEEE Signal Processing Mag.*, vol. 20, pp. 49–61, May 2003.
- [207] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli, "Super-resolution from unregistered and totally aliased signals using subspace methods," *IEEE Trans. on Signal Processing*, vol. 55, pp. 3687–3703, Jul. 2007.
- [208] J. Chung, E. Haber, and J. Nagy, "Numerical methods for coupled super-resolution," *Inverse Problems*, vol. 22, pp. 1261–1272, Aug. 2006.
- [209] D. Robinson, S. Farsiu, and P. Milanfar, "Optimal registration of aliased images using variable projection with applications to super-resolution," *The Computer Journal*, 2008.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [210] Y. He, K.-H. Yap, L. Chen, and L.-P. Chau, "Joint image registration and super-resolution using nonlinear least squares method," in *Proc. IEEE Int. Conf. on Aconst., Speech, Signal Processing*. Honolulu, HI, 2007, pp. 561–564.
- [211] —, "A nonlinear least square technique for simultaneous image registration and super-resolution," *IEEE Trans. on Image Processing*, vol. 16, pp. 2830–2841, Nov. 2007.
- [212] E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Trans. Image Processing*, vol. 12, pp. 826–837, Jul. 2003.
- [213] P. E. Eren, M. I. Sezan, and A. M. Tekalp, "Robust, object based high resolution image reconstruction from low resolution video," *IEEE Trans. Image Processing*, vol. 6, pp. 1446–1451, Oct. 1997.
- [214] A. Zomet, A. Rav-Acha, and S. Peleg, "Robust super-resolution," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*. Hawaii, HI, Dec. 2001, pp. 645–650.
- [215] M. K. Ng and N. K. Bose, "Analysis of displacement errors in high-resolution image reconstruction with multisensors," *IEEE Trans. Circuits Syst. I*, vol. 49, pp. 806–813, Jun. 2002.
- [216] M. Trimeche, R. C. Bilcu, and J. Yrjanainen, "Adaptive outlier rejection in image super-resolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 38052, 2006.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [217] N. A. El-Yamany and P. E. Papamichalis, "Robust color image superresolution: An adaptive M-estimation framework," *EURASIP Journal on Image and Video Processing*, vol. 2008, Article ID 763254, 2008.
- [218] F. Sroubek and J. Flusser, "Resolution enhancement via probabilistic deconvolution of multiple degraded images," *Pattern Recognition Letters*, vol. 27, pp. 287–293, Mar. 2006.
- [219] F. Sroubek, G. Cristobal, and J. Flusser, "A unified approach to superresolution and multichannel blind deconvolution," *IEEE Trans. Image Processing*, vol. 16, pp. 2322–2332, Sep. 2007.
- [220] F. Sroubek, J. Flusser, and G. Cristobal, "Multiframe blind deconvolution coupled with frame registration and resolution enhancement," in *Blind Image Deconvolution: Theory and Application*, P. Campisi and K. Egiazarian, Eds. Boca Raton: CRC Press, 2007.
- [221] —, "Super-resolution and blind deconvolution for rational factors with an application to color images," *The Computer Journal*, 2008.
- [222] H. Shekarforoush and R. Chellappa, "Data-driven multichannel superresolution with application to video sequences," *Optical Society of America, Journal, A: Optics and Image Science*, vol. 16, pp. 481–492, Mar. 1999.
- [223] Y. He, K.-H. Yap, L. Chen, and L.-P. Chau, "Blind super-resolution image reconstruction using a maximum a posteriori estimation," in *Proc. IEEE Int. Conf. on Image Processing*. Atlanta, GA, Oct. 2006, pp. 1729–1732.
- [224] —, "A soft MAP framework for blind super-resolution image reconstruction," *Image and Vision Computing*, vol. 27, pp. 364–373, Mar. 2009.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [225] H. Yang, J. Gao, and Z. Wu, "Blur identification and image super-resolution reconstruction using an approach similar to variable projection," *IEEE Signal Processing Letters*, vol. 15, pp. 289–292, 2008.
- [226] K. V. Suresh and A. N. Rajagopalan, "Super-resolution in the presence of space-variant blur," in *Proc. IEEE Int. Conf. on Pattern Recognition*. Hong Kong, Aug. 2006, pp. 770–773.
- [227] B. Bascle, A. Blake, and A. Zisserman, "Motion deblurring and super-resolution from an image sequence," in *Proc. European Conf. on Computer Vision*. Cambridge, UK, Apr. 1996, pp. 571–582.
- [228] M. A. Robertson and R. L. Stevenson, "Temporal resolution enhancement in compressed video sequences," *EURASIP Journal on Applied Signal Processing*, vol. 2001, pp. 230–238, Dec. 2001.
- [229] E. Shechtman, Y. Caspi, and M. Irani, "Space-time super-resolution," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, pp. 531–545, Apr. 2005.
- [230] K. Watanabe, Y. Iwai, H. Nagahara, M. Yachnida, and T. Suzuki, "Video synthesis with high spatio-temporal resolution using motion compensation and image fusion in wavelet domain," in *Proc. Asian Conf. on Computer Vision*. Hyderabad, India, Jan. 2006, pp. 480–489.
- [231] N. Maor, A. Feuer, and G. C. Goodwin, "Compression at the source for digital camcorders," *EURASIP Journal on Image and Video Processing*, vol. 2007, Article ID 41516, 2007.
- [232] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: color filter array interpolation," *IEEE Signal Processing Magazine*, vol. 22, pp. 44–54, Jan. 2005.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [233] X. Li, B. Gunturk, and L. Zhang, "Image demosaicing: a systematic survey," in *Proc. SPIE Visual Communications and Image Processing*. San Jose, CA, Jan. 2008, pp. 68 221J.1–68 221J.15.
- [234] T. Gotoh and M. Okutomi, "Direct super-resolution and registration using raw CFA images," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*. Washington, D.C., Jul. 2004, pp. 600–607.
- [235] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *IEEE Trans. on Image Processing*, vol. 15, pp. 141–159, Jan. 2006.
- [236] M. Shao, K. E. Barner, and R. C. Hardie, "Partition-based interpolation for color filter array demosaicking and super-resolution reconstruction," *Optical Engineering*, vol. 44, pp. 107 003.1–107 003.14, Oct. 2005.
- [237] M. Vega, R. Molina, and A. K. Katsaggelos, "A Bayesian superresolution approach to demosaicing of blurred images," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 25072, 2006.
- [238] M. Gevrekci, B. K. Gunturk, and Y. Altunbasak, "Restoration of Bayer-sampled image sequences," *The Computer Journal*, 2007.
- [239] X. Wu and L. Zhang, "Improvement of color video demosaicking in temporal domain," *IEEE Trans. on Image Processing*, vol. 15, pp. 3138–3151, Oct. 2006.
- [240] —, "Temporal color video demosaicking via motion estimation and data fusion," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, pp. 231–240, Feb. 2006.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [241] L. D. Alvarez, J. Mateos, R. Molina, and A. K. Katsaggelos, "High resolution images from compressed low resolution video: Motion estimation and observable pixels," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 58–66, Aug. 2004.
- [242] C. A. Segall, R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Trans. Image Processing*, vol. 13, pp. 898–911, Jul. 2004.
- [243] L. D. Alvarez, R. Molina, and A. K. Katsaggelos, "High resolution images from a sequence of low resolution observations," in *Digital Image Sequence Processing, Compression and Analysis*, T. R. Reed, Ed. Boca Raton: CRC Press, 2004.
- [244] C. Wang, P. Xue, and W. Lin, "Improved super-resolution reconstruction from video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, pp. 1411–1422, Nov. 2006.
- [245] S. C. Park, M. G. Kang, C. A. Segall, and A. K. Katsaggelos, "Spatially adaptive high-resolution image reconstruction of low-resolution DCT-based compressed image," *IEEE Trans. Image Processing*, vol. 13, pp. 573–584, Apr. 2004.
- [246] Y. Altunbasak, A. J. Patti, and R. M. Mersereau, "Super-resolution still and video reconstruction from MPEG-coded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 217–226, Apr. 2002.
- [247] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Super-resolution reconstruction of compressed video using transform-domain statistics," *IEEE Trans. Image Processing*, vol. 13, pp. 33–43, Jan. 2004.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [248] —, “Multiframe resolution-enhancement methods for compressed video,” *IEEE Signal Processing Letters*, vol. 9, pp. 170–174, Jun. 2002.
- [249] L. Shao, “Adaptive resolution upconversion for compressed video using pixel classification,” *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 71432, 2007.
- [250] B. Martins and S. Forchhammer, “A unified approach to restoration, deinterlacing and resolution enhancement in decoding MPEG-2 video,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 803–811, Sep. 2002.
- [251] R. Molina, A. K. Katsaggelos, L. D. Alvarez, and J. Mateos, “Towards a new video compression scheme using super-resolution,” in *Proc. IS&T/SPIE Electronic Imaging, Visual Communications and Image Processing*. San Jose, CA, Jan. 2006, pp. 607 706.1–607 706.13.
- [252] D. Barreto, L. D. Alvarez, R. Molina, A. K. Katsaggelos, and G. M. Callicó, “Region-based super-resolution for compression,” *Multidimensional Systems and Signal Processing*, vol. 18, pp. 59–81, Sep. 2006.
- [253] I. Foster, C. Kesselman, and S. Tuecke, “The anatomy of the grid: Enabling scalable virtual organizations,” *International Journal of High Performance Computing Applications*, vol. 15, pp. 200–222, Aug. 2001.
- [254] M. Parashar and C. Lee, *Special Issue on Grid Computing, Proceedings of the IEEE*, vol. 93, Mar. 2003.
- [255] S. Hastings, T. Kurc, S. Langella, U. Catalyurek, T. Pan, and J. Saltz, “Image processing for the grid: A toolkit for building grid-enabled image processing applications,” in *Proc. IEEE Symp. on Cluster Computing and the Grid*. Tokyo, Japan, May 2003, pp. 36–43.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [256] J. Montagnat, F. Bellet, H. Benoit-Cattin, V. Breton, L. Brunie, H. Duque, Y. Legré, I. E. Magnin, L. Maigne, S. Miguët, J. M. Pierson, L. Seitz, and T. Tweed, "Medical images simulation, storage, and processing on the European datagrid testbed," *Journal of Grid Computing*, pp. 387–400, Dec. 2004.
- [257] Y. Kawasaki, F. Ino, Y. Mizutani, N. Fujimoto, T. Sasama, Y. Sato, N. Sugano, S. Tamura, and K. Hagihara, "High-performance computing service over the internet for intraoperative image processing," *IEEE Trans. on Information Technology in Biomedicine*, vol. 8, pp. 36–46, Jan. 2004.
- [258] B. J. Liu, M. Z. Zhou, and J. Documet, "Utilizing data grid architecture for the backup and recovery of clinical image data," *Computerized Medical Imaging and Graphics*, vol. 29, pp. 95–102, Mar. 2005.
- [259] F. Ino, K. Ooyama, and K. Hagihara, "A data distributed parallel algorithm for nonrigid image registration," *Parallel Computing*, vol. 31, pp. 19–43, Jan. 2005.
- [260] Z. Shen, J. Luo, C. Zhou, G. Huang, W. Ma, and D. Ming, "System design and implementation of digital-image processing using computational grids," *Computers & Geosciences*, vol. 31, pp. 619–630, Jun. 2005.
- [261] B. Abdalhaq, A. Cortés, T. Margalef, and E. Luque, "Evolutionary optimization techniques on computational grids," in *Proc. Int. Conf. on Computational Science*. Amsterdam, Netherlands, Apr. 2002, pp. 513–522.
- [262] H. K. Ng, Y. S. Ong, T. Hung, and B. S. Lee, "Grid enabled optimization," in *Proc. European Grid Conference*. Amsterdam, Netherlands, Feb. 2005, pp. 296–304.

BIBLIOGRAPHY

BIBLIOGRAPHY

-
- [263] I. Foster and C. Kesselman, "Globus: A metacomputing infrastructure toolkit," *International Journal of Supercomputer Applications*, vol. 11, no. 2, pp. 115–128, 1997.
- [264] C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, no. 1-2, pp. 5–43, 2003.