New Techniques for Spatial Resolution Enhancement of Hyperspectral Images

by

Patel Rakeshkumar Chandulal

(200921003)

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree

of

Doctor of Philosophy

in

Information and Communication Technology

to

Dhirubhai Ambani Institute of Information and Communication Technology

October 2014

Declaration

This is to certify that

1. the thesis comprises my original work towards the degree of Doctor of Philosophy in Information and Communication Technology at DA-IICT and has not been submitted elsewhere for a degree,

2. due acknowledgment has been made in the text to all other material used.

Signature of Student

Certificate

This is to certify that the thesis work entitled "New Techniques for Spatial Resolution Enhancement of Hyperspectral Images" has been carried out by Patel Rakeshkumar Chandulal (200921003) for the degree of Doctor of Philosophy in Information and Communication Technology at this Institute under my supervision.

Thesis Supervisor Prof. Manjunath Joshi

Acknowledgments

First and above all, the endless thanks goes to Lord Almighty for all the blessings he has showered onto me, which has enabled me to write this last note in my research work. During the period of my research, as in the rest of my life, I have been blessed by Almighty with some extraordinary people who have spun a web of support around me. Words can never be enough in expressing how grateful I am to those incredible people in my life who made this thesis possible. I would like an attempt to Thank them for making my time during my research in the Institute a period I will treasure. Although the work described here was performed independently, I never would have been able to complete it if not for the support of many wonderful people. I would therefore like to offer my sincere thanks to them.

I would like to express my sincere gratitude to my supervisor Prof. Manjunath V. Joshi with whom I have learned immensely and has had a strong influence in my development as a researcher. Thank you sir for the patience you had to supervise and direct my work, and for all the hard job of guidance, fruitful discussions, and helping me throughout all the different steps of my doctoral research endeavor for the past few years. I greatly appreciate the freedom and collegial respect you have given me. As my supervisor, Prof. Joshi has constantly forced me to remain focused on achieving my goal. His observations, guidance, and comments helped me to establish the overall direction of the research and to move forward with investigation in depth. Working with Prof. Joshi has enriched my life not only in the academic sense. His achievements, his work ethics, and his keen eye for every important detail have been an inspiration throughout all the years I have worked with him. A very big thank you also goes to Prof. Mehul Raval. There was always time to talk about my research and you were always keeping an eye on how I was doing. I have appreciated your interest and support very much, thank you!

On a broader note, I would like to thank the many people who have taught me starting with my school teachers, my undergraduate and postgraduate teachers. I wish to acknowledge all the professors of DA-IICT who have inspired me directly or indirectly. I would like to thank faculty members Prof. Chetan Parikh, Prof. V. P. Sinha, Prof. Ashok Amin, Prof. Suman Mitra, Prof. Asim Banerjee, Prof Aditya Tatu, Prof. Hemant Patil, Prof.Manik Lal Das, Prof. C. Vijaykumar, Prof. V. Sunitha, Prof. L. Pillutla,

and Prof. Rahul Dubey who have shaped my thinking about this research direction and presentation of thesis work. It is always nice remembering the excellent times of the lessons by Prof. V. P. Sinha and how this sweet experience moved me toward the initiating research work.

I express my sincere gratitude to Director, Prof. S. C. Sahasrabudhe, Registrar, Dean-AP, Convener-PGC and all other faculties who directly or indirectly helped me throughout my time at DA-IICT. I would like to thank administrative and technical staff members of DA-IICT who have been kind enough to advise and help me in their respective roles. My gratitude also extend to staff of DA-IICT resource-center.

I would like to express my gratitude to Prof. Prakash P. Gajjar (EC department, Government Polytechnic, Ahmedabad) for all his technical support, valuable inputs and remarks at several occasions. I also thank Prof. Kishor P. Upla (EC department, SVNIT, Surat) for generously sharing his time and knowledge that influenced the work and resulted in cooperative joint publication. I wish to acknowledge my friends Jignesh Bhatt and Naveenkumar with whom I have shared my research experiences since it was a joy and enlightenment to me. I would like to thank anonymous reviewers of our publications and the examiners of my thesis for their constructive suggestions that have greatly improved the publications and the thesis, respectively.

I am extremely grateful to Prof. N. N. Bhuptani, Head of Instrumentation and Control (IC) Engineering Department (2009) at L. D. College of Engineering, Ahmedabad for being a constant source of inspiration to me while I was serving at L. D. College of Engineering, Ahmedabad. Without his warm encouragement and support, I would not have gone back to college to pursue my PhD. I would like to offer my special thanks to Prof. N. B. Mehta, Government Polytechnic, Ahmedabad for his warm co-operation and recommendation to pursue PhD. A special word of thanks to Commissionerate of Technical Education, Government of Gujarat, Gandhinagar, to provide me an opportunity to pursue PhD at DA-IICT. I also express my thanks to our principal Prof. M. N. Patel, Head of IC department Prof. M. J. Modi, Prof. V. P. Patel and all the faculties of IC department, L. D. College of Engineering, Ahmedabad for their patience and support when I needed them.

I made a lot of new friends at DA-IICT, who helped me in many steps of my study. I am grateful to Mr. Vinod Mall and M.Tech students of 2009 batch for assisting me

in many different ways during the coursework, thank you very much for making the atmosphere of our classroom as friendly as possible. I am indebted to my many student colleagues for providing a stimulating and fun filled environment. My thanks go in particular to Amit Patel, Nilay Khatri, Jignesh Bhavsar, Bhavesh Dharmani, and Amit Dravid with whom I started this work and many rounds of discussions on research with them helped me a lot. I admire the distinguished helping nature of Vikas Tyagi. My special appreciation goes to him for his friendship and encouragement. I am also thankful to Shrishail Gajbhar, Milind Padalkar and Sarita Agrawal for their kind co-operation. I believe that I would not be able to name every colleague and friend separately but I thank all of them for everything that they did for me directly or indirectly.

I thank Prof. Paolo Gamba (Associate Professor, University of Pavia), for availing database of hyperspectral images to conduct experiments on super-resolution. I appreciate the travel grant support received from The ISPRS Foundation (TIF) that was funded to participate in ISPRS congress 2012, Melbourne, Australia.

I would love to cordially express my gratitude to my parents and grand parents whose understanding and determination helped me to become who I am. My grand mother could not see me reach this stage of my research and I acknowledge her in front of her memory. I take this opportunity to dedicate this work to my parents and grand parents who have made me what I am, wife and children who have given consistent support throughout my research, and my guide who had a vision in my research work. I warmly thank my uncle, aunty, cousins, in-laws, brothers, sisters and their families for their understanding, encouragement, and support in many aspects of my life.

And Finally, I know that you did not want to be named, my wife, Alpa, without your supports and encouragements, I could not have finished this work, it was you who kept the fundamental of our family, and I understand it was difficult for you. I see myself unable to even express my feelings about the love and patience that I observed from you. I can just say thanks for everything and may God give you all the best in return. Special thanks to my wonderful children: Shreeya, Pranjal and our new addition Het, the true meaning of family. For the past five years, they gave up their most favorite toy, the laptop, in favor of completing the thesis and kept on reminding me saying “Pappa, take this laptop and do you study.” I love you more than anything and I appreciate all your patience and support during PhD studies.

Abstract

Vision plays a central role for human perception and interpretation of the world. With the beginning of the space age during late 1950s presented opportunities for remote sensing of earth resources [1]. Multispectral sensors image the earth in a few strategic areas of the electromagnetic spectrum using a small number of carefully chosen spectral bands (typically 3 to 10) spread across the visible and infrared regions of the electromagnetic spectrum. These bands are not contiguous and omit many wavelength ranges. The ability of a sensor to distinguish between wavelength intervals in the bands describes the spectral resolution. Higher the spectral resolution, narrower the wavelength range for a band. The spectral resolution determines the materials discrimination ability of the sensor. The high spectral resolution of multispectral imaging was found useful for ground-cover classification, mineral exploration, and agricultural assessment to name a few. In remote sensing, acquisition of image details helps accurate localization and correct identification of minerals and vegetation, and hence better classification of the landmass.

The hardware in the remote sensing sensors limits the amount of detailed information captured (i.e., spatial resolution) whenever the spectral width of the acquired image is small. The size of the ground area expressed as $meter \times meter$ represented by a single pixel in an image defines the spatial resolution of the image. Smaller the ground area per pixel means higher spatial resolution. It depends on the sensor design and the height of the sensor above the ground. To increase the spatial resolution without affecting spectral resolution, the sensor should have a small instantaneous field of view (IFOV). But this reduces the signal power falling on the detector and hence signal to noise ratio is reduced. One can increase signal to noise ratio by widening the bandwidth of the acquired spectral band, but this reduces the spectral resolution of the image. Thus, there exists a trade-off between spatial and spectral resolutions of remotely sensed images.

Multispectral images provide higher spectral resolution (of the order of 100 nanometers) but they suffer from low spatial resolution. Improved versions of these early multispectral imaging sensors known as hyperspectral imager provide spectral width of 10 nanometers for each band of hyperspectral image (HSI). Having very high spectral resolution they provide ample spectral information to identify and differentiate spectrally unique materials [1]. Hence, presently they are used in wide range of military and civilian

applications that include target detection and tracking of objects, agriculture planning, forest inventory, mineral exploration, and urban planning to mention a few. Similar to the multispectral images hyperspectral images also suffer from low spatial resolution due to very small spectral width. Many times it is not feasible to capture the spatially high-resolution (HR) images due to the limitation in implementation such as requirement of large memory, higher transmission bandwidth, high power requirement and higher camera cost. Since HR imaging leads to better analysis, classification and interpretation one may look for algorithmic approaches to obtain the HR images. Hence, we need to perform postprocessing of the hyperspectral images (HSIs) to increase their spatial resolution and hence the image details, without affecting their spectral resolution. Super-resolution enhancement refers to an algorithmic approach to increase the spatial resolution of a low spatial resolution image by using either multiple low-resolution (LR) observations or using a database of high and low-resolution images.

Many satellites like WorldView-1, 2, 3, SPOT, Landsat, Quickbird, Ikonos, etc. capture two different types of images, namely, the high spectral but low spatial resolution multispectral (MS) images and high spatial but low spectral resolution registered panchromatic (PAN) image (auxiliary image). The reason behind configuring satellite sensors this way is to reduce weight, cost, bandwidth and complexity of the satellite. In this thesis, we develop different algorithms to enhance the spatial resolution of hyperspectral images. To start with, we first address the problem of enhancing the spatial resolution of MS images by merging information from PAN image, called multiresolution fusion, using two step approach. In the first step, the high-resolution edge details of the fused multispectral image are learned in the form of an initial estimate using discrete wavelet transform and compressive sensing (CS). We know that PAN and MS images are obtained from the same geographical region with the difference that PAN image is acquired with high spatial resolution. We assume panchromatic image and the multispectral bands are in registered form. This results in high spatial correlation between the MS observation and the coarser part of the PAN image. Our approach uses CS technique to obtain the detailed wavelet coefficients of the MS image by assuming the same sparseness of MS image with the coarser level, as well as detailed level of the PAN image. This way we obtain initial estimate of the fused image. To better preserve spatial homogenity, in the second step, we regularize it further to obtain the final solution. We restrict the solution

space by using maximum a posteriori - Markov random field (MAP-MRF) approach that imposes smoothness constraint on the fused image by using first order neighborhood for MRF prior. We make use of the initial estimate to obtain the MRF parameter.

Hyperspectral images are used for the same purpose as do MS images, but they have very high spectral dimensions that enables distinguishing the spectrally unique materials. The statistical classification (clustering) methods often used with multispectral images can also be applied to hyperspectral images by handling their high dimensionality [2]. Hyperspectral sensors like AVIRIS, Hyperion, HYMAP do not capture auxiliary HR image. In such circumstances we cannot use fusion to increase the spatial resolution of HSI. Our remaining three techniques discussed in this thesis deal with spatial resolution enhancement of HSIs using the concept of super-resolution without making use of auxiliary HR image (i.e. PAN image). The goal of super-resolution (SR) is to recover high-frequency details lost during image acquisition process which in turn increases the number of pixels in the input image. This is an inverse problem wherein the original high-resolution (HR) image has to be retrieved from the observed low-resolution data. There are large number of HR images which are consistent with the LR image. Hence, while solving such an ill-posed inverse problem, knowing the forward model alone is not sufficient to obtain a satisfactory solution. We need to add proper constrains by using priors to limit the solution space. This procedure to get a solution of the inversion problem in accordance with the prior information is called regularization. Selection of appropriate model as the prior information and use of regularization helps to obtain improved solution. In our work, we have considered different kinds of priors in regularization in order to obtain improved solution.

We make use of compressive sensing theory and estimated wavelet filter coefficients to obtain SR results for HSIs. To reduce high computational load due to large number of spectral bands of HSIs, we use principal component analysis (PCA) to reduce the dimensions and work on reduced dimensional space to obtain SR results. In the first method, we use CS based approach to obtain initial SR of the most informative PCA image which represents highest spectral variance of the HSI. Here we use LR and HR raw dictionaries having large number of atoms in the CS based framework. Using the sparsity constraint, LR test patch is represented as a sparse linear combination of relevant LR dictionary elements adaptively. Assumption of same sparsity to LR and HR images with

respect to their dictionaries gives SR image as an approximate. The final SR solution is obtained using a regularization in which AR prior model parameters are obtained from the initial SR estimate obtained using CS. SR results of the other significant PCA components are obtained using the same AR parameters and using the regularization framework. While regularization, decimation process is modeled as an averaging process.

The decimation process modeled as the averaging process represents the aliased pixel in the low-resolution image by averaging the corresponding pixels in the high-resolution image. This means, the point spread function (PSF) of sensor considered is square and is same for all spatial and spectral region. However, in practice PSF depends on several factors like camera gain, zoom factor, and imaging hardware etc. This motivates us to estimate the PSF i.e., the aliasing and then perform SR. Here our CS based approach is further extended to obtain initial SR of all significant PCA components that represent most of the spectral variance (98 %) of the HSI where the aliasing is estimated for all the significant PCA components. Here we use jointly trained LR and HR dictionaries having very less number of atoms (i.e. 1000) using training algorithm called K-singular value decomposition (K-SVD). This is onetime and offline procedure. Regularization using our new prior i.e., Gabor prior preserves various bandpass features in the final SR image. Also the use of estimated entries of degradation matrix in the form of PSF represents imaging hardware more effectively in image observation model. This leads to better solution of final SR result.

Finally, we address learning based super-resolution in wavelet domain using estimated wavelet filter coefficients. In this work, we estimate the PSF in the form of wavelet filter coefficients to take care of the degradation between LR and HR images. Here we do not consider spatially varying PSF, which is quite involved as this requires the estimation of PSF at every pixel. However, the space invariant PSF is estimated for individual spectral bands. The estimated filter coefficients are also used to learn high frequency details by using the HR training images in wavelet domain. This gives us an initial estimate of SR image for each HS band and they are used in deriving the sparse coefficients that are used as priors. The final SR image is obtained using the sparsity based regularization that also has the observation model constructed using the estimated filter coefficients. Since the cost function is differentiable, a simple gradient descent optimization is used to obtain final solution. We show the computational advantage of the proposed algorithm.

Contents

Acknowledgments iii

Abstract vi

Contents xiii

List of Figures xvi

List of Tables 1

1 Introduction 1
- 1.1 Principle of Hyperspectral Imaging . . . 2
- 1.2 Image Resolution . . . 3
- 1.3 Limitations of Optical Imaging Systems . . . 5
- 1.4 Increasing Spatial Resolution . . . 7
  - 1.4.1 Dimensionality Reduction of HSI . . . 8
  - 1.4.2 Observation Model . . . 10
  - 1.4.3 Super-resolution: An Ill-posed Inverse Problem . . . 11
- 1.5 Applications of Super-resolution . . . 12
- 1.6 Contributions of the Thesis . . . 14
- 1.7 Organization of the Thesis . . . 17

2 Literature Review 20
- 2.1 Resolution Enhancement using Auxiliary Information . . . 21
- 2.2 Resolution Enhancement without using Auxiliary Information . . . 26

3 A Model Based Approach To Multiresolution Fusion Using Compressive Sensing Theory 30
- 3.1 Use of CS Theory for Initial Estimate . . . 32
  - 3.1.1 Compressive Sensing . . . 33
  - 3.1.2 Finding Initial Estimate . . . 33
- 3.2 Regularization . . . 37
  - 3.2.1 Observation Model . . . 37
  - 3.2.2 MRF Prior Model . . . 38
  - 3.2.3 MAP-MRF Formulation . . . 38
- 3.3 Experimental Results . . . 40
  - 3.3.1 Quantitative Evaluation Measures . . . 40
- 3.4 Conclusion . . . 43

4 Super-Resolution of Hyperspectral Images using Compressive Sensing Framework **45**
4.1 Previous Work . . . 47
4.2 Theoretical Background . . . 48
4.2.1 Principal Component Analysis . . . 48
4.3 Block Diagram Description of the Proposed Method . . . 50
4.3.1 Close Approximation to Super-resolution using CS . . . 51
4.3.2 Final Solution using Regularization . . . 52
4.3.3 Estimation of Autoregressive Parameters . . . 53
4.3.4 Regularization with Sparsity Coefficients and AR Parameters . . . 54
4.4 Experiments and Results Analysis . . . 54
4.4.1 Quantitative Evaluation Measures . . . 55
4.5 Conclusion . . . 61

5 Use of Learned Dictionaries and Gabor Prior **62**
5.1 Previous Work . . . 63
5.2 Block Diagram of the Proposed Algorithm . . . 66
5.3 Proposed Approach . . . 67
5.3.1 Use of PCA and CS for Dimensionality Reduction and Sparseness Estimation . . . 67
5.3.2 Generating Trained Dictionaries . . . 69
5.3.3 Initial Estimate of Super-resolution . . . 70
5.4 Final Solution using Regularization . . . 72
5.4.1 Observation Model . . . 72
5.4.2 Regularization using Gabor Prior . . . 74
5.5 Experiments and Result Analysis . . . 75
5.5.1 Quantitative Evaluation Measures . . . 76
5.5.2 Experiments on Hyperspectral Images . . . 77
5.5.3 Experiments on 31-band Natural Hyperspectral Image . . . 77
5.5.4 Experiments on AVIRIS Hyperspectral Image . . . 79
5.6 Conclusion . . . 84

6 Super-Resolution using Optimum Wavelet Filter Coefficients and Sparsity Regularization **87**
6.1 Related Work . . . 88
6.2 Block Diagram Description of the Proposed Approach . . . 91
6.3 Estimation of Wavelet Filter Coefficients . . . 94
6.4 Learning Initial SR Estimate . . . 99
6.5 Final Solution using Regularization . . . 102
6.5.1 Observation Model . . . 103
6.5.2 Sparsity as a Prior . . . 104
6.5.3 Regularization with Sparsity Coefficients . . . 106
6.6 Experiments and Result Analysis . . . 108
6.6.1 Experiments on Single band Natural Images . . . 110
6.6.2 Experiments on Hyperspectral Data Sets . . . 113
6.7 Conclusion . . . 119

7 Conclusion and Future Research Work **123**
7.1 Conclusion 123
7.2 Future Research Work 126

List of Figures

1.1 An illustration of the concept of imaging spectroscopy principle using hyperspectral image cube [3]. 2
1.2 Image formation model showing relationship between low-resolution image and high-resolution image. 11
1.3 Schematic representation of inverse problem. The forward model is a mathematical description of the image degradation process. The inverse problem addresses the issue of reconstructing the high-resolution image from the observation or observations. 12

3.1 (a) MS Image of size $M \times M$, (b) P_{wt}: One level wavelet decomposition of PAN image. Here I, II, III and IV quadrants represent the coarse, vertical, diagonal and horizontal details each of size $M \times M$. 34
3.2 Detailed understanding of CS based approach. Construction of dictionaries D_{PA}, D_{PV}, D_{PH}, and D_{PD} from coarser, vertical, horizontal, and diagonal quadrants of P_{wt}, respectively, each of size $4 \times M^2/4$. Obtaining sparse vector $\mathbf{x}$ using coarser dictionary, and use it to obtain diagonal details of fused image, assuming mean subtracted D_{MS} (hence $\mathbf{y}$) and D_{PA} (i.e. D_{MSm} and D_{PAm}). 35
3.3 Fusion results of Band 2 for $q = 2$. (a) MS image of size 128×128, (b) Approach in [127], (c) Approach in [14], and (d) Proposed approach. . . . 42
3.4 Fusion results of Band 2 for $q = 4$. (a) MS image of size 64×64, (b) Approach in [127], (c) Approach in [14], and (d) Proposed approach. . . . 42

4.1 Block diagram of HS image super-resolution. 50
4.2 SR results on PCA-I of natural HS image for $q = 2$. (a) LR image of size 100×100, (b) Groundtruth of size 200×200, (c) Bicubic interpolation [148], (d) CS based initial estimate, and (e) Proposed approach. 55
4.3 SR results on AVIRIS HS Band 100 for $q = 2$. (a) LR image of size 100×100, (b) Groundtruth of size 200×200, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], and (e) Proposed approach. 58
4.4 Plot showing detailed performance of correlation coefficients for $q = 2$ of AVIRIS HS bands 1-141 60
4.5 Surface plots showing SAM of all pixels for $q = 2$ on AVIRIS data. (a) Bicubic interpolation [148], (b) Iterative backprojection [152], and (c) Proposed approach. 60

5.1 Detailed block diagram of proposed approach for HS image super-resolution algorithm. Here blocks are not drawn as per scale. The size of LR HSI is $M \times M \times B$, primary PCA images are of size $M \times M \times K$, secondary PCA images are of size $M \times M \times (B-K)$, initial SR PCA and final SR PCA are of size $rM \times rM \times K$, and SR HSI is of size $rM \times rM \times B$, where r is super-resolution factor. 68
5.2 Gabor prior for SR image Z_m. Here the output $G_j D_m Z_m$ represents the image details at a particular frequency band which have to match with the details of Y_m when it is passed through the same filter G_j. $\{G_j\}$, $j = 1,\ldots,Q$ represents a Gabor filter bank. 75
5.3 Experimental results on PCA-1 of 31-band natural HSI for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152] (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach. 78
5.4 Experimental results on PCA-I band of AVIRIS data for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach. 80
5.5 Experimental results on PCA-II band of AVIRIS data for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach. 83
5.6 Experimental results on PCA-III band of AVIRIS data for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach. 84
5.7 Correlation coefficient Vs band number for SR on AVIRIS data for $r = 4$ 85
5.8 Spectral reflectance plots averaged over 3×3 pixels, of groundtruth and SR image obtained using different algorithms on AVIRIS data for $r = 4$. (a) Selected regions shown using the band 100, (b) Spectral reflectance at region A, (c) Spectral reflectance at region B, (d) Spectral reflectance at region C. 86

6.1 Detailed block diagram of proposed approach for HSI super-resolution . . 92
6.2 Illustration of learning of detail wavelet coefficients for $q = 2$ using a database of HR PCA images. (a) Two level wavelet decomposition of test PCA image (LR observation). Dotted lines show wavelet coefficients to be learned. (b) Three level wavelet decomposition of HR PCA training images. 100
6.3 Randomly selected training image sets from the database. (a) LR images 64×64, (b) HR images 128×128 (for $q = 2$), and (c) HR images 256×256 (for $q = 4$) 109

6.4 SR results for $q = 4$ showing importance of initial estimate. (a) LR image of size 64×64, (b) Ground truth image of size 256×256, (c) Bicubic interpolated image, (d) Learned SR image using estimated wavelet filter, (e) Regularization result when using bicubic interpolated image as initial SR estimate, and (f) Regularization result when using learned SR image as initial SR estimate (Proposed approach) 110
6.5 Experimental results on single band images for $q = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) SR image using Yang et al. method [79], (e) Initial SR image using Db4 wavelet [161], and (f) Initial SR estimate using proposed approach 120
6.6 Experimental results on PCA-1 of 31-band natural HSI for $q = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) SR image using Jiji et al. method [161], (e) SR image using Zhao et al. method [61], (f) SR image using Yang et al. method [79], and (g) SR image using the proposed approach. 121
6.7 Experimental results on 100^{th} band of AVIRIS data for $q = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) SR image using Jiji et al. method [161], (e) SR image using Zhao et al. method [61], (f) SR image using Yang et al. method [79], and (g) SR image using the proposed approach. 122
6.8 Spectral angle error for AVIRIS HSI for $q = 4$. 122

List of Tables

3.1 Performance comparison of band 2 for $q = 2$ and $q = 4$ in terms of correlation coefficient, SSIM and MSE. 43

4.1 Quantitative comparison of SR results on AVIRIS data for $q = 2$ 59

5.1 Quantitative evaluation measures for SR of 31-band Natural hyperspectral image using different techniques for $r = 4$ 79
5.2 Quantitative evaluation metrics of AVIRIS SR for $r = 4$ 85
5.3 Spectral reflectance (%) of 3×3 pixels averaged over 196 bands at different region locations shown in Figure 5.8(a) of AVIRIS SR for $r = 4$ 85

6.1 Performance comparison showing importance of initial estimate on single band "Ganapati" image for $q = 4$ in terms of MSE between groundtruth and initial SR, and groundtruth and final SR 109
6.2 Estimation of wavelet filter coefficients and comparison of MSE between true LR and reconstructed coarser images for single band images 111
6.3 Mean squared error comparison for SR results on single band images for $q = 4$ 111
6.4 Effect of different PSFs on estimation of wavelet filter coefficients for 31-band Natural hyperspectral image for $q = 4$ (Here $\kappa = 3$) 113
6.5 Quantitative evaluation measures for SR of 31-band Natural hyperspectral image using different techniques for $q = 4$ 114
6.6 Quantitative evaluation metrics of AVIRIS SR for $q = 4$ 115
6.7 Computational time in seconds of different algorithms for LR image size 64×64 and $q = 4$ 119

Chapter 1

Introduction

Space exploration started during late 1950s that resulted in the need of acquiring earth observations from the space in order to obtain useful information for managing the resources. Earth imaging became a tool for managing these resources. Imaging the earth using carefully selected filters having narrow spectral widths was found to greatly enhance the ability to identify specific crops, plant species differentiation, study the atmosphere, oceans, snow and cloud differentiation, the solid earth, the biosphere, and many more. This defined the birth of multispectral imaging sensors with four number of spectral bands (color filters), started with the first launch of Landsat satellite in 1972. Since then many satellites carrying multispectral sensors are launched. Typically, multispectral images consist of 3 to 10 noncontiguous spectral bands spread across the visible and infrared regions of the electromagnetic spectrum, each band having typically width of 100 nanometer. This width of the band determines the spectral resolution of the sensor. Higher the width lower is the spectral resolution and viceversa. Spectral resolution of the sensor determines the materials/objects discrimination ability and higher spectral resolution results in better discrimination of objects. Improved version of these early multispectral imaging sensor known as hyperspectral imager provides spectral width of the order of 10 nanometers for each band. Having very high spectral resolution hyperspectral images provide ample spectral information to identify and differentiate spectrally unique materials [1].

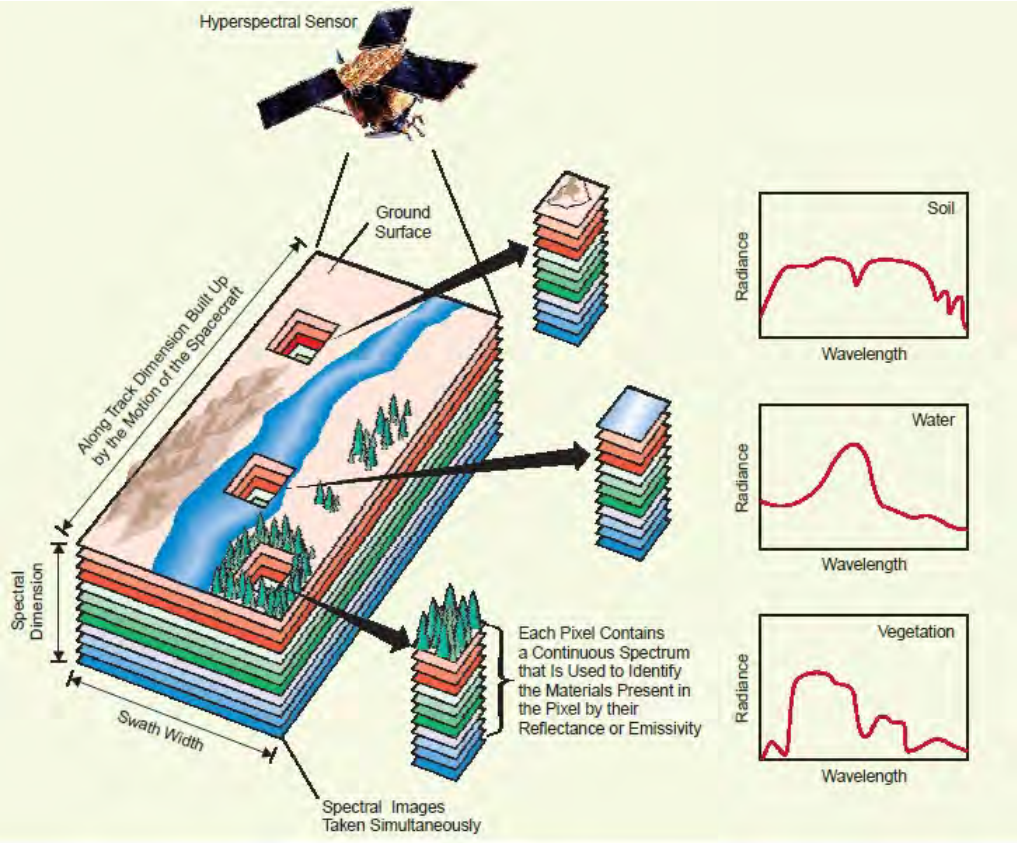


Figure 1.1: An illustration of the concept of imaging spectroscopy principle using hyperspectral image cube [3].

1.1 Principle of Hyperspectral Imaging

The materials comprising the various objects in a scene reflect, scatter, absorb, and emit electromagnetic radiation in ways characteristic of their molecular composition and their macroscopic scale and shape [3]. Hyperspectral sensors collect information in the form of reflectance spectrum in hundreds of contiguous bands (generally with a spectral range of 10 nm) simultaneously in the visible to mid infrared portion of the spectrum $0.4-2.5\mu m$, forming a three dimensional (two spatial dimensions and one spectral dimension) image cube as shown in Figure 1.1. Each pixel in the resulting image is associated with a complete spectral measurement of reflectance, called spectral signature [2], which can be interpreted to identify the material/object present in the scene. The graphs in the figure illustrate the spectral variation in reflectance for soil, water, and vegetation. They represent an evolution in technology from multispectral (MS) sensors, which typically collect spectral information in only a few discrete, non contiguous bands [1]. Being spectrally overdetermined, they provide ample spectral information to identify and

differentiate spectrally unique materials [1, 2]. Hence, they are used in wide range of military and civilian applications that include target detection and tracking of objects, precision agriculture (e.g. monitoring the types, health, moisture status and maturity of crops), agriculture planning, forest inventory, mineral exploration, coastal management (e.g. monitoring of phytoplanktons, pollution, bathymetry changes) and urban planning to mention a few.

1.2 Image Resolution

The term resolution refers to the smallest amount of change in a measured value of physical quantity that the instrument can detect. It is used to quantify the quality of various physical instruments. The high-resolution of an instrument enables us to measure the quantity with better precision. The resolution of an imaging spectrometer is classified into different types as discussed below.

- Spatial resolution: An imaging spectrometer makes spectral measurements of many small areas of the earth's surface, each of which is represented as a pixel in the hyperspectral image. The size of the ground area in $meter^2$ represented by a single pixel defines the spatial resolution of the image. Spatial resolution depends on different factors such as instantaneous field of view (IFOV), lens optics, and the height of the sensor above the earth's surface. The instantaneous field of view (IFOV) of the sensor is a measure of the ground area viewed by a single detector element (pixel) at a given instant of time. NASA's Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), has a spatial resolution of 20 $meter \times meter$ often expressed as $20m$ when flown at its typical altitude of 20 kilometers, but it has a $4m$ resolution when flown at an altitude of 4 kilometers indicating that a pixel covering $4m^2$ area represents higher resolution when compared to a pixel with $20m^2$ area. High spatial resolution allows us to capture fine details and the finer intensity transitions across different directions. Low spatial resolution of the sensor shows blockiness and increases the number of "mixed" pixels in the image. Mixed pixels present an additional challenge to hyperspectral image (HSI) data exploitation because at low spatial resolution their spectral signatures do not represent a single well-defined material.

- Spectral resolution: The ability of a sensor to distinguish between wavelength intervals in the bands describes the spectral resolution. Higher the spectral resolution, narrower the wavelength range for a band. For example, multispectral images have bandwidths of 100nm while the width of hyperspectral images are significantly less and they are of the order of 10nm. This indicates that HSI can resolve spectral features more effectively than multispectral image (MSI) because of its higher spectral resolution. High spectral resolution of hyperspectral images facilitates fine discrimination between different materials and objects based on their spectral response in each of the narrow bands.

- Radiometric resolution: Each pixel of a spectral band carry information of the image intensity in the form of binary digits called "bits". The reflected intensity at any location in a real world scene may take real values. However, it is not possible to represent these continuous range of real values in a digital computer or processor. In practice this range is divided into a finite number of levels and the intensity is quantized and the pixel value is assigned the nearest quantization level. The radiometric resolution refers to the sensitivity of a sensor to the difference in strength of the electromagnetic radiation signal and determines the smallest change in intensity level that can be distinguished by the sensing system. The increase in the brightness resolution requires more number of quantization levels and hence more number of bits for each level. A binary image has two levels; black and white, hence requires only one bit for each level. A gray scale image is usually quantized using 256 gray levels with each level represented using 8 bits. Similarly, if each color plane of an RGB image has 8 bits then at least 8 bits $\times$ 3 bands = 24 bits are needed for representing each pixel. Many multispectral sensors record data up to 8 bits in size, while the hyperspectral sensor such as AVIRIS records data using 12 bits per pixel per band. The consequences of increasing the quantization is the increase in the sensitivity of the sensor to variations in the reflected signal, thereby allowing more subtle reflectance differences from the scenes to be detected and recorded [4]. On the downside, this increases the storage space and bandwidth requirement of the imager. The images taken by a remote sensing satellite is transmitted to earth through telecommunication. The bandwidth of the telecommunication channel sets

a limit to the data volume for a scene taken by the imaging system.

- Temporal resolution: The ability of the sensor to collect imagery of the same area of the earth's surface at different periods of time is one of the most important elements for studying change detections over a period of time. Spectral characteristics of features may change over time and these changes can be detected by collecting and comparing multi-temporal imagery. For example, during the growing season, most species of vegetation are in a continual state of change and our ability to monitor those subtle changes using remote sensing depends on when and how frequently we collect imagery. By imaging on a continuing basis at different times one can monitor the changes that take place on the earth's surface, whether they are naturally occurring due to changes in vegetation cover, snow cover, or water status or due to changes caused deliberately such as urban development or deforestation etc. Temporal resolution is defined as the time interval between the acquired data for the same location. When applied to remote sensing, this amount of time depends on the orbital characteristics of the sensor platform as well as on the sensor characteristics. The temporal resolution is high when the revisiting time is low and vice-versa. Temporal resolution is usually expressed in days and is also dependent on the spatial resolution of the sensor. The higher the spatial resolution, the lower the temporal resolution is.

In this thesis, we address the problem of increasing the spatial resolution of given low spatial resolution and high spectral resolution hyperspectral image. In the rest of the thesis the term resolution is explicitly used to mean spatial resolution unless specified otherwise.

1.3 Limitations of Optical Imaging Systems

There are variery of techniques used to generate hyperspectral imagery like (i) The scanner (ii) Pushbroom, and (iii) Whiskbroom, a common one being the pushbroom imaging. Pushbroom imaging system consists of linear detector array of the charge coupled device (CCD) having number of detector elements. Each detector element projects an instantaneous field of view (IFOV) on the ground and receives photons from this IFOV. At

any instant, a row of pixels are formed and as the detector array (sensors) flies along its track, the row of pixels sweep along to generate a two dimensional image. Unfortunately, because of technical constraints, satellite remote sensing systems can only offer the following relationship between spatial and spectral resolution i.e., a high spatial resolution is associated with a low spectral resolution and vice-versa. All the reflected signals are corrupted by noise caused due to electronic circuitry in the sensor as well as due to the effects introduced during the transmission. To increase the spatial resolution the sensor should have a small instantaneous field of view (IFOV). But, this reduces the number of photons imparted to the CCD which reduces the strength of signal, while inherent noise level remains constant. Hence the signal to noise ratio (SNR) is reduced and one has to widen the sensor bandwidth for each band in order to improve SNR by allowing more photons to enter into the sensor. Unfortunately, this reduces the spectral resolution of the sensor and hence causes HSI sensor to capture images at lower spatial resolution. There remains a trade-off between the spectral and the spatial resolutions of the HSI sensor [3].

An approach to increase the spatial resolution is to use sensors with smaller detector size and use increased pixel density in the charge coupled device (CCD) through advanced sensor fabrication techniques. However, this generates shot noise that severely degrades the image quality and hence there exists a technical limitation with regard to pixel size reduction [5]. In addition, the cost of manufacturing such a high-precision sensor is high. Besides this, the reduced pixel size also decreases the number of photons arriving in a fixed exposure time, reducing the SNR. More over, the increased number of pixels per chip also increases the capacitance that results in higher noise levels and slower charge transference speed. As a result of these effects the present day CCD technology has reached the optimal level of spatial resolution, and the costs of increasing the resolution further is very high. High temporal resolution of a sensor requires larger IFOV to reduce the revisiting time which can be done by increasing the height of the sensor to increase the field of view of sensor. But increasing IFOV results in reduced spatial resolution of the sensor and viceversa. Thus, there is a kind of trade-off among spatial, spectral and, temporal resolution of the imaging sensors. Considering all these limitations, we need to perform postprocessing of the HSI in order to increase its spatial resolution.

1.4 Increasing Spatial Resolution

Hyperspectral sensors acquire images having higher spectral resolution in hundreds of bands. However, the common drawback of these sensors is relatively low spatial resolution, varying from few to tens of meters. In the process of capturing the HSI using a digital image acquisition system, there is a natural loss of spatial resolution. In addition to the optical imaging system as already discussed, there are many other phenomena that add to low spatial resolution. They include atmospheric scattering, secondary illumination effects and noise occurring within the sensor or during transmission and insufficient sensor density. The resulting low spatial resolution image has reduction in number of pixels which causes aliasing. However, in remote sensing, higher resolution image details are useful in accurate localization and correct identification of minerals and vegetation, and hence better classification of the landmass. Since HR imaging leads to better analysis, classification and interpretation of the same area one may look for algorithmic approach to obtain the HR images having larger number of pixels. Resolution improvement by applying different digital signal processing techniques has been a topic of great interest in remote sensing research community.

Fusion and super-resolution are the methods widely used to increase spatial resolution of the remotely sensed images. Many sensing platforms are equipped to capture high spectral and low spatial resolution hyperspectral or multispectral image, as well as low spectral and high spatial resolution auxiliary image (i.e., panchromatic image). The difference in spatial and spectral resolution of sensors (MS, HS, panchromatic) are due to fundamental trade-off of spatial and spectral resolutions in the design of electro-optical sensor systems. Using multiresolution fusion, information from the high spatial resolution panchromatic (PAN) image (registered) is merged with HS or MS image. As a result fused image with increased spatial resolution to that of PAN image is obtained and this has the spectral resolution of HS or MS image. The main limitation of this approach is the requirement of auxiliary image, which is not available at all time. Requirement of registered HS and PAN image poses another major difficulty in fusion.

Another kind of approach to increase the spatial resolution of HSI is known as super-resolution. The term “super-resolution” refers to such signal processing techniques that reconstruct high spatial resolution image from one or more low-resolution images. The

major objective of super-resolution techniques is to recover the high-frequency content lost during image acquisition process, without introducing considerable spectral loss. In effect, although the main concern of the super-resolution algorithms is to reconstruct high-resolution images from undersampled low-resolution observations, it gives high quality images from blurred, noisy and downsampled images. The word "super" in super-resolution represents the characteristics of the technique overcoming the inherent resolution limitation of low-resolution imaging systems. The advantages of super-resolution approach are: 1) no need of auxiliary HR image, 2) no need of additional hardware for enhancing the resolution, and 3) flexibility and reduced cost.

The super-resolution reconstruction problem is closely related to image restoration problem. The goal of image restoration is to recover an image from degradations such as blur and noise, and it does not increase the size of the image. Thus, for image restoration problem, the size of the restored image is the same as that of the observed image while it is different in image super-resolution depending on the decimation factor of the super-resolved image. Image interpolation is another problem related to super-resolution. It increases the size of image using a single low-resolution observation. Since the single image does not provide non-redundant information, the quality of the interpolated images is very much limited and the aliasing causes loss of details. Few of the interpolation methods convolve image with a filter designed to boost high-frequency contents. The drawback of such methods is that the noise is also amplified which degrades the quality. The quality of an image interpolated from an aliased low-resolution image is inherently limited even though the ideal sinc function is employed. The single image interpolation techniques cannot recover the high-frequency components lost or degraded during low-resolution sampling process. For this reason, image interpolation methods are not considered as super-resolution techniques.

1.4.1 Dimensionality Reduction of HSI

The number of spectral bands associated with the remote sensing system is referred to as its dimensionality. The main difficulty in processing hyperspectral images (HSIs) is that the number of bands can vary from few tens to several hundreds. For example, Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) hyperspectral image (HSI)

acquires 224 bands. Applying super-resolution technique to each band separately can lead to a dramatic increase of computational time of the algorithm. Besides, it increases the number of pixels stored and processed by the digital image processing system, leading to unnecessary consumption of storage and processing resources of the system. Since the hyperspectral bands are highly correlated they are inherently low-dimensional, we need to reduce dimensionality of HSI. Many methods have been developed to tackle the issue of higher dimensionality of HSIs. These methods can be grouped into two classes: (i) Feature selection, and (ii) Feature extraction. Feature selection method selects a sub-optimal subset of the original set of features while discarding the others. These methods are due to combination of a search algorithm and a criteria function [6, 7]. The search algorithm generates a subset of features and compares them on the basis of criteria function to obtain the solution to feature selection. On the other hand, feature extraction method reduce the dimensionality by mapping the feature space onto a new low-dimension subspace. These methods are more effective in representing the information content in a lower-dimensionality domain. The most common techniques to reduce the number of bands are the minimum noise fraction (MNF) transform, where an operator calculates a set of transformed features according to a signal-to-noise ratio optimization criterion, principal component analysis (PCA), in which a set of uncorrelated transformed features is generated and also independent component analysis (ICA), where a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals is used [6]. These methods perform the dimensionality reduction by retaining the components with the highest information content. The components retained are uncorrelated and the observed data is represented as linear combination of certain basis. In this thesis, we propose to use principal component analysis (PCA), a standard tool for analysis of multivariate data [8].

PCA identifies orthogonal axes for dimension reduction by performing an eigen decomposition of the spectral covariance matrix of sample data. The magnitude of the eigenvalues indicates the data variance along the component parallel to the associated eigenvector. Hence, the effective dimensionality of the data depends on the eigenvalues. We retain the first few eigen vectors corresponding to significant eigen values and high dimensional data is projected on these eigen vectors which produces a new set of decor-

related data oriented along the eigenvectors with higher eigen values and this results in a lower-dimensional multivariate data that still conveys most of the energy in the original data. This transformation incorporates most of the spectral variance of HSI data in first few principal components. Since the number of eigen vectors retained is significantly less compared to the total number of HSI bands, one cannot reconstruct original hyperspectral image exactly by inverse PCA, thus causing information loss. However, it is reasonable to assume that the spectral signature of the materials/objects of interest is present in sufficient amount in reasonable number of spectral bands. Note that the number of PCA components retained is application dependent and it can be increased at the cost of computational speed, and information loss and the reconstruction error may be made arbitrarily small in order to take care of accuracy.

1.4.2 Observation Model

In order to apply resolution enhancement algorithm, understanding of image acquisition process is required. Remote sensing camera that acquires multispectral or hyperspectral images consists of focusing lens, spectral filters, diffraction system, optical sensors, analog to digital converter (ADC), processor chip, electronic circuits and other mechanical subsystems. When the remote sensing image is captured, there is a limit to the spatial resolution at the ground due to the pixel spacing along a scan line, the line separation, and the overall optical magnification. Optical distortions such as insufficient sensor density, defocus, warping caused by variation in the track of imaging system, and diffraction limit add to degradation and reduction in spatial resolution. The observations may be blurred due to causes like optical aberration, relative motion between camera and object, limited shutter speed, and atmospheric turbulence. Furthermore, the images could be degraded by various types of noise occurring within the sensors, ADC, electronic circuits, and during transmission. To simulate all these effects for algorithm development and testing, the observed low-resolution (LR) images are modeled as warped, blurred, downsampled, and noisy version of corresponding HR image. In order to analyze the super-resolution reconstruction problem, it is required to formulate a mathematical model that represents the image acquisition process. This model, known as observation or forward model, relates the original high-resolution image to the observed low-resolution image(s). The

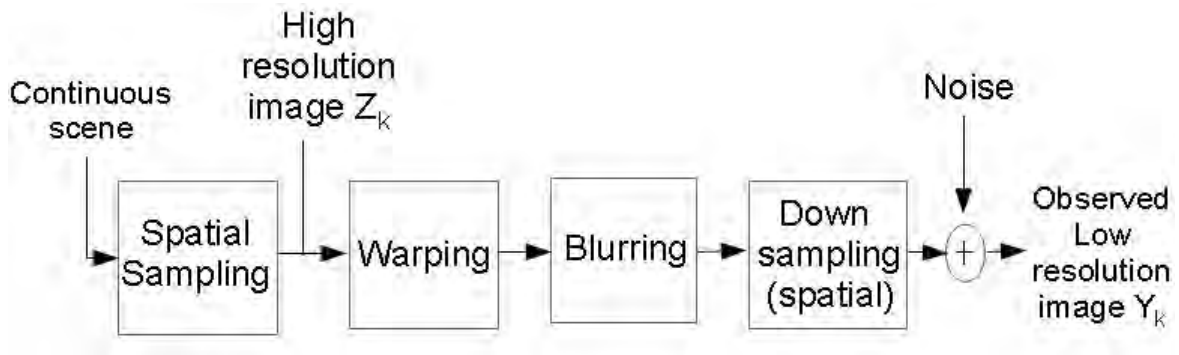


Figure 1.2: Image formation model showing relationship between low-resolution image and high-resolution image.

correct formulation of the observation model plays an important role in the success of any super-resolution approach. The most commonly used forward models for super-resolution reconstruction incorporate translation, blur, aliasing and noise in the formulation. Unfortunately, sensor point spread functions (blurring operators) are unknown, posing difficulty to construct a reliable observation model making super-resolution (SR) problem more challenging. A typical forward model is shown in Figure 1.2. Our model do not include motion or warping since we assume that the bands are registered.

1.4.3 Super-resolution: An Ill-posed Inverse Problem

The super-resolution algorithms attempt to enhance the spatial resolution of remotely sensed image corrupted due to the limitations of the optical imaging system. SR algorithm estimates high-resolution image from a single LR image or multiple LR images of the same scene, which are aligned with subpixel accuracy. The fundamental constraint for obtaining SR image is that the obtained SR image should result in the observed LR image when used in the same observation model. Obtaining SR image from LR image is an inverse problem since inverse operation has to be performed on the LR in order to obtain the high-resolution image. A schematic representation of the inverse problem is shown in Figure 1.3. The problem is ill-posed because we have more number of unknowns than known. This makes it an ill-posed inverse problem. It is difficult to invert such a forward model since there are multiple number of solutions possible. In other words, there exist infinite high-resolution images which are consistent with the original data. The forward model of high-resolution (HR) to low-resolution (LR) transformation can be a linear or non-linear transform. Because of the simplicity and mathematical tractability

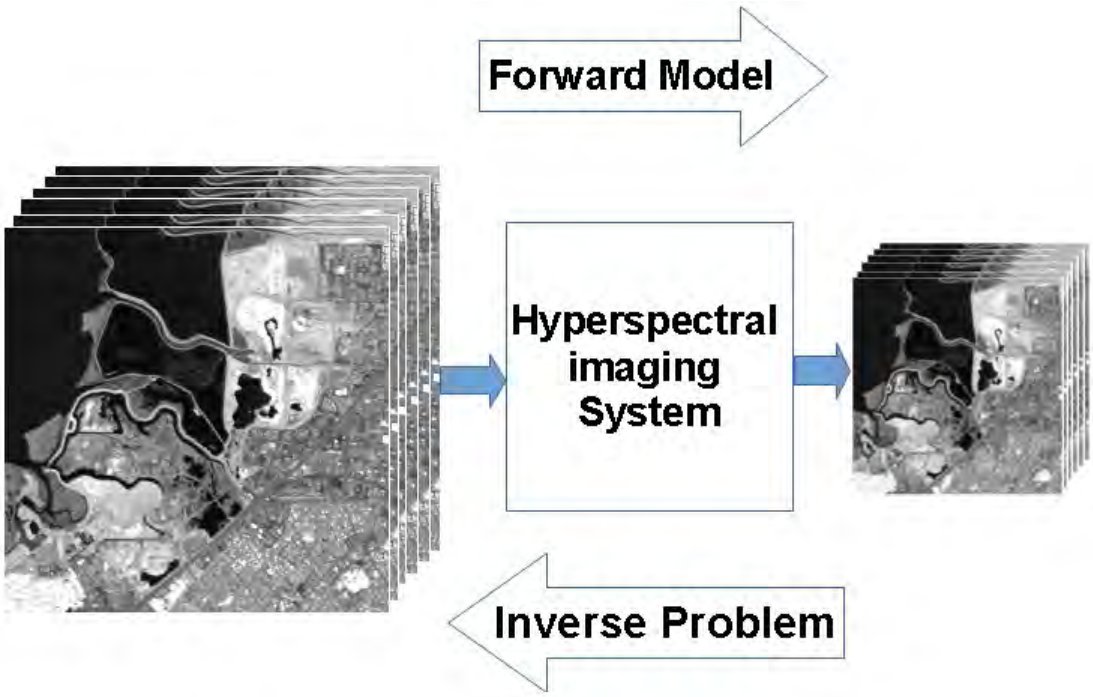


Figure 1.3: Schematic representation of inverse problem. The forward model is a mathematical description of the image degradation process. The inverse problem addresses the issue of reconstructing the high-resolution image from the observation or observations.

the researchers in the area of SR often use a linear model. While solving an ill-posed inverse problems, knowing the forward model alone is not sufficient to obtain better solution. Some form of constraints on the solution space must be included to restrict the solution space. In order to make the problem better posed i.e., to obtain better solution the researchers in SR community have used regularization. The regularization based approach solves the ill-posed inverse problems by making them better-posed using the prior information about the super-resolved image that results in favorable solution.

1.5 Applications of Super-resolution

The need of high spatial resolution images is common in many imaging applications. High-resolution images lead to better analysis, classification, and interpretation of the images. High-resolution images are of importance in many applications like remote sensing, medical diagnostics, forensic, surveillance, military information gathering, and many more. Many applications require zooming of the specific area of interest in the image wherein high-resolution becomes essential, e.g. satellite imaging, military surveillance,

medical, and forensic applications. Long range target detection has been a military priority, especially for small targets such as enemy vehicles. Images taken by an aircraft over a geographical area of interest provides the most up-to-date information about the area. Accurate detection of targets is dependent on the spatial resolution of the images. For example, if the target size is of 10 $meter \times$ 10 $meter$ and spatial resolution of image captured is itself 20 $meter \times$ 20 $meter$ then the spectral response of target represented by single pixel is mixed with other objects/materials in the vicinity of vehicle. In this case, we obtain mixed pixel and it becomes necessary to perform subpixel target detection. However, subpixel target detection is more challenging [9], and it is difficult to conclude about the presence of vehicle. Now if the spatial resolution of the image is 10 $meter \times$ 10 $meter$ or better, then a pixel value represents the target information and the target detection algorithm can determine the presence of target (vehicle) with a greater accuracy. The medical field has long been a user of image processing technology. Hyperspectral imaging has found application for the detection of the tissue and tumor [10, 11], particularly useful in surgery, clinical procedures, medical evaluation, etc. The images captured using a spectroscopy are enhanced using SR techniques which enables the doctor to focus on the area of interest that helps in better diagnosis. Hyperspectral imaging has also found applications in forensic science to detect different types of paint and latent fingerprints on the surface. Super-resolving the areas of interest in such images aids in the investigation process which leads to correct identification of a criminal.

In many applications where we are interested in saving the resources such as memory, bandwidth, power and cost, super-resolution techniques are very useful. In such applications it may not be feasible to capture HR images from the imager. Consider a hyperspectral imager mounted on the aircraft to monitor dynamic phenomena such as vegetation cover, of a wide region on the earth. Use of high-resolution imagery helps to study the dynamic behavior of vegetation. But, this may need installation of high power and costly imager on the aircraft or satellite. The transmission of the large volume of data generated by this imager needs large transmission bandwidth and power. Super-resolution techniques are very helpful in such a situation where one can employ a low-cost sensor to capture the images at lower spatial resolution. This data can be transmitted over the channel having limited bandwidth. Alternately, we can increase altitude of satellite to cover wide areas which increases the IFOV. This results in reduction of

the spatial resolution but the flight time to capture the image of the region of interest is also reduced, thereby reducing time and cost of data collection. The data received at the base station can be applied to a super-resolution technique to reconstruct high resolution images. This method effectively produces HR image similar to that of the sensor platform flown at lower height.

1.6 Contributions of the Thesis

Super-resolution is one of the most important techniques used in a remote sensing system. The remotely sensed data having multiple bands can make use of either fusion or super-resolution to improve the spatial resolution of hyperspectral images. Both these techniques have been used under different circumstances. For fusion, onboard sensors are required to capture high-resolution image of the same geographical area in addition to the LR observation. That is approaches based on fusion techniques require registered HR image, which may not be possible to capture many times. When the auxiliary HR image is not available, super-resolution enhancement provides the necessary solution of increasing the spatial resolution of the image. SR techniques enhance the spatial resolution of a low spatial resolution image by using either multiple low-resolution (LR) observations or using a database of high and low-resolution images. Many of the approaches based on multiple LR observations, extract non-redundant information from multiple observations that have subpixel shift between them. These approaches are very useful in super-resolving images captured in the optics and they require accurate registration of the observations. The quality of the super-resolved image is highly dependent on the accuracy of registration. The task of accurate registration is very difficult and complex, and is computationally intensive. In addition multiple LR images with sub-pixel shift are not available many times. In this thesis, we first propose an approach for the multiresolution fusion for enhancing the spatial resolution of multispectral image. Subsequently, a number of SR algorithms for enhancing the spatial resolution of hyperspectral data are considered. Our fusion based approach needs auxiliary HR PAN image in registered form while SR methods do not require auxiliary HR image and hence the registration is not required. The contributions of this thesis are summarized as below.

- In recent years, compressive sensing (CS) theory [12] has attracted a great deal of

attention in signal and image processing research community. Our fusion work in the thesis is based on this. Given the LR MS image and a panchromatic image having high spatial resolution, we use CS based approach in discrete wavelet transform domain to fuse HR PAN image details and LR multispectral image of the same geographical area thus obtaining what is known as multiresolution fusion or pan-sharpening in remote sensing community. A two step approach is used to perform fusion, assuming that LR and HR images are registered. These steps include: (i) A CS based approach in wavelet domain to obtain the initial estimate, and (ii) regularization using MAP-MRF approach to obtain final solution. In the first step, the required edge details of the multispectral image to be fused are learned in the form of an initial estimate using wavelet transform and compressive sensing. To do this the PAN image is decomposed using discrete wavelet transform. Since both the PAN and MS images cover the same area of the landmass it results in high spatial correlation between the MS observation and the coarser part of the wavelet transformed PAN image. Our approach uses patch based compressive sensing (CS) technique to derive an initial estimate of fusion by first computing the sparseness of observed MS patches and then using them for obtaining detail coefficients. Taking the inverse discrete wavelet transform gives us the initial estimate of the fused image. In the second step, we regularize the solution further to obtain improved solution. A maximum a posteriori - Markov random field (MAP-MRF) approach that imposes smoothness constraint on the fused image is used for regularization. The MRF model parameter is obtained using the initial estimate. A simple gradient descent approach is used to minimize the cost function. Experiments are conducted on images captured by Landsat 7 satellite sensors. The quantitative and perceptual comparison is carried out to evaluate the proposed technique w.r.t the recently proposed image fusion techniques [13, 14].

- In many cases registered PAN image is not available especially in HSIs. Besides this, HSI consists of hundreds of spectral bands (e.g., AVIRIS acquires 224 bands). Considering these aspects, we propose a new technique for super-resolving hyperspectral bands using a training database and using CS based method. We first reduce dimensionality of HSI by representing the HS observations of different wave-

lengths, as weighted linear combination of a small number of basis image planes (BIPs) using principal component analysis (PCA) [8]. Then, SR is performed on the significant PCA components. This is advantageous since most of the information is contained in few PCA images and super-resolving them do not degrade the spatial characteristics of the SR image but reduces the computational burden. Our approach uses LR and HR dictionaries formed using large number of raw images obtained by random sampling of raw patches having similar properties in order to obtain an approximation to SR by using CS framework. Since the approximation to SR image is patch based, dependency among neighboring patches is not considered while obtaining the initial solution. Hence further regularization is carried out in order to obtain a better solution. We obtain the final solution using a regularization framework in which the sparse coefficients obtained by CS and the autoregressive (AR) parameters obtained from the initial SR estimate are used as the priors. Remaining significant PCA components are bicubically interpolated and regularized using the AR parameters obtained from the most significant PCA component. Inverse PCA on the SR of significant PCA components gives us the SR for the entire hyperspectral cube. Experiments have been conducted on two different kinds of hyperspectral images. First one consisting of 31 band natural HSI captured in controlled laboratory environment and the other is real 224 band HSI captured by Airborne Visible/Infrared Imaging Spectrometer (AVIRIS).

- Learned dictionary provides more compact representation of the signal compared to raw dictionary which simply samples large number of patches. Use of learned dictionary results in substantial reduction in computation while estimating the initial approximation to SR. This motivates us to consider the use of learned dictionary in our next work. Our CS based approach is further extended here to obtain a better solution by learning initial SR estimates for each of the significant PCA components. Unlike the previous approach, in this case we estimate decimation using the initial SR estimate and use it while regularizing the solution. Here we use a new prior called "Gabor prior" as a prior terms that imposes the condition that the degraded SR image and initial SR image should possess the bandpass features close to that of the LR test image and SR image, respectively. Experiments are conducted on

two different kinds of HSI data sets having 31 bands and 224 bands, respectively. Quantitative comparisons indicate that the proposed method enhances the spatial details with minimum spectral distortions.

- Most of the research on SR of HSI assumes implicitly or explicitly that each LR pixel in a band is obtained as a equally weighted sum of pixels of corresponding HR pixels [15, 16]. That is the degradation consists of averaging of HR pixels. In practice, PSF which represents the degradation between LR and HR images depends on various factors such as fill factor of CCD array, camera gain, zoom factor, imaging wavelength etc [17]. This results in spatially and spectrally varying PSF of the degradation function. In our next work, we estimate the spectrally varying PSF in the form of low-pass wavelet filter coefficients to take care of the degradation between LR and HR images. Our work do not consider spatially varying PSF, which is quite involved as this requires the estimation of PSF at every pixel. We estimate the space invariant PSF for individual spectral bands. Wavelet decomposition on a set of HR training images is then used to learn the high-frequency details that uses our estimated filter coefficients. The final SR image is obtained using the sparsity based regularization that has the observation model consisting of the estimated PSF corresponding to low-pass wavelet filter coefficients. Effectiveness of the proposed method is tested by conducting experiments on three different data sets: (i) Single band natural images (ii) Natural hyperspectral image acquired in controlled laboratory environment, having 31 bands, and (iii) Remotely sensed HSI (AVIRIS) having 224 bands. Visual and quantitative comparison show that our method enhances spatial resolution of HSI without introducing considerable spectral distortions.

1.7 Organization of the Thesis

In this thesis, we propose methods to increase spatial resolution of remote sensing (multispectral/hyperspectral) images using multiresolution fusion as well as super-resolution based approaches. We use compressive sensing framework in these approaches. We propose a technique to estimate the decimation due to low-resolution sampling and use the same in the forward model. Use of estimated wavelet filter coefficients in learning based

super-resolution and observation model is also demonstrated. The super-resolution problems in our work is solved using regularization framework. We make use of priors such as autoregressive (AR), Gabor and sparseness to obtain the final solution. Organization of the thesis is as follows.

Many researchers have attempted to improve spatial resolution of hyperspectral images using different methodologies such as multiresolution fusion using auxiliary PAN or MS image, super-resolution using multiple LR observations, single LR observation and, spectral unmixing based approaches followed by super-resolution. Chapter 2 provides a review of the existing spatial resolution enhancement approaches proposed in the remote sensing literature.

In chapter 3, we propose a model based multiresolution fusion using registered PAN image. We suggest a novel approach based on CS in the discrete wavelet transform (DWT) domain for obtaining initial fused image. Use of regularization to obtain the final fusion result using MAP-MRF framework is demonstrated. Model based algorithms require knowledge of the imaging processes in order to formulate a model for spatial analysis. Here we use a linear observation model.

We introduce a CS based approach to obtain the SR of primary (significant) principal components of LR observations in chapter 4. Raw dictionaries constructed by using LR and HR images are used in CS based approach to obtain close approximation to SR of the first principal component. We reconstruct the final SR image using regularization where we make use of autoregressive and sparsity based priors. The AR model parameters are derived from the close approximation (initial estimate). Super-resolution of other primary principal components is obtained without use of CS but using interpolation and making use of the AR parameters of the first principal component.

Super-resolution technique using CS based approach is extended further to all primary PCA components in chapter 5. Here jointly trained LR and HR dictionaries are used in CS based approach to obtain initial SR for each primary PCA component. Using initial approximation we estimate the degradation that also includes aliasing in the image formation model. The problem of SR is solved once again by using regularization framework. A new prior called Gabor prior which is based on Gabor filters is used to preserve different frequency components in the final SR.

In chapter 6, we address the problem of super-resolution using wavelet based learning

in which wavelet filter coefficients are estimated from the training database of hyperspectral images. In this work, we estimate the PSF in the form of low-pass wavelet filter coefficients to take care of the degradation between LR and HR images. Here also the problem is solved using regularization in which we make use of sparsity based prior.

Finally, we summarize our work and conclude in chapter 7. We discuss the further challenges and directions for future research.

Chapter 2

Literature Review

In practice, few satellite sensors capture high spatial resolution images ($< 5m \times 5m$) onboard. For example Quickbird satellite collects panchromatic (PAN) image of spatial resolution $0.7m$ and four multispectral bands of $2.8m$ resolution. Recently launched $WorldView-2$ satellite carry an imaging instrument specifically designed to meet the requirements of very high spatial resolution and more number of spectral bands. It provides PAN images of $0.46m$, and eight multispectral bands having spatial resolution of $1.84m$. Although it acquires high spatial resolution images, it fails to provide better discrimination of many surface materials and environmentally relevant information due to their limited spectral resolution. This limitation resulted in evolution of high spectral resolution senors known as the hyperspectral imaging spectrometers. Hyperspectral sensor such as AVIRIS provides very large number of narrow spectral bands, but their spatial resolution is limited. For airborne hyperspectral sensors such as AVIRIS, HYDICE, and HyMap, the spatial resolution is dictated largely by the height of the aircraft [18]. As the height increases the extent of coverage also increases but this decreases the spatial resolution. At lower spatial resolutions, the averaging effect on the pixels degrades the performance of spectral detection algorithm [19]. Due to hardware cost, fabrication limitations, and trade-off between spectral and spatial resolutions it is expensive and difficult to acquire the hyperspectral images with higher spatial resolution. Also it is difficult to improve upon the current resolution level by using the approach of hardware modification. Hence we need to perform post processing of the HSIs in order to improve the spatial resolution. The magnification of an image captured with the current resolution level of the sensors by using simple interpolation techniques introduces visible artifacts. Spatial

resolution enhancement using algorithmic approach is a topic of great interest among researchers [20, 21, 15, 22, 16, 23, 24, 25, 26, 27, 28]. The pioneer work on resolution enhancement was proposed in 1984 by Tsai and Huang [20]. They reconstructed HR image from a set of aliased LR images. Since then various approaches for increasing spatial resolution of different kinds of imagery like single band natural image, multispectral image and hyperspectral image have been proposed [29, 30, 31, 32, 21]. The approaches include: deterministic and stochastic based regularization, nonuniform interpolation, projection onto convex sets, iterative backprojection, adaptive filtering, multi-resolution fusion, and learning based approaches. The difference among these approaches is based on the reconstruction method, observation model, processing domain used (frequency or spatial), use of auxiliary information and so on. We categorize different resolution enhancement methods of remotely sensed images into two main categories. They include the methods which require registered auxiliary information of the same geographic area and, those which do not require the same. The techniques based on auxiliary information use the multiple LR observations with subpixel shifts or HR observation of the same geographic area to extract spatial information for resolution enhancement of target images. The techniques based on learning or compressive sensing do not require auxiliary information of the same geographic area and they enhance the spatial resolution from single LR observation. In this case, only a single undersampled and degraded input image is available. Here, the task of improving spatial resolution comprises of recovering the missing spatial information using a database of high-resolution images which may not be of the same geographical area.

2.1 Resolution Enhancement using Auxiliary Information

Over the last three decades many researchers have attempted to solve the problem of the spatial resolution enhancement of the HS images using auxiliary information. These methods make use of additional information in the form of (i) multiple LR observations [21, 33, 34, 5, 35], (ii) HR multispectral images [36, 22, 37, 38], and (iii) HR panchromatic image [39, 15, 16, 23, 40, 28, 41, 38, 42]. Methods based on multiple LR observations

are generally referred to as multiframe resolution enhancement methods. These methods enhance resolution using several subsampled and subpixel shifted low-resolution images of the desired geographic location, can be either obtained as a set of images captured over different times by using a single imaging device, or taken at the same time with multiple imaging devices. For example, the sensing system can be mounted on a moving platform (e.g. aircraft), and the vibrations associated with the platform can be used to generate the subpixel shifted images. Since, the observations have subpixel shifts and are aliased, the available non-redundant information in each of these observations can be used to construct a high-resolution image for the given low-resolution observation. These approaches mostly comprised of two stages: (i) registration, and (ii) reconstruction [20, 43, 44, 33]. In the registration stage, motions between observations are estimated with sub-pixel accuracy [45, 46, 47, 48]. Reconstruction stage is based on an appropriate model that makes HR image to undergo appropriate shifts, blur and downsampling operations, to generate the corresponding LR images. Modeling of imaging process requires the knowledge of blur and downsampling process which are often unknown. Few techniques perform blind super-resolution enhancement in which the knowledge of blur is not required [49, 50]. Also the performance of motion based approaches are highly dependent on accuracy of motion estimation or the registration which is a difficult task. Due to iterative nature of these algorithms, computational cost is high. Towards this end algorithms with reduced computational cost are also proposed in [51].

Wilson et al. [21] used multiple LR HSIs which are assumed to be registered. Here, the solution space is restricted by using visual cue. In [33] a set-theoretic method i.e., projection onto convex sets (POCS) is used to combine the information from multiple LR HSIs to obtain HR HSI. Another method proposed by Chan et al. [34] used multiangular LR HSIs that are registered using thin plate spline nonrigid transform and reconstruct HR HSI using Delaunay triangulation-based nonuniform interpolation method. A different approach by Zhang et al. [5] uses a maximum a posteriori (MAP) based multi-frame SR algorithm. They used principal component analysis in order to reduce the dimensionality of the HSIs in order to reduce the timing complexity. Xu et al. [35] utilized MAP model to integrate the complementary information available in multiple low-resolution images having sub-pixel shifts. They reconstruct a classification map of higher resolution from the multiple images having low-resolution. Using three different prior models namely Lapla-

cian, total variation, and bilateral total variation they regularize the sub-pixel mapping in order to improve the solution.

Few imaging devices like compact airborne spectrographic imager (CASI) and earth observing-1 (EO-1) capture high spatial resolution multispectral images in addition to hyperspectral images. Orbview-4 (warfighter) captures high-resolution PAN and MS images in addition to hyperspectral images. Researchers have proposed methods for joint processing of these images in order to obtain a fused hyperspectral image having spatial resolution of the multispectral and spectral characteristics of the observed hyperspectral images. The Landsat Thematic Mapper (TM) sensor acquires seven band multispectral images. But the resolution of band 6 is low when compared to other bands. This difference of resolutions causes difficulties in analyzing TM data, since band 6 has to be discarded due to resolution mismatch. Nishii et al. [36] proposed a statistical approach based on multivariate normal distribution to enhance the spatial resolution of band 6 using remaining bands. They use global predictor for band 6 based on conditional distribution to obtain the high-resolution raw image. Then using linear combination of high-resolution data correction is applied to obtain final HR image. In a different approach Gomez et al. [22] proposed the fusion of HS and MS images using the wavelet based method in order to enhance resolution of HSI. The method reported in [37] by Eismann and Hardie employs generalized MAP framework that makes use of stochastic mixing model to obtain high-resolution HSI. Explicit spectral relationship between MS and HS image is not required in this method. Recently, the approach proposed by Simoes et al. [38] performed fusion of HSI and MSI as well as HSI and PAN in reduced dimensional subspace by minimizing a convex objective function that has two quadratic data-fitting terms and an edge-preserving term. Their data-fitting terms account for blur, downsampling, and noise. A Total Variation based regularization aligns discontinuities across the reconstructed HR hyperspectral bands.

Remote sensing imaging devices such as Australian resource information and environmental satellite (ARIES), Naval earth map observer (NEMO), and Orbview-4 (Warfighter) collect PAN image having the high spatial resolution and many spectral bands having low spatial resolution. Various techniques have been proposed to increase the spatial resolution of hyperspectral images using the high-resolution PAN image. These algorithms fuse the high frequency details of HR image into the LR HS image to enhance its spa-

tial resolution. Zomet and Peleg [52] suggested multisensor SR by considering statistical redundancy among the data of sensors. They use non-registered images captured by different sensors and the spatial resolution is improved by using a local photometric affine alignment along the edges. Winter and Winter [15] replace first PCA component of LR HSI with HR PAN image for resolution enhancement, but performance of this method decreases when there is less correlation between the spectral responses of the PAN and HSI images. First PCA component represents the intensity component, hence the resulting HR HSI has only the spatial intensity variations. The disadvantage is that the spectral characteristics are not better preserved. In addition to resolution enhancemnet the authors in [15] also proposed a physical model based on linear mixing and used it to sharpen HS imagery using high spatial resolution panchromatic (PAN) image. An alternative approach fuses the high-resolution PAN image to the HSIs and then computes a mixture model based on the fused image [53]. Another algorithm proposed in [16] uses MAP estimation method and exploits the use of localized correlations using spatially varying statistical model to obtain HR HSIs using co-registered PAN image.

Eismann and Hardie described MAP formulation in conjunction with stochastic mixing model (SMM) to restrict the solution space in order to obtain HSI with increased spatial resolution in [23]. In [40] Capobianco et al. fused PAN data with Hyperion HSI using two different linear injection models, namely single spatial detail (SSD) and the band-dependent spatial detail (BDSD) models. In the SSD model same PAN image is used to enhance all the bands of HSI data, while in BDSD model an optimum image detail is extracted from the PAN image and the same is fused with the HSI, thereby giving more accurate results than SSD model. A different approach in [28] performs fusion of PAN image and HS images by considering minimum spectral angle distortion as a criterion. This helps to reduce spectral distortion in HR HSIs. Bar et al. [41] combined spectral and spatial analysis for detection and classification, respectively. In the detection stage they use high spectral resolution HSIs to locate the target and in the classification stage high spatial resolution PAN image is fused with low spatial resolution HSIs to reduce the false alarms.

Spectral mixture analysis (SMA) is one of the soft-classification approaches which models the total reflectance in a pixel as the linear combination of reflectance from each endmember (material) using linear mixing model (LMM). It predicts the proportion of

each endmember within each pixel using this model. A set of proportional images are produced which gives subpixel information for each member. However, location of each endmember in the mixed pixels is not decided by SMA. Super-resolution mapping predicts the location of endmembers within a pixel based on proportional images produced using SMA, so that the SR pixels corresponding to the similar kind of materials are placed in the neighbourhood. Generally spectral mixture analysis (SMA) based techniques improve spatial resolution without using auxiliary information. Few of the researchers utilized auxiliary information before performing either spectral mixture analysis or super-resolution mapping to improve performance of their algorithms [54, 53, 55, 42]. These techniques utilize auxiliary image hence they cannot be strictly classified as SMA methods. Gaecia *et al.* [54] proposed a technique to enhance resolution of HSI using fusion and spectral mixture analysis (SMA). They fuse detailed information of HR image into HS data to obtain higher resolution of HSI and then use it in the linear mixing model. In [53] Winter and Winter used a LMM based joint end-member determination and the unmixing algorithm to combine HR PAN with LR HSI in order to create HR HSI. In a different approach, Nguyen et al. [55] exploited the spectral information of multispectral image and spatial information of HR panchromatic image. They first obtain proportion images corresponding to endmembers. Then using fusion as an additional source of information, super-resolution mapping is performed using the Hopfield neural network (HNN) optimization. The method by Zhao et al. [42] proposed recently, combines SMA results and spatial information of PAN image into the SR process using sparse regularization to preserve edge sharpness and spectrum consistency in the super-resolved HSI.

A common limitation of the techniques based on auxiliary information is the requirement of supplementary spatial information associated with test HSI. Besides this, requirement of accurate co-registration of images to achieve acceptable results is also a limiting factor of these methods.

2.2 Resolution Enhancement without using Auxiliary Information

Availability of auxiliary information can be expensive or sometimes it may not be captured onboard. Hence many indirect approaches are proposed to increase spatial resolution of HSIs. In these techniques HSIs exploit the use of their own information content and they are generally based on spectral mixture analysis (SMA) [56, 57, 58, 59, 60, 61, 62, 63, 64], and/or they use learning based approaches [65, 24, 26, 66, 67]. These methods are also referred to as single frame super-resolution techniques since they make use of the single LR observation as the acquired data while performing SR. Note that resolution enhancement without using auxiliary data is severely underconstrained since we have the limited information about the geographic area. However, it can be considered as a more general approach because the resolution improvement can be achieved with out the use of underlying groundtruth.

Hyperspectral images consist of pixel intensities that may be due mix of several materials because of their limited spatial resolution. The acquired pixels within an imagery correspond to discretely sampled reflectance. These pixels are constituted as a result of reflections from different targets hence are named as mixed pixels. The resultant reflectance constitutes the spectral reflectances for the separate component materials, each weighted according to its relative abundance on the surface leading to linear mixing model (LMM) for reflectance measurements. The linear mixture model has been used by many researchers to analyze hyperspectral images [68, 69, 70, 71, 72, 73]. The model represents the image pixel intensities as linear combination of constituent spectra known as endmembers. Additionally the abundances are constrained by non-negativity and sum to one. The set of endmembers should include all material types present in the image. Linear SMA involves finding the spectrally unique signatures of pure ground components referred to as endmembers, and to obtain abundance map by expressing mixed pixels as linear combinations of endmembers. Several algorithms have been proposed to extract endmembers from hyperspectral data automatically, like pixel purity index (PPI) [69], automated morphological endmember extraction (AMEE) [74], optical real-time adaptative spectral identification system (ORASIS) [70], N-finder algorithm (NFINDR) [72], iterative error analysis (IEA) [73], etc. Plaza et al. have rigorously compared different

endmember extraction algorithms using a unified framework in [75]. A variety of approaches based on spectral mixture analysis using LMM have been proposed to address the problem of spatial resolution enhancement of HSIs. Brown et al. [58] estimate landcover components by sub-pixel processing using linear support vector machines (SVM). Here, SVM automatically selects the relevant pure pixels and determines number of endmembers in the region of interest. These techniques provide more accurate representation of landcovers than the original LR HSI. The advantage of SMA based methods is that there is no need of auxiliary spatial information. They provide the abundances of the endmembers within a pixel which is very useful in determining the presence of object of subpixel scale in remote sensing applications. The accuracy of SMA is increased by using accurate method for selecting pure endmember spectra using visual and semi-automated approaches in [59] in order to obtain detailed view of their study area.

In SMA based subpixel processing the spatial dependencies of materials in mixed pixels is not considered at subpixel level. Thus SMA based processing acts as an initial stage for the spatial resolution enhancement of HSIs. These methods do not exploit spatial and spectral informations to their full capacity. Hence super-resolution mapping (SRM) is performed to enhance the spatial resolution of the HSI by exploiting the spatial dependency of pixels. Atkinson et al. [56] presented an algorithm for spatial resolution enhancement using spectral mixture analysis followed by sub-pixel target mapping. High resolution pixels are placed based on spatial correlation among them using a distance-weighted function. In the algorithm proposed by [62] the spectral unmixing is first performed to determine proportion of endmembers in each pixel. Subpixels are then located by SR mapping performed either by simulated annealing or pixel swapping in unsupervised way. However, the limitation of these kinds of algorithms is the requirement of high computational load because of the large number of spectral bands of HSIs.

Some of the approaches perform super-resolution mapping using learning based methods like Hopfield neural network (HNN) [65, 55, 76] and backpropagation neural network (BPNN) [24, 26, 66]. Here, given the remotely sense data, a set of training images is used for learning the SR attributes. A method proposed by Gu et al. [24] first obtains abundance map using linear spectral mixture analysis (LSMA). Then based on spatial correlation of landcovers, SR mapping is performed using BPNN to enhance the spatial resolution. They use LR training images to determine the parameters used in SRM and

BPNN. In a similar work, Mianji et al. [26] used LR images and its downsampled versions to train the BPNN in learning based SR. A mean filter is used as the downsampling operator and the super-resolution is performed by considering spatial correlation of different materials present in the HSI. Here, they used test HSIs themselves as the training data to achieve better coherence between the results of the enhanced and the original HSI. In a different algorithm Zhao et al. [67] proposed example based super-resolution mapping that uses support vector regression. They do not use explicit formula to describe prior information about subpixel spatial pattern, rather they learn the nonlinear relationships between the coarse fractional pixels and corresponding labeled subpixels from the best-matching high-resolution training data.

In recent years, sparse representations of signals have attracted a great deal of attention among signal and image processing researchers. Olshausen and Field [77] showed that a natural image can be represented with a relatively small number of basis functions chosen from over-complete set. Compared to methods based on orthonormal transforms or direct time domain processing, sparse representation usually offers better alternate for efficient signal modeling [78]. The use of sparsity for SR of single wideband natural [79, 80, 81, 82, 83, 84] as well as multi-band remote sensing images [85, 86, 61] is explored by many researchers . CS based approaches use trained dictionaries of HR and LR patches and assume same sparseness for LR and HR patches in order to obtain SR image. Here, sparse coefficients for every LR patch are found using trained LR patch dictionary, and these coefficients are used in generating the HR output. To eliminate the need of searching for pure signatures in the input data and overcome limitations of endmember extraction algorithms, Iordache et al. [86] proposed semisupervised approach using linear SMA to unmix the pixels of LR observation. They represent the mixed pixels as the linear combinations of a number of pure spectral signatures from publicly available library, using sparsity based l_1-minimization algorithm to increase number of pixels, thereby increasing resolution. Zhao et al. [61] utilize sparseness property of HS images in spatial and spectral domains. They use the sparse representation and linear spectral mixing model for obtaining the initial SR. To maintain spectral consistency regularization based on linear spectral mixing model is performed in order to obtain final SR HSIs. Their method is implemented on HS data directly without using dimensionality reduction algorithm leading to high computational cost.

Our proposed approaches in this thesis are based on learning, CS theory as well as wavelets. Hence we also discuss few research works of SR of HSI using wavelet transforms. There are considerable number of techniques in which wavelet decomposition is used to increase the spatial resolution of remote sensing images [87, 22, 76, 27, 88]. These methods are based on the decomposition of the image into multiple levels based on their local frequency contents. The wavelet transform decomposes images into a number of new images each having different spatial resolution. Need of the registered HR auxiliary information is the main limitation of these methods. Besides, these methods use fixed wavelet basis like Db4 in their implementation and they require accurate co-registration to achieve acceptable results. Mertens et al. [76] proposed use of predicted wavelet coefficients to obtain SR image. They learn relation between approximate and detail coefficients using training data in neural network, without making any assumption about data distribution. In a recent approach, Li et al. [88] characterized the wavelet coefficients by a mixed Gaussian distribution and the dependencies between the coarser and the finer scale wavelet coefficients were modeled as prior by using the universal hidden Markov model and the problem was solved as an maximum a posteriori (MAP) framework. Recently, learning based SR approaches for single wideband and multiband images have been explored by the researchers to solve the super-resolution problem [89, 90, 91, 92, 93, 94, 95, 96, 26, 79, 97]. These methods use a database of HR images or LR-HR image pairs in order to learn the high frequency details for SR.

Currently, spatial resolution enhancement of remote sensing images has become one of important research areas and researchers use different approaches such as fusion, super-resolution, or combination of both in order to obtain the enhanced HSI. Several research articles including special issues [98, 99, 100, 101] have been published for multi-frame resolution and single frame super-resolution for single wideband images. In addition many books are published in the are of fusion of remotely sensed images [102, 103, 104, 105, 106] and super-resolution techniques [107, 108, 109, 110].

Chapter 3

A Model Based Approach To Multiresolution Fusion Using Compressive Sensing Theory

Many remote sensing applications like urban-area monitoring, map updating, precision farming, land use classification, and hazard monitoring require images with both high spatial and high spectral resolution. Given technological limitations related to spatial and spectral resolution of imaging sensors, MS images having high spectral resolution are acquired with a larger IFOV i.e., lower spatial resolution than panchromatic images having lower spectral resolution. Many of the earth observation satellites like Advanced Land Imager (ALI), IKONOS, Indian Remote Sensing satellites (IRS), Landsat, Systeme Probatoire d' Observation de la Terre (SPOT), and recently launched Worldview-2 provide data consisting of a panchromatic channel of high spatial resolution and several multispectral channels at a lower spatial resolution. Acquisition of panchromatic image enables the accurate geometric analysis of materials and the MS channels provide the spectral information, required for accurate discrimination between classes of different ground materials. A combination of spatial features of high-resolution PAN image and spectral features of MS can add valuable information in these applications. Image fusion refers to the algorithmic approach to generate images with high spectral as well as spatial resolution by combining high spatial resolution panchromatic (PAN) image with the high spectral resolution multispectral (MS) images. An image obtained using a fusion tech-

nique improves the interpretation of the image. An ideal fusion technique has to account for the enhancement of high spatial resolution as well as minimization of the spectral distortion.

Many researchers have investigated strategies for the fusion of multiresolution satellite images. Among the existing multiresolution fusion techniques, the methods most widely used are: intensity-hue-saturation (IHS)-based methods [111, 112], high-pass filter (HPF) [113], principal component analysis [114], Brovey transform [115] and the wavelet-based approach [116]. In these methods, first of all multispectral images are upscaled before the fusion to the spatial resolution of the high-resolution PAN image using bicubic interpolation. IHS based method is one of the most commonly used fusion techniques for sharpening of the multispectral images. IHS fusion converts a color MS image from the RGB space into the IHS color space which decomposes it to intensity and color components. In the IHS space, spectral information is reflected on the hue and the saturation while intensity change has little effect on the spectral information. The fusion is obtained by replacing the intensity (I) component in IHS image by the PAN image and converting the IHS back to RGB domain. In the high-pass filter (HPF) based fusion, a high-pass filter kernel is used to filter the high-resolution PAN data. Subsequently HP filtered image is added to each MS band by taking care of mean, standard deviation and spatial resolution ratio. It gives acceptable results, but the edges are too much emphasized. The PCA transform based fusion approach first obtains the principal components of the multispectral bands where the first principal component contains the most information (i.e. variance) of the MS image. After that, the first principle component is substituted by the panchromatic image. Applying inverse PCA results in new RGB (Red, Green, and Blue) bands of multispectral image having spatial resolution of PAN image. Being statistical method it is sensitive to the area to be sharpened hence the fusion results may vary depending on the selected MS image. Brovey transformation is due to the result of combination of arithmetic operations and normalizes the spectral bands before they are multiplied with the panchromatic image. The spectral properties, however, are usually not well preserved by this method. In a different approach based on multiresolution fusion, wavelet transform is widely used due to its compactness, orthogonality and the availability of directional information. In this method, a wavelet transform is applied to the PAN image, resulting in a four components namely coarser, horizontal, vertical, and

diagonal. Coarser component contains low-resolution information while remaining three components have the spatial details. The LR coarser component is replaced by the MS band and the process is repeated for all MS bands. Applying inverse wavelet transform results in the fused MS image. More recent works on the multiresolution fusion can be found in [117, 118]. Recently, the authors in [14] have proposed a learning based approach for multiresolution fusion using contourlet based learning approach. Zhu *et al.* [119] have proposed fusion using compressive sensing theory. They construct dictionary of LR-HR pairs using MS image and PAN image and obtain sparse representation of image patches by estimating sparse coefficients and finally reconstruct HR multispectral image.

In this chapter, we propose a new two step approach for multiresolution fusion in remotely sensed multispectral images. The edge details of the fused multispectral images are learned using wavelet transform and compressive sensing. Given the registered panchromatic (PAN) image and a multispectral (MS) image, we first decompose the PAN image using discrete wavelet transform. Compressive sensing (CS) technique is then applied to derive an initial estimate of fusion by first computing the the sparseness of observed MS patches and then using them for obtaining detail coefficients. Taking the inverse discrete wavelet transform gives us the initial estimate of the fused image. In the second step, we model the image acquisition process using a linear system and solve the fusion problem by formulating it as a maximum a posteriori (MAP) framework that enforces smoothness constraint on the fused image. The cost function consisting of a data fitting term and the prior term is minimized using simple gradient optimization technique. The quantitative and perceptual comparison is carried out to evaluate the proposed technique w.r.t the recently proposed image fusion techniques.

3.1 Use of CS Theory for Initial Estimate

Compressive sensing presents a new approach in signal processing for sparse signal recovery [12] which directly measures a compressed representation of the signal. The key assumption is that the most signals that arise in nature are sparse whose digital representation requires few nonzero coefficients. CS exploits this sparsity, by allowing a digital signal to be reconstructed using only few linear measurements when compared to the size of the original signal. As long as the measurement matrix satisfies a Restricted Isometry

Property (RIP), the exact signal recovery is possible from these measurements [120].

3.1.1 Compressive Sensing

Suppose $\mathbf{x}$ is an unknown digital image vector in R^N that we want to acquire. Normally this should require N pixels. But if we know apriori that $\mathbf{x}$ is compressible in certain transform domain (e.g. wavelet, Fourier), then as per CS theory we can acquire $\mathbf{x}$ by measuring only M linear projections rather than all N pixels. If these projections are properly chosen (projection matrix or the measurement matrix satisfy the RIP), the size of M can be smaller than N, the size of image. If the image is K sparse, random projection works if $M = O(Klog(N/K))$ [121]. Mathematically, under the sparsity assumption, a signal $\mathbf{x} \in R^N$, can be recovered by solving the l_1-minimization using standard optimization tool such as linear programming [122] i.e., the problem can be posed as

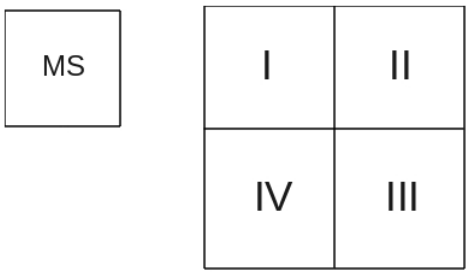
$$\min_{\mathbf{x}\in R^N} ||\mathbf{x}||_{l_1} \text{subject to } \mathbf{y} = D\mathbf{x},$$
$$\text{where } ||\mathbf{x}||_{l_1} = \sum_{i=1}^{N} |x_i| \tag{3.1}$$

In proposed approach, observation $\mathbf{y} \in R^M$ is represented as a linear combination of few number of elements of dictionary D using sparse vector $\mathbf{x}$ i.e., $\mathbf{x}$ has few number of nonzero elements. The main component in the l_1-minimization is the dictionary D. The choice of dictionary is dependent on the application. In proposed approach we construct the dictionaries from the empirical data and use them in finding initial estimate of fusion.

3.1.2 Finding Initial Estimate

We construct various dictionaries from PAN image and use them in CS based multiresolution approach to obtain initial estimate of fused image. Our method fuses high frequency details from panchromatic image into the multispectral bands assuming that they are registered. As a powerful multiresolution analysis tool, we use wavelet transform to obtain high frequency details in the fused image without altering the spectral contents of the multispectral images. Considering a spatial resolution difference of 2 between PAN and MS images, the learning process can be explained as follows.

As shown in Figure 3.1, single level wavelet decomposition of panchromatic image



(a) (b)

Figure 3.1: (a) MS Image of size $M \times M$, (b) P_{wt}: One level wavelet decomposition of PAN image. Here I, II, III and IV quadrants represent the coarse, vertical, diagonal and horizontal details each of size $M \times M$.

results in coarse coefficients in the top-left quadrant I, having the same dimension as that of MS image. Remaining three quadrants (II-IV) show vertical, diagonal and horizontal edge details, respectively. These edge details are used to obtain finer details of fused MS image. We know that PAN and MS images are obtained from the same geographic region with the difference that PAN image is acquired with high spatial resolution covering wide spectral features. This results in high spatial correlation between the MS image and the coarser coefficients of the PAN image. Hence patches of MS image can be well-represented as a sparse linear-combination of coarse elements of the PAN image having high spectral width. Since the finer details fused MS and PAN are similar, we make use of the same sparseness in order to obtain the finer details of fused MS image. Thus our method fuses fine details from PAN image to MS bands by deriving the sparseness of low-resolution observation using the coarser part of PAN image and the same sparsity is used to obtain the detail coefficients of the initial fused estimate. This reduces spectral distortion of the fused image. The proposed algorithm to obtain initial fused image is described below.

Suppose we have the MS image of size $M \times M$, and the corresponding PAN image of size $qM \times qM$, where q is the resolution factor. Here, we describe the algorithm by considering a value of $q = 2$.

1. Obtain one level wavelet decomposition of PAN image P_{wt} as shown in Figure 3.1(b).

2. For each 2×2 patch of MS image choose corresponding patches from all four quadrants of P_{wt}. Arrange them in lexicographical order to form the vectors $\mathbf{v}_{MS}$, $\mathbf{v}_{PA}$, $\mathbf{v}_{PV}$, $\mathbf{v}_{PH}$ and $\mathbf{v}_{PD}$. Here, $\mathbf{v}_{MS}$ corresponds to vector of size 4×1 formed by using MS image, while $\mathbf{v}_{PA}$, $\mathbf{v}_{PV}$, $\mathbf{v}_{PH}$ and $\mathbf{v}_{PD}$ are the corresponding vectors

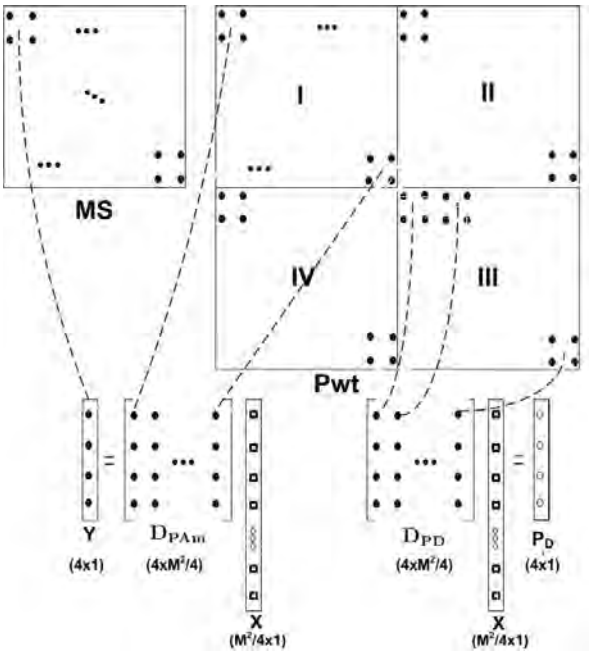


Figure 3.2: Detailed understanding of CS based approach. Construction of dictionaries D_{PA}, D_{PV}, D_{PH}, and D_{PD} from coarser, vertical, horizontal, and diagonal quadrants of P_{wt}, respectively, each of size $4 \times M^2/4$. Obtaining sparse vector $\mathbf{x}$ using coarser dictionary, and use it to obtain diagonal details of fused image, assuming mean subtracted D_{MS} (hence $\mathbf{y}$) and D_{PA} (i.e. D_{MSm} and D_{PAm}).

formed by using the coarse, vertical, horizontal and diagonal quadrants of the PAN image, as shown in Figure 3.2.

3. Create dictionaries D_{MS}, D_{PA}, D_{PV}, D_{PH} and D_{PD} each of size $4 \times M^2/4$ using the corresponding vectors formed in step 2. Construction of dictionaries D_{PA} and D_{PD} are shown in Figure 3.2. Dictionaries D_{PV} and D_{PH} are constructed in a similar way.

4. Find the mean subtracted dictionaries of MS image and the coarse quadrant of the PAN image i.e.,

$$\text{for } j = 1 \text{ to } M^2/4$$
$$D_{MSm}(:, j) = D_{MS}(:, j) - \mathbf{m_{MS}},$$
$$D_{PAm}(:, j) = D_{PA}(:, j) - \mathbf{m_{PA}},$$

where D_{MSm} and D_{PAm} represent the mean subtracted dictionaries. Here, m_{MS} and m_{PA} represent the mean of the MS and PAN image patches.

5. Consider a mean subtracted vector $\mathbf{y}$ of MS test image taken from D_{MSm}.

6. Now solve for sparse vector $\mathbf{x}$ using compressive sensing based l_1-minimization problem found in [12] i.e.,

$$\min_{x \in R^N} ||\mathbf{x}||_{l1} \text{ such that } \mathbf{y} = D_{PAm}\mathbf{x}$$

Here, $\mathbf{x}$ represents the sparseness vector of size $N \times 1$, where $N = M^2/4$. It gives sparse representation of test patch $\mathbf{y}$ in terms of dictionary (D_{PAm}) elements adaptively. Remember $\mathbf{y}$ and D_{PAm} resembles observation vector and sensing matrix, respectively of the compressive sensing theory.

7. Obtain the finer detailed patches of the MS image i.e., the patches of horizontal, vertical and diagonal by using the sparsity vector $\mathbf{x}$, found in step 6.

$$\mathbf{P}_H = D_{PH}\mathbf{x}; \qquad \mathbf{P}_V = D_{PV}\mathbf{x}; \qquad \mathbf{P}_D = D_{PD}\mathbf{x}$$

Convert these into a block of size 2×2. This indicates that the finer details of initial fused MS image are obtained as a linear combination their respective dictionary elements (also referred to as patches) where the sparseness obtained using LR MS image is used. Steps 2-7 are explained graphically in Figure 3.2.

8. Create the fused decomposed patch

$$Fp = [P_M, P_V; P_H, P_D],$$

where P_M is the patch of MS image of size 2×2; P_V, P_D and P_H are the patches corresponding to vertical, diagonal and horizontal edges each of size 2×2.

9. Repeat steps 6-8 for all elements of D_{MSm} and append the MS patches and the obtained finer detailed patches to Fp created in step 8 to their respective locations.

10. Take the inverse wavelet of derived initial estimate Fp of size $2M \times 2M$ to obtain the initial fused image Z.

These steps give us the initial fused image Z having size of $2M \times 2M$.

3.2 Regularization

Since we are constructing the fused MS image by using patch based approach, the spatial homogenity is not taken into account. Hence regularize it further to obtain the final solution. We restrict the solution space for the fused image by using maximum a posteriori - Markov random field (MAP-MRF) approach. The MAP-MRF approach requires data fitting term. This can be done by considering an observation model that represents the MS image formation.

3.2.1 Observation Model

Let Y be the observed MS image of the size $M \times M$ pixels and Z be the fused high-resolution (HR) image, then the forward model for the image formation can be written as,

$$\mathbf{y} = D\mathbf{z} + \mathbf{n}, \tag{3.2}$$

where $\mathbf{y}$ and $\mathbf{z}$ represent the lexicographically ordered vectors of size $M^2 \times 1$ and $q^2M^2 \times 1$, respectively. Here, $\mathbf{n}$ is the independent and identically distributed (i.i.d.) noise vector with zero mean and variance σ_n^2 and has the same size as $\mathbf{y}$. It is given by

$$P(n) = \frac{1}{(2\pi\sigma_n^2)^{\frac{M^2}{2}}} e^{-\frac{1}{2\sigma_n^2}n^2}, \tag{3.3}$$

where D is the downsampling matrix which takes care of aliasing caused due to downsampling. For an integer downsampling factor of q, the matrix D consists of q^2 non-zero elements along each row at appropriate locations [123].

$$D = \frac{1}{q^2} \begin{pmatrix} 1\,1\ldots 1 & & \mathbf{0} \\ & 1\,1\ldots 1 & \\ \ldots & \ldots & \ldots \\ \mathbf{0} & & 1\,1\ldots 1 \end{pmatrix}, \tag{3.4}$$

Now our problem is to estimate fused $\mathbf{z}$ given $\mathbf{y}$, which is an ill-posed inverse problem and can be solved by using regularization methods.

3.2.2 MRF Prior Model

The Markov Random Field (MRF) has emerged as a popular stochastic model for images due to its ability to capture local dependencies. MRF provides a convenient and consistent way of modeling context dependent entities. This is achieved through characterization of mutual influence among such entities. The practical use of MRF models is largely ascribed to the equivalence between the MRF and the Gibbs Random Fields (GRF). In proposed approach we characterize the final fused image using the MRF model. It finds out pixel intensity at current site using neighbourhood pixel intensities of that site and not anything else. This is justified because the changes in intensities in a scene is gradual and hence there is a local dependency.

Let Z be a random field over regular $N \times N$ lattice of sites $L = \{(i,j)|0 \leqslant i,j \leqslant N-1\}$. The equivalence between MRF and GRF is established in Hammersley-Clifford theorem, hence we have

$$P(Z = z) = \frac{1}{Z_z} e^{-U(z)}, \tag{3.5}$$

where z is a realization of Z, Z_z is the partition function given by $Z_z = \Sigma_z e^{-U(z)}$ and U(z) is the energy function given by $U(z) = \sum_{c \in C^z} V_c^z(z)$. Here, $V_c(z)$ denotes the potential function of clique c and C is the set of all cliques. The lexicographically ordered fused image z satisfying Gibbs density function is now written as,

$$P(z) = \frac{1}{Z_z} exp\{-\sum_{c \in C^z} V_c^z(z)\}. \tag{3.6}$$

3.2.3 MAP-MRF Formulation

We use MAP-MRF framework for regularization approach in order to obtain the final fused image. The use of MAP estimation for fusion requires a suitable prior for the same. A method of specifying MRF prior on fused image involves considering cliques c on a neighborhood. By using first order neighborhood, the energy function corresponding to the MRF prior for fused image can be written as,

$$\sum_{c \in \mathcal{C}} V_c(\mathbf{z}) = \gamma \sum_{k=1}^{N_1} \sum_{l=1}^{N_2} [(z_{k,l} - z_{k,l-1})^2 + (z_{k,l} - z_{k-1,l})^2], \tag{3.7}$$

where γ represents the penalty for departure from smoothness in $\mathbf{z}$. $\mathcal{C}$ is the set of all cliques and $V_c(\mathbf{z})$ is the clique potential. Here, the MRF parameter γ is estimated using the initial fused image and this avoids the tuning of the parameter.

The MRF model on the fused image serves as the prior for the MAP estimation. The MAP estimate of the high-resolution fused image comes about by an application of Bayes theorem,

$$P(z|y) = \frac{P(y|z)P(z)}{P(y)}. \tag{3.8}$$

The left hand side is known as the posterior distribution over z and y represents observed data. Here, $P(y)$ may be considered as a normalization constant. We apply this to our problem. Given the LR observation y, the MAP estimate $\hat{z}$, using Bayesian rule, is given by,

$$\hat{z} = \underset{z}{argmax} P(z|y) = \underset{z}{argmax} P(y|z)P(z). \tag{3.9}$$

Taking the log of the posterior probability we can write,

$$\hat{z} = \underset{z}{argmax} \Big[logP(y|z) + logP(z) \Big]. \tag{3.10}$$

Since n is independent. The above MAP formulation allows us to incorporate prior knowledge about z for improving robustness during reconstruction. Using equations 3.2 and 3.3, we obtain

$$P(y|z) = P(n)|_{n=y-Dz} = \frac{1}{(2\pi\sigma_n^2)^{\frac{M^2}{2}}} e^{-\frac{||y-Dz||^2}{2\sigma_n^2}}. \tag{3.11}$$

Thus for MAP-MRF approach, the final cost function to be minimized can be expressed as,

$$\hat{\mathbf{z}} = \underset{z}{^{argmin}} \left[\frac{\| \mathbf{y} - D\mathbf{z} \|^2}{2\sigma_n^2} + \sum_{c \in \mathcal{C}} V_c(\mathbf{z}) \right]. \tag{3.12}$$

Convexity of this cost function allows us to use the simple gradient descent optimization technique, which quickly leads to the minima. Since the optimization process is iterative the choice of initial solution fed to the optimization process determines the speed of convergence. Use of the available fused approximation as an initial solution speed-up the convergence. It may be mentioned here that we obtain initial fused approximation separately for each of the MS observations and the optimization is carried out independently

for every low-resolution MS band.

3.3 Experimental Results

In this section, we present the results of the proposed method for fusion. The experiments are conducted on real images captured using Landsat-7 Enhanced Thematic Mapper plus (ETM+) satellite. The original PAN image and the MS images are of size 512×512 and 256×256 pixels, respectively with ground resolution of each pixel as $30m \times 30m$ and $15m \times 15m$, respectively. These are considered as ground truth. In order to make the quantitative comparison, we downsampled both images by a factor of 2 and 4 and conducted the experiments using the downsampled versions. It is to be noted that the resolution difference between the MS image and PAN image captured by Landsat-7 satellite is of 2. But to evaluate the performance of the proposed algorithm for the higher resolution factor we conducted experiments by downsampling by a factor of 4 also. In both cases, the size of the fused MS image is 256×256. We compare the performance of the proposed method with other methods on the basis of qualitative as well as quantitative measures. Different quantitative measures used in our experiments are described in the following section.

3.3.1 Quantitative Evaluation Measures

For quantifying the results of fusion we used correlation coefficient (CC), structural similarity (SSIM), and mean squared error (MSE) as a evaluation index. These metrics have been widely used in the multiresolution fusion techniques in order to measure the spatial and spectral fidelity of the fused MS images. What follows is a brief review of these measures.

1. Correlation Coefficient (CC) [124]: The correlation coefficient is the most popular measure for checking spatial fidelity between the fused and the groundtruth (original) MS image. It shows the similarity between the fused and the groundtruth MS image band. CC between groundtruth and fused image bands $F, \hat{F} \in R^{M \times N}$ is

defined as

$$CC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{i,j} - \bar{F})(\hat{F}_{i,j} - \bar{\hat{F}})}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{i,j} - \bar{F})^2 \sum_{i=1}^{M}\sum_{j=1}^{N}(\hat{F}_{i,j} - \bar{\hat{F}})^2}}, \tag{3.13}$$

where $F_{i,j}, \hat{F}_{i,j}$ are pixel values at location (i, j) of groundtruth and fused images F and $\hat{F}$, respectively and $\bar{F}, \bar{\hat{F}}$ are the mean values. The CC has a value lying between zero and one, with zero representing the lowest correlation.

2. Structural similarity (SSIM) [125] : The SSIM combines a comparison of luminance, contrast, and structure. It is applied locally to 8×8 square window. This window is moved pixel-by-pixel over the entire image and the SSIM is calculated within the window having range of 0 to 1. Values close to 1 show the highest correspondence of fused image with the groundtruth image. SSIM between groundtruth (X) and fused (Y) image windows is defined as below

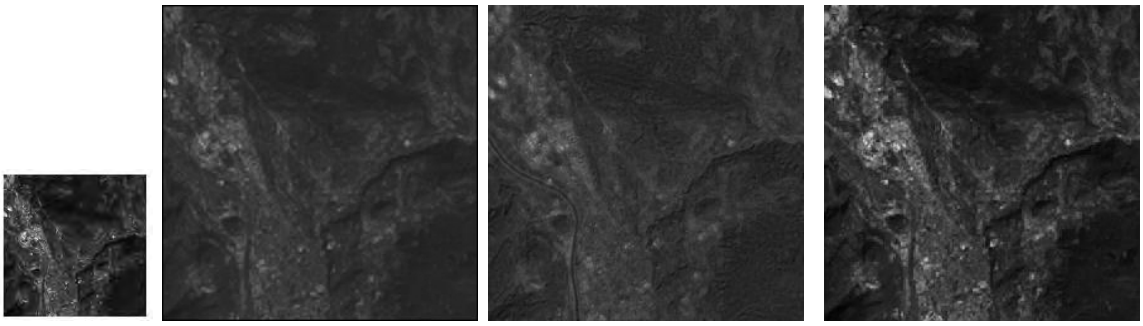
$$SSIM(X, Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{3.14}$$

where μ_x and μ_y represent the mean intensities of X and Y, respectively, while σ_x and σ_y represent the standard deviation of X and Y, respectively. Here, σ_{xy} is the correlation between X and Y. The constant C_1 is included to avoid instability when $\mu_x^2 + \mu_y^2$ is very close to zero. Similarly, the constant C_2 is included to avoid instability when $\sigma_x^2 + \sigma_y^2$ is very close to zero. SSIM is averaged over all windows to obtain average over entire fused image.

3. Mean squared error (MSE): It is a common measure to estimate the squared error between two entities, measuring the ratio of the power within the error to the signal power. It is given by [126]

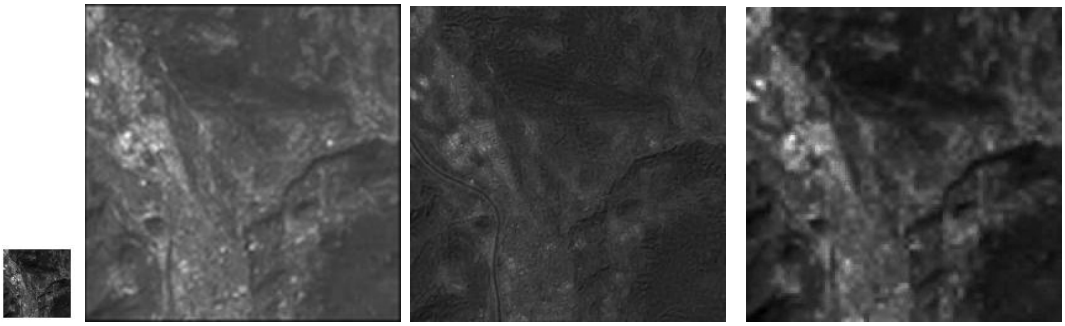
$$\text{MSE} = \frac{\sum_{i,j}[F_{i,j} - \hat{F}_{i,j}]^2}{\sum_{i,j}[F_{i,j}]^2}, \tag{3.15}$$

where $F_{i,j}$ and $\hat{F}_{i,j}$ represent the true HR (groundtruth) and the fused images, respectively.



(a) (b) (c) (d)

Figure 3.3: Fusion results of Band 2 for $q = 2$. (a) MS image of size 128×128, (b) Approach in [127], (c) Approach in [14], and (d) Proposed approach.



(a) (b) (c) (d)

Figure 3.4: Fusion results of Band 2 for $q = 4$. (a) MS image of size 64×64, (b) Approach in [127], (c) Approach in [14], and (d) Proposed approach.

Figure 3.3 show the results of fusion for Band-2 using different approaches. Table 3.1 shows the quantitative comparison using correlation coefficient, structural similarity (SSIM) [125] and mean squared error (MSE) [127] for Band-2. We can see that proposed approach gives better quantitative values as compared to other approaches. We also observe perceptual improvement in the fused images obtained using the proposed method. Method in [127] uses interpolation of MS image and [14] uses patch matching criterion between MS image and coarse level of the PAN image to obtain initial estimate of the fused image, while proposed approach uses sparsity coefficients to represent high frequency details, which leads to better initial estimate compare to other methods. This results in better regularized final fused image because results of regularization depends on initial estimate also.

BAND 2	Approach in [127]	Approach in [14]	Proposed Approach
	Correlation coefficient		
q = 2	0.9462	0.8473	0.9483
q = 4	0.8693	0.8244	0.8879
	SSIM		
q = 2	0.8989	0.8290	0.9396
q = 4	0.8529	0.8132	0.8736
	MSE		
q = 2	0.0068	0.0842	0.0061
q = 4	0.0169	0.0901	0.0108

Table 3.1: Performance comparison of band 2 for $q = 2$ and $q = 4$ in terms of correlation coefficient, SSIM and MSE.

3.4 Conclusion

We have presented a new technique to recover the high spatial and high spectral resolution fused MS image using compressive sensing based learning and MAP-MRF approach. Since the initial estimate is obtained using DWT and compressive sensing theory, the suggested method gives finer details present in different directions with reduced spectral distortion. We considered MRF model to enforce smoothness constraint while regularizing. True MRF parameter can be known only if the fused image is known. Since this is not available, we make use of the close approximation of fusion (initial estimate) to obtain the same. A simple approximation method called maximum pseudo likelihood is used for estimating this parameter [128]. The quantitative results demonstrate that the proposed technique yields better solution as compared to those obtained using the recent approaches.

The proposed approach to increase spatial resolution of multispectral images uses auxiliary panchromatic image in registered form. The high frequency details in fusion result are obtained from high spatial resolution panchromatic image. The proposed approach recovers these high frequency details by exploiting the similarity in sparse representation of coarser resolution of PAN and observed MS image. Note that the accuracy of results is also dependent on registration. Many times auxiliary high-resolution image is not available onboard. Based on this discussion, one may conclude that spatial resolution enhancement without using auxiliary information is difficult. Contrary to this, another

alternative is to use a database of HR and LR training images unassociated with test image to form the dictionaries and recover the finer details from them to enhance the spatial resolution. In next chapter, we discuss the use of LR-HR dictionaries to increase spatial resolution of hyperspectral data where the auxiliary data is seldom available.

Chapter 4

Super-Resolution of Hyperspectral Images using Compressive Sensing Framework

Over the past decade hyperspectral (HS) image analysis has turned into one of the most powerful and growing technologies in the field of remote sensing. Hyperspectral sensors collect information in the form of reflectance spectra in very narrow contiguous bands simultaneously in the visible to mid infrared portion of the spectrum i.e., 0.4-2.5μm. They represent an evolution in technology from multispectral sensors, which typically collect spectral information in only a few discrete, noncontiguous bands [1]. In general HS images have hundreds of spectral bands [129]. Being spectrally overdetermined, they provide ample spectral information to identify and differentiate spectrally unique materials [2]. The high spectral resolution of hyperspectral sensors preserves important properties of the spectrum and makes possible better discrimination of different materials on the ground [1]. Hence they have been proven to be a powerful source for the monitoring of the Earth surface and the atmosphere on global as well as local scales [130]. Nowadays, HS images are widely used in wide range of military and defence applications which include target detection and tracking of objects [131], agriculture planning [132], forest inventory [133], and urban monitoring [134] to mention a few. These applications require high spectral and high spatial resolution data for accurate determination of object properties.

In practice, few satellites have high spatial resolution ($< 5m \times 5m$) sensors available

onboard. For example Quickbird satellite collects panchromatic (PAN) image of spatial resolution $0.7m$ and four multispectral bands of $2.8m$ resolution. Recently launched WorldView-1 and WorldView-2 satellites carry an imaging instrument specifically designed to meet the requirements of very high spatial resolution and more number of spectral bands. WorldView-1 provides a single PAN image of half a meter resolution, while WorldView-2 provides a high spatial resolution ($0.46m$) PAN image and eight spectral bands having resolution of $1.84m$. Although they have high spatial resolution data, they do not provide better discrimination of many surface materials and environmentally relevant information due to their limited spectral resolution. An improved space borne hyperspectral sensors (e.g. Hyperion and Chiris) provides very large number of narrow spectral bands, but their spatial resolution is very less ($17 - 34m$). For airborne hyperspectral sensors (e.g. AVIRIS, HYDICE, HyMap) the spatial resolution is dictated largely by the height of the aircraft [18]. As the height increases the extent of coverage also increases but this decreases the spatial resolution. At low spatial resolutions, the averaging effect on the pixels degrades the performance of spectral detection algorithm, which is used to identify the materials present in the scene [24]. Though the HS images cover large area at fine spectral resolution, their spatial resolutions are often limited for the use in various applications. Hence improving their resolution has a high payoff. This chapter presents a novel approach for super-resolution (SR) of HS images using compressive sensing (CS). Besides ill-posedness of SR problem, the main challenge in HS super-resolution is to preserve spectral contents among all bands while increasing their spatial resolutions.

In this work, given the hyperspectral (HS) images we first obtain an initial estimate of the super-resolution on a reduced dimension HS data. The dimensionality reduction is obtained by using principal component analysis (PCA). Our approach uses CS based method to super-resolve the most informative PCA transformed image representing highest spectral variance (i.e. the first principal component). We make use of low and high spatial resolution dictionaries of patches generated by random sampling of raw patches of PCA transformed images that are generated using the training images of LR and HR having similar statistical properties. Using the sparsity constraint, low-resolution test patch is represented as a sparse linear combination of relevant dictionary elements. Finally, assuming that same sparseness holds for LR and corresponding HR patches an

initial estimate of super-resolved PCA is obtained. Since SR is an ill-posed problem, we obtain the final solution using a regularization framework considering the sparse coefficients obtained by the CS approach and the autoregressive (AR) parameters obtained from the initial estimate. The SR for remaining PCA images is obtained by performing bicubic interpolation and regularization, considering the same AR parameters which were obtained from the initial SR estimate of first PCA component. Application of inverse PCA results in SR of HSI bands in original spatial domain. Experiments are conducted on two different kinds of HS images. Visual inspections and quantitative comparison confirm the effectiveness of the proposed method.

4.1 Previous Work

Super-resolution enhancement refers to an algorithmic approach for increasing the spatial details [135]. Many researchers have attempted to increase the spatial resolution of the HS images by fusing the PAN image and the hyperspectral data [15, 16, 23, 40]. Winter et al. [15] replaced first PCA component of LR hyperspectral image (HSI) with HR PAN image for resolution enhancement, but performance of this method decreases when correlation between the spectral response of the PAN and HSI decreases. First PCA component represents the intensity component, hence the resulting HR HSI has only the intensity variations at finer resolution. The main limitation of all this method is that the spectral characteristics are not preserved. In a different approach Bar et al. in [41] combined spectral and spatial analysis for detection and classification, respectively. In the detection stage they used high spectral resolution HSI to locate the target and in the classification stage high spatial resolution PAN image is fused with low spatial resolution HSI to reduce the false alarms. The limitation of all these algorithms is that they require the images to be registered. Besides this, many times high-resolution PAN imager is not available onboard, hence auxiliary high-resolution image of the same geographic area is not captured. In such circumstances our proposed method provides the solution of resolution enhancement. Recently, compressive sensing (CS) theory has drawn major attention in various image processing applications [80, 136, 81, 82, 137, 138]. Compressive sensing theory has many potential applications in signal and image processing applications. It is primarily concerned with the recovery of a vector $\mathbf{x}$ that is sparse in some transform

domain. In this work, we present a novel approach for super-resolution of HS images from the perspective of CS which is not dependent on the registration of the different HS images. Our method needs only HR and LR registered HS images of any compatible scenery to create training dictionaries which is one time offline procedure.

4.2 Theoretical Background

In this work, we are using compressive sensing (CS) and principal component analysis (PCA). Since compressive sensing has been discussed in chapter 3 section 3.1.1 we only give the necessary details of CS and briefly discuss PCA here.

A fundamental ingredient to deploy CS theory in applications is the dictionary D in equation 3.1 mentioned in section 3.1. There are three different ways to construct the dictionaries [122]: (1) Preconstructed dictionaries, like wavelets [139], contourlets [140] etc. They are generally used for "cartoon-like" images, assumed to be piecewise smooth having smooth boundaries [141, 142]. (2) Tunable dictionaries, in which a basis or frame is generated under the control of particular parameter (discrete or continuous): wavelet packets [143] (parameter is time-frequency subdivision) or bandelettes [144] (parameter is spatial position). (3) A training database of signal instances similar to those anticipated in the application, and build an empirically learned dictionary [80]. Here the entries in dictionary are chosen from the empirical data rather than from some theoretical model. In proposed approach a training database of signal instances similar to those anticipated in the application is used to build an empirically learned dictionary. Thus the entries in dictionaries are chosen from the empirical data rather than from some theoretical model. Validity of such dictionaries is examined by [80]. Such dictionary has the potential to outperform commonly used predetermined dictionaries [145]. Such a dictionary can then be used in the application as a fixed and redundant dictionary. In our application we explore the third option.

4.2.1 Principal Component Analysis

Hyperspectral images are composed of large number of spectral bands (e.g., AVIRIS acquires 224 bands). Hence, applying super-resolution technique to each band separately is prohibitive because of time complexity. In addition individual band SR does not make

use of the information present across the bands [33]. Information is present across these bands in the form of spectral signatures and the identification of ground materials of interest is based on their unique spectral signatures [1]. Besides the ill-posed nature of SR problem, the HS image SR task becomes difficult since preservation of spectral correlation combined with the SR is more challenging. The spectral content of HS images are inherently low dimensional, hence this must be exploited. Principal component analysis (PCA) plays a central role in the analysis of multivariate data [8].

Suppose we have a dataset of B hyperspectral bands with size of each band as $M \times M$ pixels. Assume that $\mathbf{F}$ represents the pixel vector of size $B \times 1$ along spectral dimension of HS image. A set of principal components of dimension $M^2 \times K$ are computed from the first K eigen vectors E=$[\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_K}]^T_{B\times K}$, which in turn is computed from the covariance matrix Σ of size $B \times B$ from the given data set

$$\Sigma = \frac{1}{M^2 - 1}\sum_{i=1}^{M^2}(\mathbf{Fi} - \mathbf{m_F})(\mathbf{Fi} - \mathbf{m_F})^T, \tag{4.1}$$

$$\text{where } \mathbf{m_F} = \frac{1}{M^2}\sum_{i=1}^{M^2}\mathbf{Fi}.$$

Here $\mathbf{m_F}$ is the vector representing average band intensity. Here the top $K \ll B$ eigen vectors corresponding to maximum eigenvalues contribute to maximum information of HS observations and the remaining $B - K$ eigen vectors cover very less information. By projecting HS observations on these eigen vectors we obtain K number of primary PCA components and $B-K$ number of secondary PCA components each of dimension $M \times M$. The most obvious feature of the principal components is that the maximum spectral variability of hyperspectral bands is contained in the first few principal components. Each component has variability in decreasing order of magnitude. Percentage of variability included in i^{th} principal component is given by,

$$Variability = \frac{\lambda_i}{\sum_{i=1}^{B}\lambda_i}, \tag{4.2}$$

where λ_i is eigen value of the i^{th} eigen vector. Since K is less than B, one cannot reconstruct hyperspectral image exactly. However, the hyperspectral bands are highly correlated, and hence only a small value of K suffices to reconstruct HS image to obtain

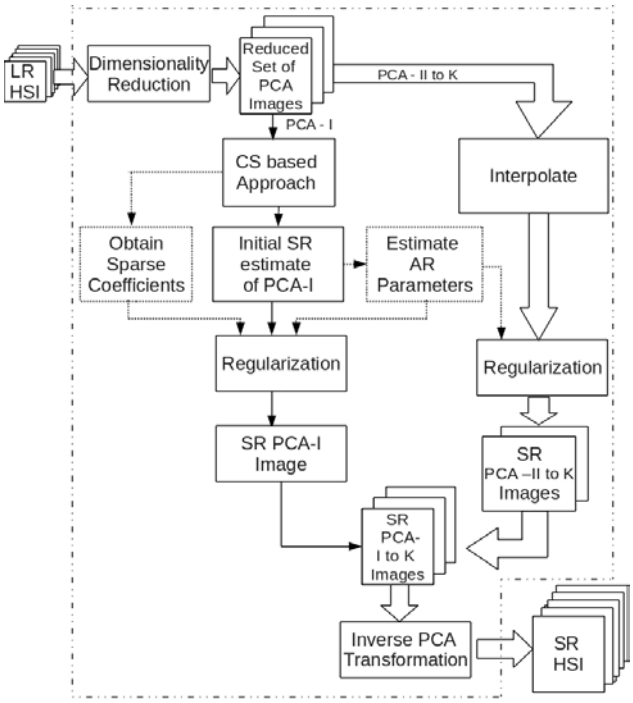


Figure 4.1: Block diagram of HS image super-resolution.

the required details. Our algorithm works on K number of primary PCA components to obtain super-resolution of HSI cube.

4.3 Block Diagram Description of the Proposed Method

Block diagram of the proposed approach is shown in Figure 4.1. In our approach, we first represent the HS observations from different wavelengths, as a weighted linear combination of small number of basis image planes (BIPs) using PCA transform described in section 4.2.1. The super-resolution is applied on the reduced set of PCA transformed images. To reduce computational burden of l_1-minimization, we first apply the CS based approach only to first PCA component to obtain initial SR estimate of PCA-I. Regularization based on AR parameters and sparsity priors is performed on this image to obtain the final SR PCA-I image. Remaining images corresponding to K principal components are interpolated and regularized using AR parameters obtained from the same initial SR image. In our work we have used $K = 3$. Here our assumption is that the AR parameters learned from the initial SR PCA-I are valid for all K PCA images, as we are learning the spatial dependence and not their explicit values. Applying inverse PCA on this set of SR PCA images yields super-resolved HS image.

4.3.1 Close Approximation to Super-resolution using CS

Here, a dictionary based approach is used on the first LR principal component to learn the HR details. Training database of the same class is used to create a set of the HR and LR patches, so that BIPs of PCA better represent the materials of interest. Thus the choice of BIPs depends on the application. Note that dictionaries are not created directly from all bands of HS image, instead first PCA transformed image is used.

Suppose LR HS image of size $M \times M$ need to be super-resolved to a size of $qM \times qM$, where q is the super-resolution factor. Due to similar statistical properties, HS test image patches can be represented as a sparse linear combination of LR dictionary elements using equation (3.1). The same sparsity holds good for the corresponding HR unknown image [146]. Hence we can recover the HR image using the HR dictionary. The proposed algorithm to obtain initial estimation of SR PCA-I image is described below.

1. Generate mean subtracted LR and HR training HS images of B bands with $L_m = [\mathbf{L}_1, \mathbf{L}_2, .., \mathbf{L}_B]$ and $H_m = [\mathbf{H}_1, \mathbf{H}_2, .., \mathbf{H}_B]$, respectively. Here all bands of LR and HR HS images are arranged lexicographically to convert them in vectors of size $M^2 \times 1$ and $q^2M^2 \times 1$, respectively.

2. Determine basis eigen vectors corresponding to LR and HR training HS images. Retain basis eigen vectors $\mathbf{e_h}$ and $\mathbf{e_l}$ corresponding to maximum spectral variability of data (i.e., highest variance), each of size $1 \times B$.

3. Create HR and LR transformed images B_H and B_L by projecting H_m and L_m on their corresponding basis eigen vectors generated in step (2). $\mathbf{B}_H = \mathbf{e_h} \times H_m^T$; $\mathbf{B}_L = \mathbf{e_l} \times L_m^T$; Convert $\mathbf{B}_H$ and $\mathbf{B}_L$ into matrices of size $qM \times qM$ and $M \times M$ respectively.

4. For each $b \times b$ patch of B_L choose the corresponding patch from B_H. Arrange them in lexicographic order to form the vectors $\mathbf{V}_l$ and $\mathbf{V}_h$ of size $b^2 \times 1$ and $q^2b^2 \times 1$, respectively.

5. Repeat step (4) to create dictionaries D_L and D_H of size $b^2 \times M^2/b^2$ and $q^2b^2 \times M^2/b^2$, respectively.

6. Project the mean subtracted LR hyperspectral test image on basis eigen vector $\mathbf{e}_l$ and obtain the transformed LR test image Y of size $M \times M$.

7. Consider a patch of size $b \times b$ from the transformed LR test image Y, convert it in lexicographic order to obtain a vector $\mathbf{y_p}$.

8. Solve CS based l_1-minimization optimization problem
   $min||\mathbf{x_p}||_1$ such that $\mathbf{y_p} = D_L\ \mathbf{x_p} + e$.
   Here $\mathbf{x_p}$ gives sparse representation of test patch $\mathbf{y_p}$ in terms of LR dictionary (D_L) elements adaptively.

9. Obtain SR patch using $\hat{\mathbf{z}}_\mathbf{p} = D_H\ \mathbf{x_p}$. This is SR patch in transform domain.

10. Repeat steps (7) to (9) for all patches of Y to obtain transform domain SR PCA-I image $\hat{Z}$. Note that the spatial dependency still exists within the pixels of this image. However, different PCA components of HS image are uncorrelated. Hence there is no spatial dependency of pixels among different PCA components.

The above steps gives us the initial estimate of the SR PCA-I image Z having a size of $qM \times qM$. In our experiments, we considered patch size of $b = 2$ for $q = 2$. We used a dictionary of 100000 raw patches. Here raw patches belonging to the same class were used in our experiment.

4.3.2 Final Solution using Regularization

Since we are estimating the initial SR image using patch based approach, the spatial correlation is not considered whenever there are discontinuities at the patch boundaries. Hence we need to regularize it further to obtain a better solution by restricting the solution space by using AR and sparsity priors. For regularization purpose one needs to have data fitting term and regularization term. Hence we first model the image formation in order to get the data fitting term.

Let Y be the PCA transformed LR HS image of size $M \times M$ and and Z be the corresponding HR HS image of size $qM \times qM$, then the model of image formation is represented as:

$$\mathbf{y} = D\mathbf{z} + \mathbf{n}, \tag{4.3}$$

where $\mathbf{y}$ and $\mathbf{z}$ represent the lexicographically ordered vectors of size $M^2 \times 1$ and $q^2M^2 \times 1$, respectively with $\mathbf{z}$ representing the SR vector. D is the downsampling matrix taking care of aliasing caused as a result of downsampling. For an integer downsampling factor of q, matrix D consists of q^2 non-zero elements along each row at appropriate locations. Here $\mathbf{n}$ is the independent and identically distributed (i.i.d.) noise vector with zero mean and variance σ_n^2 and has same size as that of $\mathbf{y}$. In this model the LR intensity is the average of the HR intensities over a neighborhood of q^2 pixels corrupted with additive noise.

4.3.3 Estimation of Autoregressive Parameters

We consider learned super-resolved image as the initial estimate and regularize it further to obtain the final solution. We characterize the statistical dependence of pixel values on its neighbors, by using an AR model, where the pixel value at a location is expressed as a linear combination of its neighborhood pixel values and an additive noise [13]. We use initially estimated SR PCA-I image using CS approach as an AR model. Note that, PCA components are mutually uncorrelated across bands but there exist spatial dependency among the pixels in each of the PCA components. We estimate the AR parameters from the PCA-I component and use them for other bands. We use a homogeneous AR model and derive a set of parameters for entire SR image. Suppose $z(s)$ is the gray level value of the image pixel at location $s = (i, j)$ in an $N \times N$ image, where $i = 1, 2, ..., N$ and $j = 1, 2, ..., N$. The AR model for $z(i, j)$ can be expressed as [147]

$$z(i,j) = \sum_{r \in Ns} B_r z(s+r) + \sqrt{\rho} n(s) \tag{4.4}$$

where N_s is the neighborhood of pixel s, r being a neighborhood index with $r \in N_s$, and ρ are unknown parameters, $n(.)$ is an independent and identically distributed noise sequence with zero mean and unit variance; ρ is the variance of the white noise that generates the spatial data for the given AR parameters. Here we use fifth order neighborhood as a compromise between local and global texture representation, that requires to estimate a total of eight AR model parameters B_r using the iteration scheme given in [147]. We are considering same neighborhood size around each pixel. The extracted AR parameters are also used to regularize other PCA images.

4.3.4 Regularization with Sparsity Coefficients and AR Parameters

Super-resolution is an ill-posed inverse problem. Prior information can enhance the quality of the solution considerably. Hence we obtain the final solution using regularization framework. Sparse coefficients obtained during the CS framework and AR parameters obtained from the initial super-resolved image are considered as prior informations. The prior knowledge about sparsity coefficients is used in determining weightage of dictionary atoms to represent HR patches. It provides the constraint of sparsity in final solution. The AR parameters plays the role of maintaining the contextual constraint used to regularize the solution. Using a data-fitting term, sparsity and AR prior terms, the final cost function is written as

$$\epsilon = ||\mathbf{y} - D\mathbf{z}||^2 + \beta\| \mathbf{z} - D_H\mathbf{x} \|^2 + \sum_i \sum_j \left((z(i,j) - \sum_{l,k \in Ns} B_r z(i+l, j+k) \right)^2 \quad (4.5)$$

Here β represents the regularization parameter. The above cost function is convex. Hence it can be minimized by using a simple optimization technique such as gradient descent. In order to provide a good initial guess and to speedup the convergence, the result obtained by CS based technique is used as the initial estimate for $\mathbf{z}$.

4.4 Experiments and Results Analysis

In order to evaluate the performance of the proposed SR technique, experiments are conducted on two different hyperspectral image data sets. The first data set is comprised of 31 band reflectance image of natural scene, corresponding to wavelengths between $0.4\mu m$ and $0.7\mu m$ in steps of 10 nm all acquired under the direct sunlight in clear or almost clear sky [149]. Here we used "Scene 5" of Hyperspectral images of natural scenes 2002[1]. The second data set is comprised of 224-band real hyperspectral image of Moffett Field acquired by AVIRIS hyperspectral imaging system[2]. First dataset is

[1]http:// personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_02.html

[2]Aviris Free Data, Jet Propulsion Lab., California Inst.Tech., Pasadena, http://aviris.jpl.nasa.gov/html/aviris.freedata.html

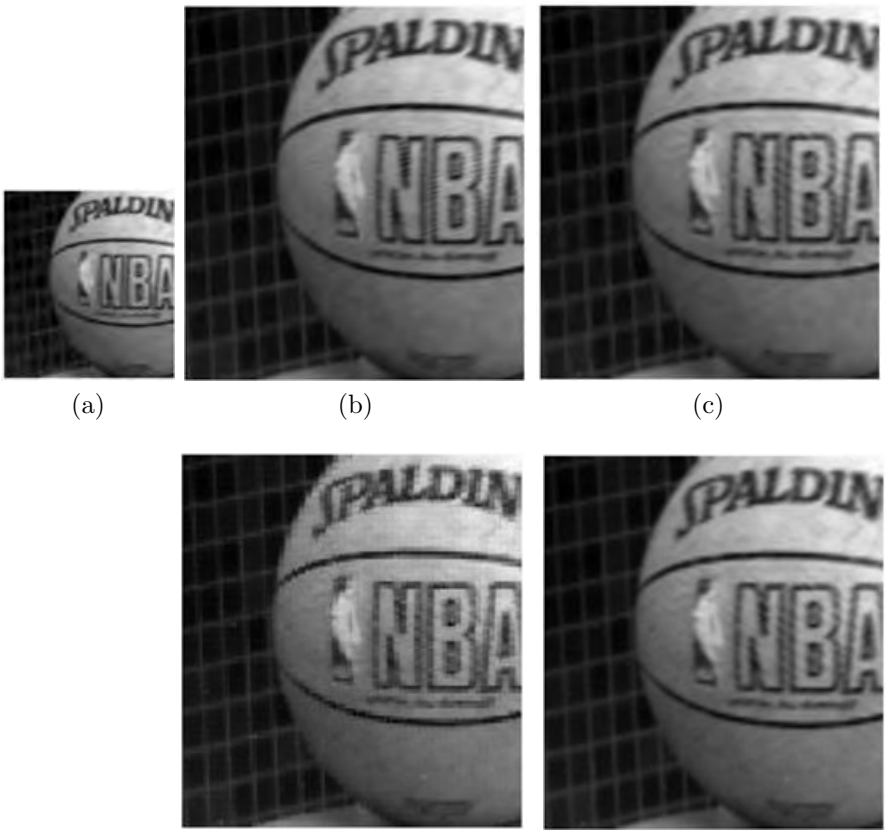


Figure 4.2: SR results on PCA-I of natural HS image for $q = 2$. (a) LR image of size 100×100, (b) Groundtruth of size 200×200, (c) Bicubic interpolation [148], (d) CS based initial estimate, and (e) Proposed approach.

used for checking effectiveness of different steps of algorithm by visual comparison and second dataset is used to test the performance using visual comparison as well as different quantitative measures.

4.4.1 Quantitative Evaluation Measures

Detailed quantitative evaluation of spatial and spectral fidelity of super-resolved AVIRIS hyperspectral image is performed using different measures such as correlation coefficient (CC), spectral angle mapper (SAM), and erreur relative globale adimensionnelle de synthese (ERGAS). Note that these metrics are used by the multiresolution fusion researchers in order to measure the spatial and spectral fidelity of the fused MS images. What follows is a brief review of these measures.

1. Correlation Coefficient (CC) [124]: The correlation coefficient is the most popular

measure for checking spatial fidelity between the SR and the original HSI. It shows the similarity between the super-resolved and the groundtruth HSIs for each band. CC_k between two k^{th} image bands $F, \hat{F} \in R^{M \times N}$ is defined as

$$CC_k = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{i,j} - \bar{F})(\hat{F}_{i,j} - \bar{\hat{F}})}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{i,j} - \bar{F})^2 \sum_{i=1}^{M}\sum_{j=1}^{N}(\hat{F}_{i,j} - \bar{\hat{F}})^2}}, \tag{4.6}$$

where $F_{i,j}, \hat{F}_{i,j}$ are pixel values at location (i, j) of groundtruth and super-resolved images F and $\hat{F}$, respectively and $\bar{F}, \bar{\hat{F}}$ are the mean values. The CC has a value lying between zero and one, with zero representing the lowest correlation. Since we have a number of spectral images, we average the correlation coefficient values computed over all the bands to obtain the CC_{avg} i.e.,

$$CC_{avg} = \frac{1}{B}\sum_{k=1}^{B} CC_k, \tag{4.7}$$

where B is the total number of bands i.e. 141 in this experiment.

2. Spectral Angle Mapper (SAM): Since this measure is insensitive to variable gain resulting from the topographic illumination effects [150], we have chosen it to measure spectral fidelity of super-resolved images. It is defined as the angle between two vectors. Low value of SAM indicates less spectral distortions. Spectral angle between ground truth and super-resolved image is defined as:

$$SAM(\mathbf{v}, \hat{\mathbf{v}}) = arccos\left(\frac{\mathbf{v}, \hat{\mathbf{v}}}{||\mathbf{v}||_2 . ||\hat{\mathbf{v}}||_2}\right), \tag{4.8}$$

where $\mathbf{v}$ and $\hat{\mathbf{v}}$ are the pixel vectors of groundtruth and super-resolved image, respectively.

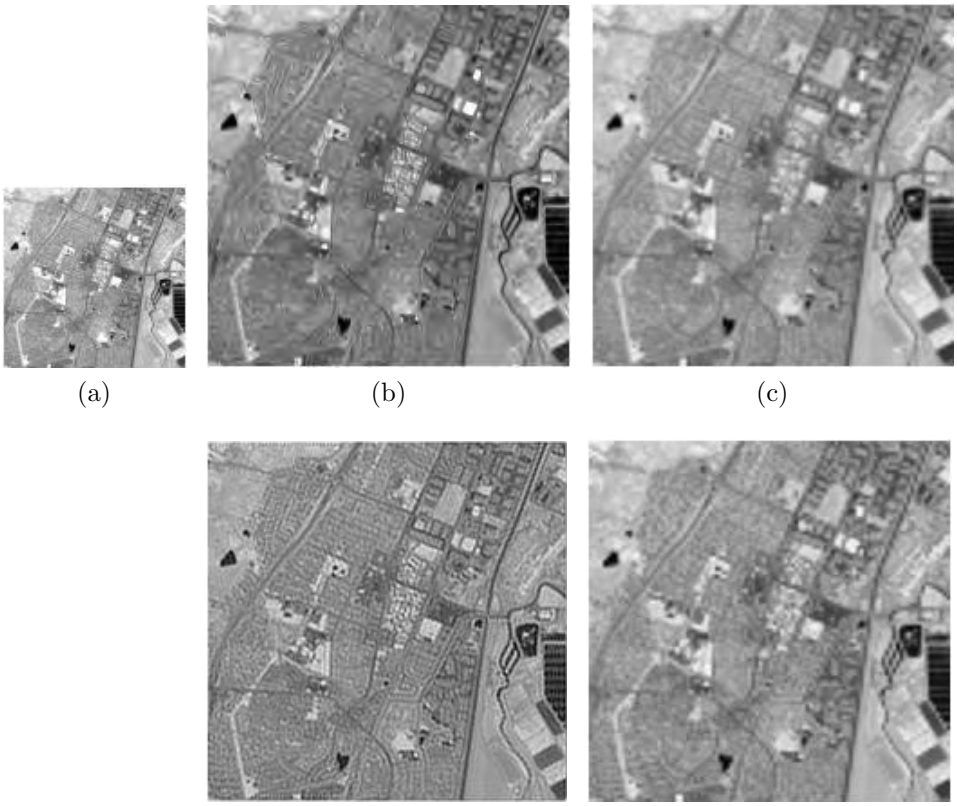
For example for SR HSI with spatial size of 200×200 and spectral size of 141 bands we obtain $200 \times 200 = 40000$ values for SAM indicating spectral fidelity between groundtruth pixel and SR pixel along spectral dimension. These values are averaged over the entire image band to get a global measure of spectral distortion of the super-resolved image. Ideally this value has to be zero.

3. Relative Dimensionless Global Error in Synthesis (ERGAS): This is an indicator of the overall error in super-resolved HSI. The ERGAS value is defined as [151]

$$ERGAS = 100\frac{h}{l}\sqrt{\frac{1}{B}\sum_{k=1}^{B}\left(\frac{RMSE(k)}{\mu(k)}\right)^2}, \tag{4.9}$$

where h/l is the ratio of number of pixels in HR and LR images i.e., r in our case. $RMSE(k)$ and $\mu(k)$ are the root mean squared error and mean of the k^{th} band, respectively. This value has to be small for better performance.

Both datasets have high spatial dimensions, hence specific regions are cropped from them and experiments are carried out on these regions. For the purpose of quantifying the results we consider original HS images as groundtruth and generated synthetic LR HS images by applying downsampling by a factor of $q = 2$, in both the spatial directions of HS data. We then applied SR algorithm to the synthetically generated LR HS images and compared the results against the original HS image. We limit AVIRIS HSI data in the wavelength range of $0.4\mu m$ and $1.79\mu m$ to conduct the simulation to reduce the time complexity, as this will not make any significant difference in the results of our proposed algorithm. After removing few bands having low signal to noise ratio (SNR), 141 bands from the original HS image of AVIRIS were super-resolved by a factor of 2. The band removal was based on visual inspection of the images. Effectiveness of the proposed algorithm on the natural scene is presented in Figure 4.2. Note that, here results are presented only on the PCA-I image instead of HS bands. Figures 4.2(a) and (b) display the LR test image and the ground truth image, respectively. The result obtained using bicubic interpolation is shown in Figure 4.2(c). To demonstrate the effectiveness of proposed approach, here also we show the initial SR estimate obtained using CS based approach in Figure 4.2(d) and final super-resolved PCA-I image in Figure 4.2(e). Figure 4.2(c) indicates that the borders of text and crossed lines appear blurred in the bicubic interpolated image. Visual inspection of Figure 4.2 (d) shows that initial estimate of SR PCA-I using CS based approach gives a quality, comparable to the groundtruth. Since the neighborhood relations are not considered in patches while obtaining the initial estimate, we observe shading of letters written on a ball and also observe little blockiness. Applying regularization to initially estimated SR PCA-I helps to reduce this effect considerably as



(a) (b) (c)

(d) (e)

Figure 4.3: SR results on AVIRIS HS Band 100 for $q = 2$. (a) LR image of size 100×100, (b) Groundtruth of size 200×200, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], and (e) Proposed approach.

shown in Figure 4.2(e), which is closer to the groundtruth. This shows effectiveness of regularization in proposed algorithm.

In Figure 4.3 we display the SR results on band 100 of AVIRIS data. The LR test image and the groundtruth image are displayed in Figures 4.3(a) and (b) respectively. High resolution image obtained using bicubic interpolation [148] is shown in Figure 4.3(c). The results obtained using iterative backprojection [152] and proposed approach are shown in Figures 4.3(d) and (e), respectively. From Figure 4.3(c) we can see that high-resolution image obtained using bicubic interpolation is blurred and the high frequency spatial details are not preserved. We can see in Figure 4.3(d) that pure white patches in LR image is converted to grayish patches in SR image obtained using iterative backprojection method. One can see that the SR image obtained using the proposed method displayed in Figure 4.3(e) compares well with the groundtruth. The proposed method provides better visual quality compared to other approaches.

In order to compare the results on quantitative basis we use score indices such as correlation coefficient (CC) [124], relative dimensionless global error in synthesis (ERGAS) [151], and spectral angle mapper (SAM) [150]. These are generally used in the multiresolution fusion techniques in order to measure the spatial and spectral fidelity of the fused multispectral images. Since SAM measure is insensitive to variable gain that results from the topographic illumination effects [150] we choose it for measuring the spectral fidelity. Table 4.1 shows quantitative comparison among bicubic interpolation, IBP and proposed approach. Results are listed for different amount of spectral variability (I) retained after transformation. CC is averaged over all bands of HS image to obtain a global measurement of spatial distortion and SAM is averaged over all pixels to yield a global measurement of spectral distortion. As seen from the Table 4.1, our method provides scores that are more closer to reference values compared to bicubic interpolation and IBP. Lower value of ERGAS in the proposed method indicates less global distortion in super-resolved HS image. CC for all the HS bands plotted in Figure 4.4 shows that proposed method gives better spatial fidelity compared to bicubic interpolation and iterative backprojection. The surface plots in Figures 4.5(a), (b), and (c) represent spectral fidelity of each specific pixel in super-resolved HS images obtained using bicubic interpolation, iterative backprojection, and proposed approach, respectively for $q = 2$. Bicubic interpolation and iterative backprojection method give a maximum of 144.89 degree and 120.10 degree of SAM, while proposed method gives a maximum of 116.22 degree SAM as seen from surface plots of Figures 4.5(a), (b), and (c) respectively. Lower values of maximum as well as average SAM indicate that proposed method provides better spectral fidelity.

Method	q	I	CC	SAM	ERGAS
Bicubic [148]	2	98	0.886	5.117	7.446
IBP [152]	2	98	0.865	6.312	7.214
Proposed	2	98	0.902	4.863	7.065
Bicubic [148]	2	80	0.838	9.682	10.464
IBP [152]	2	80	0.830	10.153	10.125
Proposed	2	80	0.847	9.397	10.147
Reference			1.000	0.000	0.000

Table 4.1: Quantitative comparison of SR results on AVIRIS data for $q = 2$

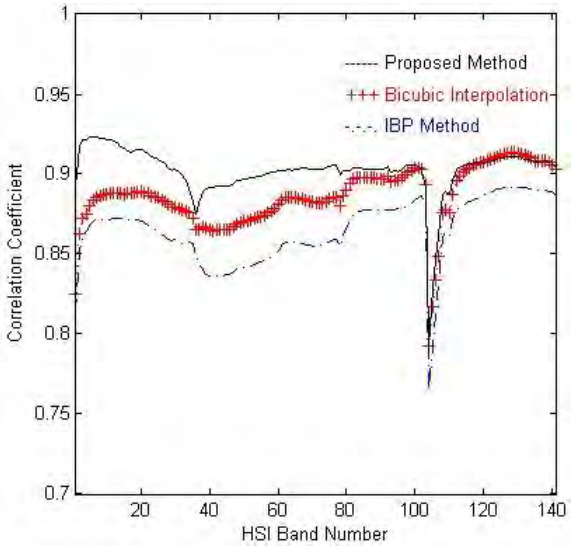
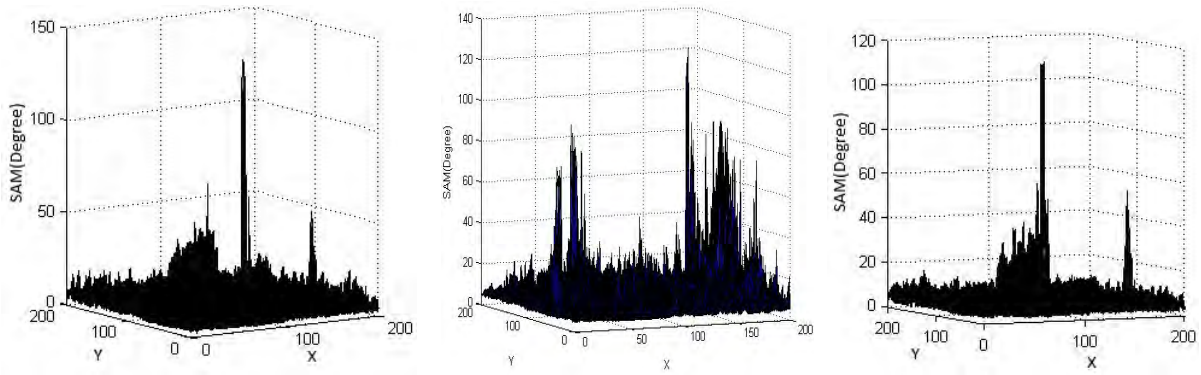


Figure 4.4: Plot showing detailed performance of correlation coefficients for $q = 2$ of AVIRIS HS bands 1-141



(a) (b) (c)

Figure 4.5: Surface plots showing SAM of all pixels for $q = 2$ on AVIRIS data. (a) Bicubic interpolation [148], (b) Iterative backprojection [152], and (c) Proposed approach.

The proposed method achieves higher spatial correlations indicated by the CC and ERGAS, which are more closer to reference value. In IBP the choice of back-projection filter is arbitrary and incorporation of prior information is difficult, which results in higher spatial as well as spectral distortion in SR image compared to proposed approach. In the proposed approach the use of regularization helps us to achieve better spectral fidelity in terms of lesser value of SAM. As we increase the variability retained in PCA components, spatial and spectral distortions are reduced considerably. Here we can extend the CS based approach to all significant PCA components in order to improve the performance.

4.5 Conclusion

We have presented a novel approach to recover the high spatial resolution and high spectral resolution HS image using CS based learning and global AR prior model. The advantages of the proposed technique are: 1) no need of supplementary spatial information in registered form, 2) has high spatial fidelity and low spectral distortions, and 3) once the low-resolution and high-resolution dictionaries are created from training dataset, the HS images captured by a low-resolution sensor can be super-resolved. Quantitative comparison of score indices show that our method enhances spatial information without introducing significant spectral distortion.

It is necessary to point out that proposed approach is using CS based learning on first PCA component only. Use of raw dictionary for CS based approach needs large number of patches, which in turn increases computation time while obtaining initial SR estimate. Another drawback is the use of an observation model which assumes that the LR pixel intensity is the average of the corresponding HR pixels intensities i.e., we assumed an averaging as degradation for all bands of hyperspectral image. In practice, many factors like diffraction, shape, location, physical construction and electronic response of the detectors contribute to the degradation (i.e., PSF) of any hyperspectral imager. Hence the degradation considered in proposed approach is not optimum for all spectral bands. It is typical for the degradation to degrade as distance from the center of the FOV is increased. Hence a better way is to estimate degradation that can optimally represent image formation for all spatial locations and spectral bands. But the spatially varying PSF requires estimation of PSF at each observed pixel which is quite involved. The work presented in the next chapter involves extension of CS based approach to multiple PCA components instead of single one. Instead of raw dictionaries jointly trained dictionaries that have few number of atoms are used. This reduces timings in initial estimates of SR. Besides this, it estimates PSF for each spectral component separately to represent optimum image observation model for each spectral band to obtain final SR results.

Chapter 5

Use of Learned Dictionaries and Gabor Prior

In the work of previous chapter, significant PCA components are used to obtain super-resolution of hyperspectral images. Due to large number of atoms in raw dictionaries of LR-HR, the required computational time is very high in order to obtain initial SR estimation. Hence in chapter 4 we restricted our CS based approach to obtain initial estimate on first PCA component only. Besides this, we assumed degradation as an averaging effect which is not true in practice. This chapter presents a novel approach to increase the spatial resolution of HS images using the compressive sensing (CS) and a new prior called "Gabor prior". The novelty of the proposed approach lies in the use of : (i) jointly learned CS dictionaries, (ii) estimated degradation matrix, and (iii) a new prior called "Gabor prior" in order to super-resolve the significant PCA transformed images. Given the hyperspectral images, we first represent the HS observations as linear combination of small number of basis image planes (BIPs) using principal component analysis (PCA) and the data of reduced dimension is used in our work. In order to obtain SR image for for each HS band we first obtain the initial estimates of the super-resolution on this reduced dimension. Since SR is an ill-posed problem, the final solution in PCA domain (SR PCA components) is obtained by using a regularization framework. Similar to the previous work, applying inverse PCA to significant SR PCA components results in super-resolved hyperspectral bands in spatial domain. Experiments are conducted on two different HS data sets namely 31-band natural hyperspectral image (HSI) collected under controlled laboratory environments and 224-band real HS images collected by remote sensing sen-

sor Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Visual inspections and quantitative comparison confirm that our method enhances spatial information without introducing significant spectral distortion.

5.1 Previous Work

Super-resolution enhancement refers to an algorithmic approach to overcome the inherent spatial resolution limitation of imaging systems [153]. Tsai and Huang [20] were first to propose the SR technique in frequency domain. They reconstructed HR image from a set of aliased LR images. In the last three decades many researchers have attempted to increase the spatial resolution of the HS images using auxiliary information in the form of (i) multiple LR observations [21, 33, 154, 34, 5], (ii) HR multispectral images [22, 37], and (iii) HR panchromatic image [15, 53, 16, 23, 40, 28, 41, 54]. These algorithms are generally referred to as fusion algorithms.

Wilson et al. [21] used combination of multiple LR HSIs in order to obtain a subset of HR images while maintaining the visual information necessary for human analysis. They assume that LR HSIs are registered and are acquired from a single sensor or multiple sensors. To reduce the computations, [33] modeled HS image acquisition process as weighted linear combinations of a small number of basis image planes. A set-theoretic method is used to combine the information from multiple LR HSIs to obtain HR HSI. Another method proposed by [34] used LR multiangular HSIs that are registered using thin plate spline nonrigid transform to reconstruct HR HSI using Delaunay triangulation-based nonuniform interpolation method. Zhang et al. [5] proposed a maximum a posteriori (MAP) based multi-frame SR algorithm utilizing principal component analysis in order to reduce the complexity. Gomez et al. [22] proposed the fusion between HS and MS images using the wavelet based method. The method reported in [37] employs generalized MAP approach that makes use of stochastic mixing model in order to obtain high-resolution HSI. Explicit spectral relationship between MS and HS image is not required in this method.

Many algorithms use HR panchromatic image to fuse details into the LR HS image to enhance its the spatial resolution. Winter and Winter [15] replaces first PCA component of LR HSI with HR PAN image for resolution enhancement, but performance of this

method decreases when correlation between the spectral response of the PAN and HSI decreases. First PCA component represents the intensity component, hence the resulting HR HSI has only the intensity variations at finer resolution. The main limitation of this method is that the spectral characteristics are not preserved. A different approach presented by Hardie et al. [16] used MAP estimator for enhancing the spatial details using co-registered PAN image to obtain enhanced HSI. The method allows for any number of spectral bands in primary and auxiliary image. On a similar line Eismann and Hardie [23] used MAP estimation framework combined with a stochastic mixing model (SMM) for reconstructing subpixel spatial information. Here SMM is used to provide the constraint in estimation of HR HSI. Capobianco et al. [40] fused PAN data with Hyperion HSI using two different linear injection models, namely single spatial detail (SSD) and the band-dependent spatial detail (BDSD) models. In the SSD model same PAN image is used to enhance all the bands of HSI, while in BDSD model an optimum detail image is extracted from the PAN data and the same is fused with the HSI, thus providing more accurate results than SSD model. Garzelli et al. [28] proposed constraint spectral angle (CSA) fusion to preserve spectral properties in HR HSI while increasing its spatial resolution. A method proposed by Bar et al. in [41] extracts anomalies from LR HSI captured using Compact army Spectral Sensor (COMPASS) and a subregion from the HR PAN image is extracted to match each anomaly resulting in HR HSI. Few researchers have proposed methods that use fusion as preparation stage and uses them in linear mixture model to improve the performance of their algorithm [54, 53]. All these methods based on fusion of HR MS or PAN data require accurate coregistration of LR HSI and HR image acquired over the same area.

Availability of auxiliary information can be very expensive or sometimes impossible, hence indirect approaches based on spectral mixture analysis and learning [24, 26, 60, 62], and compressed sensing [61, 79] were proposed by the researchers. A method proposed by Gu et al. [24] first obtains abundance map using linear spectral mixture analysis (LSMA). Then based on spatial correlation of landcovers, learning based SR mapping is performed by using back propagation neural network (BPNN) to enhance the spatial resolution of HSI. Villa et al. In a different approach Mianji et al. [26] used LR test image and its downsampled version to train the BPNN. They perform learning based SRM after SMA by considering spatial correlation of different materials present in the HS image.

Villa et al. [62] proposed the algorithm in which first the spectral unmixing is performed to determine proportion of endmembers in each pixel, then subpixels are located by SR mapping performed either by simulated annealing or pixel swapping in unsupervised way. However, the limitation of these algorithms is the requirement of high computational load because of large number of spectral bands of HSIs. In a method proposed by Zhao et al. [61] used trained dictionaries created from different PAN images which are rich in edges and texures. By utilizing the sparse representation and spectral regularization based on linear mixing model (LMM) they obtained HR HSI. Here all bands are super-resolved individually without applying dimensionality reduction, which increases computational complexity of the algorithm.

In this chapter, we present a novel approach for super-resolution of HSIs that uses CS theory and Gabor prior. The compressive sensing theory is used for obtaining a close approximation to SR which in turn is used to obtain the final solution by using regularization framework in which we use a new prior called "Gabor prior" which is based on a bank of bandpass filters. Our method makes use of HR and LR registered HS training images to create CS dictionaries corresponding to LR and HR data. This is one time offline process. The training set consists of the HS data of an HR HSI sensor that can also be used for capturing the LR data. One approach to super-resolve HSI is to obtain SR for each spectral band separately. But this results in two major problems (i) Due to hundreds of spectral bands computational load is increased manifold, (ii) HS bands are highly correlated, hence considering each band independently do not exploit the correlation among them explicitly which in turn results in changing the pure spatial colors resulting in spectral distortion. Hence in this chapter, we first used principal component analysis (PCA) to reduce the dimensionality of the HSI. Since most of the information is contained in first few principal components, we apply SR reconstruction to these few PCA components to obtain their HR counterparts, thereby greatly reducing the computational complexity. Using CS based approach on this reduced set we obtain an initial SR image. Since SR is an ill-posed problem, we improve the solution by using regularization that uses a suitable prior. Our prior called "Gabor prior" consists of outputs of bandpass filters which correspond to bandpass features of LR image and initial SR estimate. Use of this prior restricts the solution space of final SR image. While regularizing the solution, degradation matrix entries are not assumed to be fixed rather

estimated from the initial SR image separately for significant PCA bands. Experimental results are validated using 31-band natural HSI captured under controlled laboratory environment and 224-band remotely sensed HSI. Experimental results show that our method improves spatial resolution without introducing considerable spectral distortions. Visual and quantitative comparison validates the effectiveness of proposed algorithm.

5.2 Block Diagram of the Proposed Algorithm

A block diagram showing the procedural flow of the proposed approach of CS based SR of hyperspectral image is illustrated in Figure 5.1. Given LR test image and a training database of LR-HR HS images, the proposed technique is implemented using the following steps:

1. Reduce the dimensionality of the LR test HSI using PCA. We now have PCA transformed LR primary images consisting of most of the information of HSI.

2. Apply CS based approach to primary PCA components to obtain initial SR images in PCA domain.

3. Use the initial SR image to estimate the degradation matrix for each of the primary PCA components obtained in step 2 that represents the observation model.

4. Regularize using Gabor priors and observation model estimated in step 3 to obtain the final super-resolved images in PCA domain.

5. Apply inverse PCA to obtain final SR HSI in spatial domain.

The spectral content of HSIs are inherently low dimensional, hence we exploit it by using PCA, a standard tool for analysis of multivariate data [8]. In our approach, we first represent the HS observations from different wavelengths, as a weighted linear combination of small number of basis image planes using principal component analysis (PCA) transform. The first few principal components referred to as primary components contain most of the information of HS observations and remaining PCA components referred to as secondary components contain very less information. In our work we do not consider secondary components as they represent very small portion of total information.

The proposed super-resolution algorithm is applied on the reduced set of primary PCA images to decrease computational burden of the algorithm.

In the next step, we need trained LR and HR dictionaries in CS based approach to obtain initial SR estimate. LR and HR raw dictionaries of respective PCA components, consisting larger number of atoms are generated using training database of LR and HR hyperspectral images. These raw dictionaries of the PCA components are jointly trained using K-singular value decomposition (K-SVD) algorithm [145], obtaining optimum number of atoms in each dictionary. We train a pair of dictionaries for each of the PCA band. An initial SR is obtained by using these dictionaries in CS based reconstruction.

We assume linear image observation model for the proposed SR algorithm. The LR image is modeled as the aliased and noisy version of the corresponding HR image. Using initial SR PCA image and LR test PCA image, degradation matrix entries are estimated for each PCA component. Estimated degradation matrices are used to further regularize the initial solution.

Our method of obtaining initial SR estimate do not consider the contextual dependencies among pixels as it is patch based. This results in artifacts in initial SR image around the patch boundaries. Hence prior knowledge about HS imagery is utilized in order to obtain better solution. Regularization based on Gabor prior is performed in order to obtain final solution for each of the initial SR PCA components. A simple optimization technique like gradient descent is used to minimize the cost function. This results in final SR HSI in PCA domain. Applying inverse PCA transformation on primary SR PCA components yields final super-resolved HS image.

5.3 Proposed Approach

5.3.1 Use of PCA and CS for Dimensionality Reduction and Sparseness Estimation

Hyperspectral image consists of large number of spectral bands and the spectral content of HS images are inherently low dimensional, hence this must be exploited. In this work, we first use principal component analysis (PCA) on LR test image as well as on LR-HR pairs of training HSIs to construct dictionaries. After learning these dictionaries using

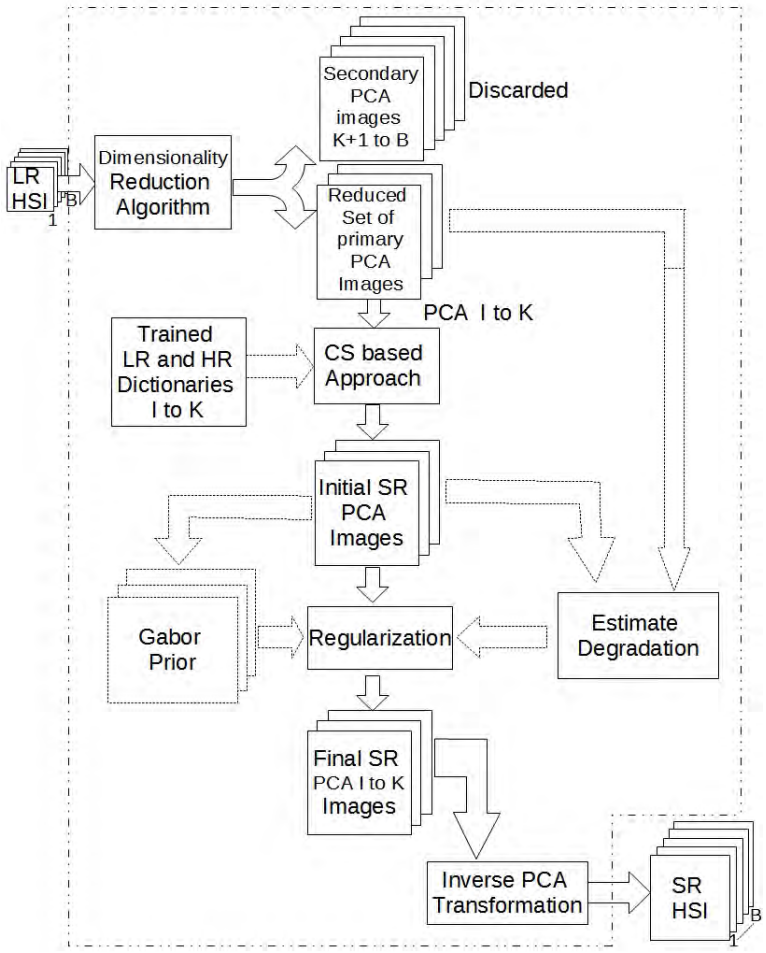


Figure 5.1: Detailed block diagram of proposed approach for HS image super-resolution algorithm. Here blocks are not drawn as per scale. The size of LR HSI is $M \times M \times B$, primary PCA images are of size $M \times M \times K$, secondary PCA images are of size $M \times M \times (B-K)$, initial SR PCA and final SR PCA are of size $rM \times rM \times K$, and SR HSI is of size $rM \times rM \times B$, where r is super-resolution factor.

K-SVD we use CS framework to obtain initial SR of all significant PCA components. In chapter 4, we already discussed on PCA in section 4.2.1. Here it is discusses to maintain continuity in discussion of present work.

Suppose we have a dataset of hyperspectral bands, represented by the matrix $L = [\mathbf{L}_1, \mathbf{L}_2, ..., \mathbf{L}_B]_{M^2 \times B}$, where $\mathbf{L}_i, i = 1, 2, ..., B$ is the i^{th} hyperspectral band of size $M \times M$ arranged in lexicographical order and B is the total number of bands of HSI. A set of eigen vectors $E = [\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_B]^T_{B \times B}$ are computed from the covariance matrix, $\sum = \sum_{i=1}^{B} (\mathbf{L}_i - \mathbf{m}_L)^T (\mathbf{L}_i - \mathbf{m}_L)$, where $\mathbf{m}_L$ is the average image intensity defined by $\mathbf{m}_L = \frac{1}{B} \sum_{i=1}^{B} \mathbf{L}_i$. Here the top $K \ll B$ eigen vectors corresponding to maximum eigenvalues contribute to maximum information of HS observations and the remaining $B - K$ eigen

vectors cover very less information. By projecting HS observations on these eigen vectors we obtain K number of primary PCA components and $B - K$ number of secondary PCA components each of dimension $M \times M$. Here we use compressive sensing (CS) framework which is already described in chapter 3 in section 3.1.1. In present work, we use dictionaries constructed from the available empirical data as discussed in chapter 4 sections 4.2 and 4.3.1. The difference lies in the use of joint dictionary learning using K-SVD algorithm on all raw dictionaries of significant PCA components.

5.3.2 Generating Trained Dictionaries

Learned dictionary provides more compact representation of the signal compared to raw dictionary which simply samples large amount of patches. This results in substantial reduction in computation while estimating the initial approximation to SR. CS based approach used in our earlier work is further extended here by using joint learning of dictionaries. Here the CS is applied on all the primary PCA bands and the initial estimates are derived for each of the primary bands. We create a set of HR and corresponding set of LR patches using a training database. Here we choose training images of the same class so that basis vectors of PCA better represent the materials of interest. The choice of basis vectors depends on the field of application. Note that dictionaries are not created directly from all bands of HS image, rather we use primary PCA images corresponding to maximum variability of HSI.

We first apply PCA on the training HR and LR HSIs and retain only primary PCA components of HR and LR both, then work on the first three primary components to obtain SR in PCA domain. We construct the joint dictionaries by randomly choosing raw patches from the HR and corresponding LR PCA components. Considering a resolution factor of $r = 4$, we select HR patch of size 8×8 (i.e. atom of 64×1 vector) and corresponding LR patch of size 2×2 (i.e. atom of 4×1 vector). Appending the LR to HR vector we obtain a joint vector of size 68×1. This way, we obtain a joint raw dictionary for each PCA component, having large number of patches (100000). We now have three dictionaries of raw patches corresponding to three primary components each of size 68×100000.

By assigning different weightages to HR and LR patches, two dictionaries in the high

and low-resolution spaces are balanced while training to achieve better initial estimate. These dictionaries were trained using K-SVD algorithm [145], obtaining optimum number of atoms in each dictionary. During dictionary training, we keep arbitrary number of atoms to represent each signal until a specific representation error is reached. It is to be noted that this kind of dictionary learning reduces the reconstruction error while obtaining the initial estimate. This way the number of atoms of joint dictionary are reduced to 1000 and the size of each dictionary becomes 68×1000. From the joint dictionary we separate out HR and LR trained dictionaries to obtain two dictionaries each of size 64×1000 and 4×1000, respectively. Now we have HR and LR dictionaries D_{Hm} and D_{Lm}, $m = 1, .., K$ corresponding to K number of primary PCA components. It is to be noted that dictionary training is one time and offline procedure. Using these dictionaries and LR HSI test images as inputs, initial SR PCA estimates are generated using CS based approach. Corresponding to each primary PCA component, we have a pair of LR-HR dictionaries. These PCA dictionaries are used in CS based approach to obtain initial SR estimates in PCA domain.

5.3.3 Initial Estimate of Super-resolution

Here we use dictionary based approach on the primary principal components to obtain the initial SR estimates in PCA domain. Due to similar statistical properties, HS test image patches can be represented as a sparse linear combination of LR dictionary elements using equation (3.1). We assume that the same sparsity holds good for its corresponding HR image which is unknown. Hence one may recover HR image using the HR dictionary.

Given LR HS test image cube L of size $M \times M \times B$ and trained dictionaries D_{Hm} and D_{Lm}, $m = 1, .., K$, the proposed algorithm to obtain initial SR images (primary SR PCA) of size $rM \times rM$, where r is the super-resolution factor, is described below.

1. Generate mean subtracted LR HS test image $L_{ms} = [\mathbf{L_{1ms}}, \mathbf{L_{2ms}}, .., \mathbf{L_{Bms}}]_{M^2 \times B}$, where B is the total number of bands of HSI. Here all mean subtracted bands of LR HS image are arranged lexicographically to convert them in vectors of size $M^2 \times 1$.

2. Obtain basis eigen vectors of the covariance matrix $C = [L_{ms}^T L_{ms}]_{B \times B}$. Retain primary basis eigen vectors $\mathbf{e_{lm}}, m = 1, ..., K$, $K \ll B$ corresponding to maximum variability of data (i.e., highest variance), each of size $1 \times B$. The percentage of

information retained in primary PCA components is given by

$$\text{Information retained} = \frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{B} \lambda_i} * 100. \tag{5.1}$$

3. Create LR transformed images $Y_m, m = 1, ..., K$ by projecting L_{ms} on primary basis eigen vectors generated in step 2

$$\mathbf{Y_m} = \mathbf{e_{lm}} \times L_{ms}^T;$$

$\mathbf{Y_m}$ is still a vector representing m^{th} image. Compute $\mathbf{Y_m}, m = 1, ..., K$ and obtain matrix of size $M \times M$ for each test image. We now have K number of primary PCA transformed test images each of size $M \times M$.

4. Consider a patch of size $b \times b$ from the transformed m^{th} LR test image Y_m. Convert it into lexicographic order to obtain a vector $\mathbf{y}$ of size $b^2 \times 1$.

5. Solve CS based l_1-minimization optimization problem i.e.,

$$min||\mathbf{x}||_1 \text{ such that } \mathbf{y} = D_{Lm}\ \mathbf{x}.$$

Here $\mathbf{x}$ gives sparse representation of test patch $\mathbf{y}$ in terms of LR dictionary (D_{Lm}) atoms.

6. Obtain SR patch using $\mathbf{z} = D_{Hm}\ \mathbf{x}$. This is the SR patch in the PCA transform domain.

7. Repeat steps (4) to (6) for all patches of Y_m to obtain transform domain HR image Z_m. The same procedure is repeated for other primary images that gives initial SR approximation to all primary LR components. Note that the spatial dependency still exists within the pixels of these transform domain SR images. However, different PCA components of HS image are uncorrelated. Hence there is no spatial dependency of pixels among different PCA components.

Thus we obtain the initial estimate of the SR image for each test image having a size of $rM \times rM$. We considered patch size of $b = 2$ for $r = 4$. Note that the inverse PCA using $Z_m, m = 1, 2, ..., K$ can be used to obtain super-resolved HSIs in the spatial domain.

5.4 Final Solution using Regularization

Since we are estimating the initial SR image using patch based approach, the spatial correlation is not considered, as the sparse representation of patches is done independently. Hence we need to regularize it further to obtain a better solution. We restrict the solution space for the SR image by using our proposed Gabor prior. For regularization purpose one needs to have data fitting term and regularization term. Hence we first model the image formation in order to get the data fitting term.

5.4.1 Observation Model

A linear image observation model is used to relate the desired HR image to the observed LR image for decimation factor of r. Continuing with the transformed components the observed LR HSIs are modeled as decimated and noisy versions of the corresponding HR HSIs. Let $Y_m(m = 1, 2, ..., K)$ be the LR PCA image of m^{th} PCA band of size $M \times M$ and and Z_m be the corresponding HR PCA image of size $rM \times rM$, then the model of image formation is represented as:

$$\mathbf{y_m} = D_m \mathbf{z_m} + \mathbf{n}, \quad m = 1, 2, ...K, \tag{5.2}$$

where $\mathbf{y_m}$ and $\mathbf{z_m}$ represent the lexicographically ordered vectors of size $M^2 \times 1$ and $r^2M^2 \times 1$, respectively with $\mathbf{z_m}$ representing the SR vector to be estimated. Here $\mathbf{n}$ is the independent and identically distributed (i.i.d.) noise vector with zero mean and variance σ_n^2 and has the same size as that of $\mathbf{y}$. D_m is the downsampling/decimation matrix taking care of aliasing caused as a result of downsampling. For an integer downsampling factor of r, matrix D_m consists of r^2 non-zero elements along each row at appropriate locations. It models the integration of light intensity that falls on the HR detectors of corresponding spectral bands.

In most of the earlier research, either implicitly or explicitly the same degradation matrix D_m with fixed entries is considered to construct degradation model for all bands of the multispectral and HSIs [23, 13, 61]. This clearly means that LR pixel of any band is considered as equally weighted sum of corresponding r^2 HR pixels for all bands, i.e., the ideal squared response optical point spread function (PSF) is considered. Generally,

the decimation matrix used to model aliased pixel intensities from the corresponding HR pixels for a decimation factor of r, has the form [126]

$$D_m = \frac{1}{r^2}\begin{pmatrix} 1\,1\ldots 1 & & \mathbf{0} \\ & 1\,1\ldots 1 & \\ \ldots & \ldots & \ldots \\ \mathbf{0} & & 1\,1\ldots 1 \end{pmatrix}, \tag{5.3}$$

In practice, many factors like diffraction, shape, location, physical construction and electronic response of the detectors, and the electronics of the amplifications contribute to the PSF of any spaceborne radiometer. The effect of diffraction is significant at higher wavelength in the HS imager. This results in spatially and spectrally varying PSF of degradation function. For more details on this readers may refer to the [17].

In our work, we do not consider LR pixel as sum of equally weighted HR pixels, rather we estimate the alias by estimating the entities of matrix D_m. For the estimation of aliasing we need true HR image which is not available. Since we have availability of initial estimate of SR image, we estimate decimation matrix D_m using the the available LR PCA test component and initial SR PCA component. Then the form of decimation matrix for m^{th} LR-HR pair is modified as below

$$D_m = \begin{pmatrix} d_1^m\ d_2^m \ldots\ d_{r^2}^m & & \mathbf{0} \\ & d_1^m\ d_2^m \ldots\ d_{r^2}^m & \\ \ldots & \ldots & \ldots \\ \mathbf{0} & & d_1^m\ d_2^m \ldots\ d_{r^2}^m \end{pmatrix}, \tag{5.4}$$

where $0 < d_i^m < 1, i = 1, 2, ...r^2$ are unknown. Here we use a simple least squares approach to estimate the decimation coefficients d_i. It is worth to mention that for each primary PCA component we estimate the decimation matrix D_m i.e., different PSF is considered for each primary spectral basis. Hence the estimated D_m matrix for each PCA component is close to the true degradation of HSI, hence incorporation of this degradation model leads to better solution.

5.4.2 Regularization using Gabor Prior

Hyperspectral images contain various textured regions having different frequency contents. Hence it is necessary that these frequency details are preserved in the SR image. This can be achieved using a prior that incorporates the information about the details at various frequencies. In computer vision community, a linear filter named Gabor filter, is widely used for feature extraction at various bandpass frequencies. Frequencies and orientation representation of this filter are similar to those of the human visual system and they have been found very useful for texture representation and discrimination. This motivates us to use Gabor prior for regularization in our work. The impulse response of Gabor filter is given by [155]

$$G\left(p,q,f,\theta,\sigma_p,\sigma_q\right) = e^{-\frac{1}{2}\left(\frac{p'^2}{\sigma_p^2}+\frac{q'^2}{\sigma_q^2}\right)} cos(2\pi f p'), \tag{5.5}$$

where (p,q) represents spatial coordinates, $(p',q') = (pcos\theta + qsin\theta, -psin\theta + qcos\theta)$, σ_p and σ_q are the spatial extent of the filter in p and q directions, respectively. Here f is the center frequency of sinusoidal carrier wave, and θ is its orientation.

Using a data-fitting term, and Gabor prior terms, the final cost function to be minimized for each PCA band image $m = 1, 2, ..., K$ can be written as,

$$\epsilon_m = ||\mathbf{y_m} - D_m\mathbf{z_m}||^2 + \lambda_1 \sum_{j=1}^{Q} ||G_j\mathbf{y_m} - G_j(D_m\mathbf{z_m})||^2 + \lambda_2 \sum_{j=1}^{Q} ||G_j\hat{\mathbf{z}}_\mathbf{m} - G_j\mathbf{z_m}||^2, \tag{5.6}$$

where $\mathbf{y_m}$ is m^{th} band of LR, D_m represents degradation matrix for m^{th} band, estimated using the LR and initial SR image. Here $\hat{\mathbf{z}}_\mathbf{m}$ and $\mathbf{z_m}$ are the SR images of m^{th} band of initial estimate and final estimate, respectively. G_j is j^{th} Gabor filter matrix representing the impulse response given in equation (5.5), and Q is the total number of filters in the Gabor filter bank. λ_1 and λ_2 represent the weightages given to the second and third term, respectively, chosen empirically. This way we obtain K number of primary SR PCA components. It is to be noted that while solving the HSI SR problem degradation matrix is usually considered as fixed entries for all spectral bands [13, 5]. However, in this work a matrix with different entries is considered for all spectral bands. This way an optimum linear observation model is considered for all primary bands. Our prior in second term of equation (5.6) imposes the condition that degraded SR image should possess features

similar to that of the LR test image when viewed at different frequencies. This means we look for a solution i.e., SR image whose downsampled version has the same Gabor features as that of LR input image when passed thorough the same Gabor filter bank. This is illustrated in Figure 5.2. Similarly, final term in equation (5.6) indicates that features of different frequency contents in final SR image should be identical to that of initial SR estimate. Use of the available initial SR estimate as an initial solution speed-up the convergence. Applying inverse PCA to all primary SR PCA images results in SR hyperspectral image in spatial domain.

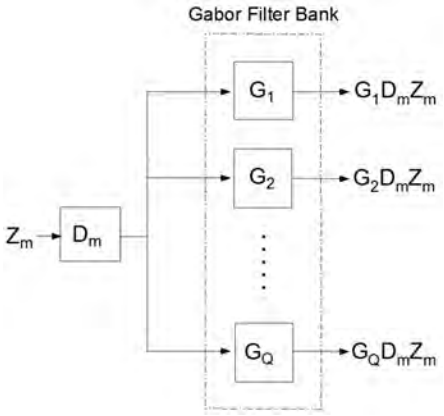


Figure 5.2: Gabor prior for SR image Z_m. Here the output $G_j D_m Z_m$ represents the image details at a particular frequency band which have to match with the details of Y_m when it is passed through the same filter G_j. $\{G_j\}$, $j = 1,\ldots,Q$ represents a Gabor filter bank.

5.5 Experiments and Result Analysis

In this section, we show the effectiveness of the proposed method by conducting experiments on two different data sets: (1) Natural hyperspectral images, and (2) Remotely sensed HSIs (AVIRIS). The data sets used in the experiments constitute images with 31 and 224 spectral bands, respectively. Detailed analysis of the results is performed on 224 band AVIRIS HSI. We performed experimentations for $r = 2$ as well as for $r = 4$. Due to the space constraint, we are demonstrating results only for the case of $r = 4$. For all our experiments, the step size for gradient descent algorithm was chosen as 0.01. The weightage to second and third terms were set as $\lambda_1 = 0.14$ and $\lambda_2 = 0.17$, respectively in the regularization equation (5.6). These were chosen by trial-and-error procedure for all experiments. Here, one can use a generalized cross-validation technique [156] to identify

the optimum values of λ_1 and λ_2, but it is computationally expensive and it is specific to a given image only. We show the visual as well as quantitative comparison for experiments on both data sets. Different quantitative measures used in our experiments are described in the following section.

5.5.1 Quantitative Evaluation Measures

For quantifying the results on 31-band natural HSI we used mean squared error (MSE) as a preliminary evaluation index which is discussed in chapter 3 section 3.3.1 (see equation 3.15), where $F_{i,j}$ and $\hat{F}_{i,j}$ represent the true HR (groundtruth) and the SR images, respectively. Detailed quantitative evaluation of spatial and spectral fidelity of super-resolved hyperspectral images is performed using different measures such as correlation coefficient (CC), erreur relative globale adimensionnelle de synthese (ERGAS), and $Q2^n$. CC and ERGAS are defined in chapter 4 section 4.4.1 (see equations 4.6, 4.7 and 4.9). What follows is the description of $Q2^n$.

$Q2^n$ [157] index is derived from the theory of hyper-complex numbers of 2^n-ons (pronunciation: two-to-the-any-ons) [158]. It takes into account the correlation, mean of each spectral band, intra-band local variance, and the spectral angle. Both spectral and spatial distortion metrics are encapsulated in this index. It takes a real value in the interval 0 to 1, with 1 being the best value. It is defined between k^{th} super-resolved (F) and groundtruth ($\hat{F}$) bands as:

$$Q2^n_k = \frac{cov(F,\hat{F})}{\sigma_F \sigma_{\hat{F}}} \cdot \frac{2||\bar{F}||||\bar{\hat{F}}||}{||\bar{F}||^2 + ||\bar{\hat{F}}||^2} \cdot \frac{2\sigma_F \sigma_{\hat{F}}}{\sigma_F^2 + \sigma_{\hat{F}}^2}, \tag{5.7}$$

where $k = 1, ..., B$ and $cov(F, \hat{F})$ is the covariance between bands F and $\hat{F}$. σ_F^2 and $\sigma_{\hat{F}}^2$ are the variances of F and $\hat{F}$, respectively. Here we averaged $Q2^n$ over the all HSI bands to get a global measure of spatial and spectral distortion of the super-resolved HSI i.e.,

$$Q2^n = \frac{1}{B} \sum_{k=1}^{B} Q2^n_k, \tag{5.8}$$

where B is the total number of bands in hyperspectral image, which is 196 in our case.

5.5.2 Experiments on Hyperspectral Images

The first set of our experiment consists of 31-band reflectance images of natural scene, having spectral range of $400nm - 700nm$ all acquired under the direct sunlight in clear or almost clear sky [149]. Our second HSI data set is comprised of 224 bands of AVIRIS HSI cube [1]. After discarding few bands having low signal to noise ratio (SNR), 196 bands were used for super-resolving by a factor of 2 and 4, respectively. The band removal was based on visual inspection of the images. The above data sets have high spatial dimensions and hence specific regions are cropped from them and experiments are carried out on the cropped regions. Here, we do not have the true LR-HR pairs of HSIs. Hence the low spatial resolution (LR) images were created from these cropped images by using filtering and downsampling operations. The whole HSI cubes are used to generate raw dictionaries by random selection of patches. Use of same cube to create raw dictionaries ensures inclusion of large number of materials and objects of interest. Note that we need LR-HR pairs to construct dictionaries. If these images are acquired offline they can be utilized to form the pairs. One may also use the LR and HR images of the same scene captured by using different sensors but after applying the radiometric and geometric (registration) corrections. In order to evaluate the performance of our approach using quantitative measures, we need the groundtruth images. Since these images are not available, we consider original cropped HSIs of size 256×256 as ground truths and generated the LR HSIs of size 128×128 and 64×64 by applying downsampling operation by a factor of $r = 2$ and $r = 4$, respectively. The SR algorithm was then applied on these LR HSIs. In order to restrict the maximum spatial frequency in the image we use low pass filtering operation before downsampling. The low pass filtering operation was performed and tested using Gaussian filter with standard deviation of 0.5. For this purpose we used filter mask of size 5×5. While performing joint training of dictionaries weightages given to HR patches and LR patches were 0.65 and 0.35, respectively.

5.5.3 Experiments on 31-band Natural Hyperspectral Image

In this section, effectiveness of the proposed algorithm on the 31-band natural HSI is evaluated. Here we use cropped region of "Scene 5" of hyperspectral images of natural

[1] Aviris Free Data, Jet Propulsion Lab., California Inst.Tech., Pasadena, http://aviris.jpl.nasa.gov/html/aviris.freedata.html

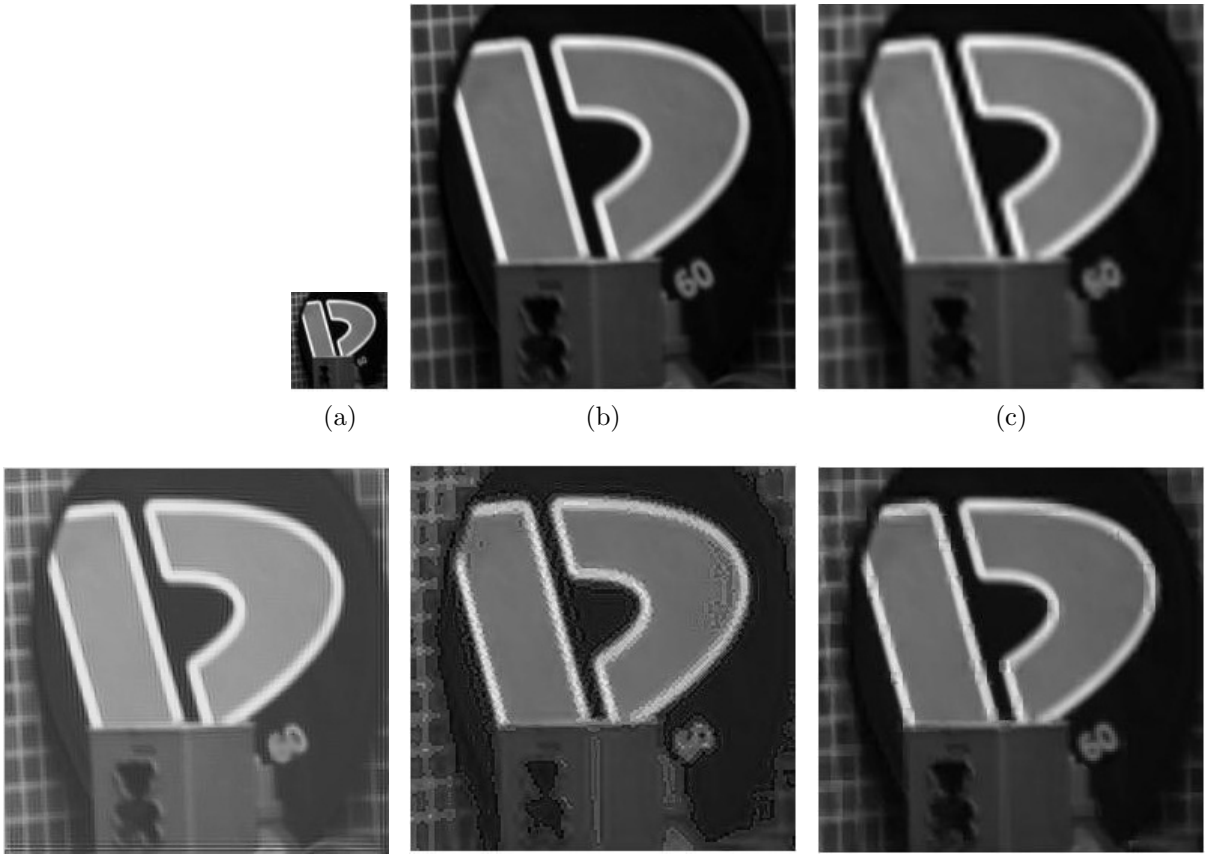


Figure 5.3: Experimental results on PCA-1 of 31-band natural HSI for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152] (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach.

scenes 2002 [2] as test data. In this case we found that 99.4% of spectral variance is covered by first three principal components. Hence we retained these three PCA components and applied SR algorithm in our experiment. Figure 5.3 displays the results for the PCA-I component. Figures 5.3(a) and (b) display the LR test image and the ground truth image of size 64×64 and 256×256, respectively. The results obtained using different methods are shown in Figures 5.3(c-f). Visual inspection of images in Figures 5.3(c-e) indicate that the white borders of "D" shape mounted on the box appear blurred in the bicubic interpolated image, iterative backprojection (IBP), and the method proposed by Yang et al. [79]. One can see in Figure 5.3(f) that the SR image obtained using the proposed method compares well with the groundtruth. The white border of "D" shape is sharper

[2] http:// personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_02.html

Table 5.1: Quantitative evaluation measures for SR of 31-band Natural hyperspectral image using different techniques for $r = 4$

Quantitative Measures	Bicubic interpolation [148]	Iterative backprojection [152]	Yang et al. [79]	Proposed approach
MSE-PCA-1	0.0072	0.0069	0.0065	0.0053
MSE-PCA-2	0.0388	0.0368	0.0563	0.0308
MSE-PCA-3	0.7487	0.6716	0.5940	0.5935
CC_{avg}	0.9681	0.9806	0.9827	0.9857
ERGAS	6.6613	5.2086	4.9888	4.6388
$Q2^n$	0.9643	0.9789	0.9796	0.9801

and closer to groundtruth. Similarly, the number "60" in the image appear more clear in Figure 5.3(f) compared to others shown in Figures 5.3(c-e). Quantification of this experiment is provided in Table 5.1. From the table we can see that the MSEs between the true and the estimated SR PCA components are significantly less for the proposed method. The use of CS based approach using learned dictionaries and regularization using Gabor prior improves the results in our approach as evident from quantitative evaluation measures such as CC_{avg}, ERGAS, and $Q2^n$. Note that the MSE is computed on PCA bands directly while all other measures are computed on 31 bands after performing the inverse PCA. These measures show that proposed approach better preserves spatial and spectral fidelity in the super-resolved hyperspectral images.

5.5.4 Experiments on AVIRIS Hyperspectral Image

We now discuss the SR results for 224-band AVIRIS hyperspectral image. This data set is comprised of 224-band real hyperspectral image of Moffett Field acquired by AVIRIS hyperspectral imaging system. In this experiment, cropped region of an urban area in Moffett Field is used as test data. This cropped region is specifically chosen to include various bandpass components in the image to evaluate the performance of the proposed method. The SR results on remotely sensed data acquired using AVIRIS hyperspectral imager is shown with reduced dimension. Figures 5.4, 5.5 and 5.6 show the SR results of first, second and third PCA bands, respectively. Quantitative results are listed for the first three PCA bands that include 97.56% of spectral variance of HSI. Here the measures listed in Table 5.2 are computed over 196 bands. Figures 5.4-5.6 (a) show the LR PCA test images of size 64×64 and the original PCA bands of size 256×256 are displayed in

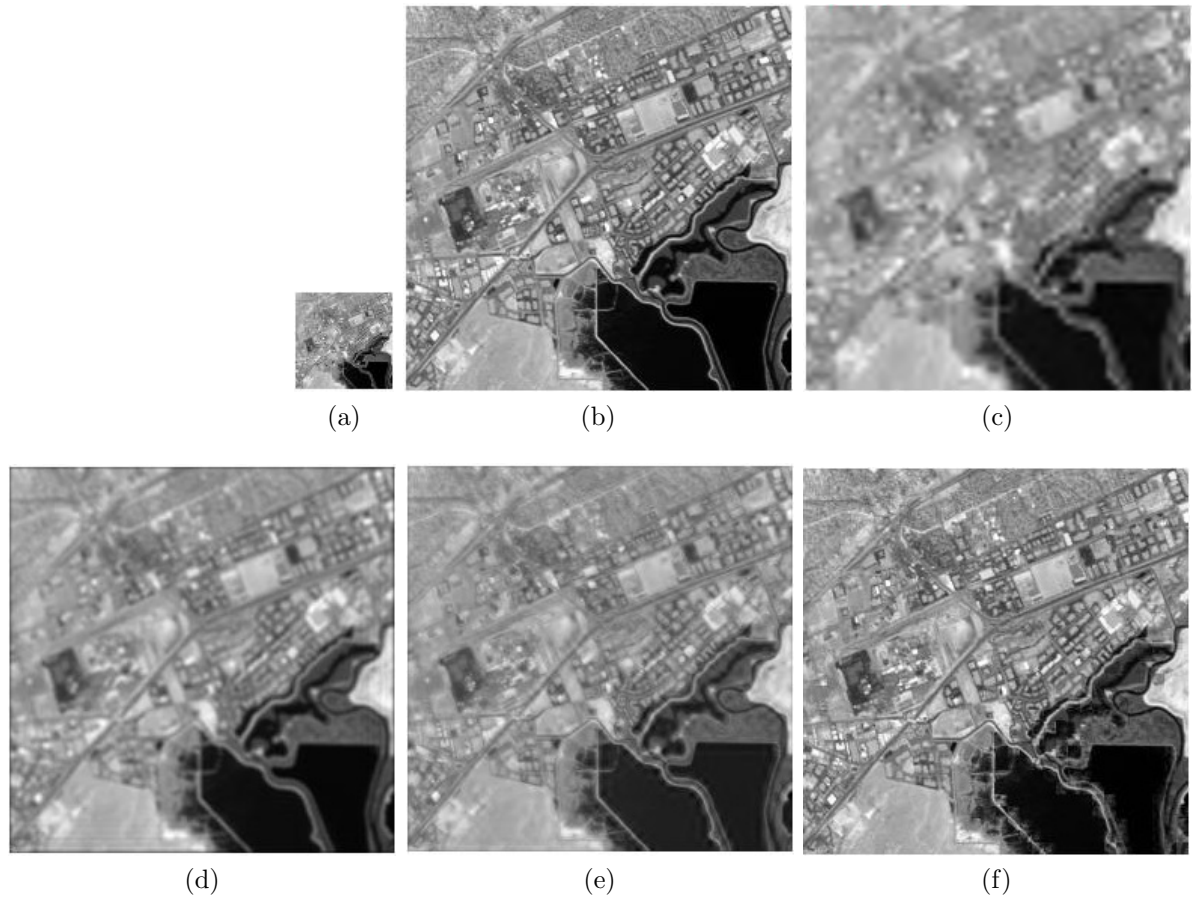


Figure 5.4: Experimental results on PCA-I band of AVIRIS data for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach.

Figures 5.4-5.6 (b).

From Figures 5.4-5.6(c) we can see that when the PCA images are upsampled using bicubic interpolation they become blurred and the high frequency spatial details are lost. Roads and buildings are no longer visible in bicubic interpolated image in the PCA-I and PCA-III results displayed in Figure 5.4(c) and Figure 5.6(c), respectively. One may notice that bicubic interpolation in Figure 5.3(c) appears better than the result in Figure 5.4(c). This is because there is significant high frequency content in AVIRIS data when compared to natural HSI displayed in Figure 5.3(c) and the interpolation fails to preserve the high frequency details. This indicates that the interpolation techniques are not suitable for solving the SR problem, hence they are not considered as SR techniques. Quantitative comparisons for bicubic interpolation given in Table 5.1 and Table 5.2 further proves this observation. SR result on PCA-I of IBP [152] shown in Figure 5.4(d)

is less blurred compared to bicubic interpolated image but the overall contrast of the image is not preserved by this method. Visibility of roads and buildings has improved over bicubic interpolation in all super-resolved PCA components as seen from Figures 5.4-5.6(d). Borders of objects appear blurred in SR PCA images obtained using IBP method (Figures 5.4-5.6(d)). Sparsity based SR results of [79] method shown in Figures 5.4-5.6(e) are visually better than bicubic interpolation and IBP method, but it fails to preserve high frequency details as evident from top half portions of the images. As seen from Figures 5.4-5.6(f), the use of compressed sensing and Gabor priors regularization results has reduced artifacts and also takes care of preservation of different frequency details. Sharpness of different objects such as roads and buildings has improved over all other methods, particularly noticeable in PCA-I and PCA-II SR images shown in Figures 5.4(f) and 5.5(f), respectively. The visual quality of SR PCA images in Figures 5.4-5.6(f) is closely matching with the groundtruth. We can see that the white patches visible in LR PCA-1 observation appear grayish in Yang et al. [79] method (See Figure 5.4(e)), but the result is improved in the proposed approach as seen from Figure 5.4(f). One can clearly discriminate the road lines joining the top right ($\frac{1}{4}^{th}$ way down) to bottom left corner of the image in Figure 5.4(f) indicating that edge details are well preserved in the proposed approach.

As far as the quantitative comparison is concerned, it is clear from Table 5.2 that the proposed method provides scores that are closer to the reference values shown in the same table when compared to bicubic interpolation [148], IBP [152], and Yang et al. approach [79]. Note that the quantitative measures CC_{avg} and $Q2^n$ are averaged over the 196 HSI bands. Lower value of ERGAS in the proposed method indicates lesser global distortion in super-resolved HSI. Generally a value of ERGAS below 3 is believed to be an image with good quality [124]. We see that when compared to other approaches CC_{avg} and $Q2^n$ are also better for the proposed method. Lower value of $Q2^n$ in Table 5.2 indicates minimum spatial as well as spectral distortions by the proposed approach. To further support the performance improvement using our approach, we show the plot of bands Vs CC in Figure 5.7 for various methods. From the average CC value listed in Table 5.2 and the plots in Figure 5.7, we can see that the proposed method better preserves the spatial details.

In order to compare the performance in terms of spectral fidelity, we show the spectral

reflectances of groundtruth and outputs of different SR algorithms at different regions. We have chosen three different regions to show the performance of various algorithms at different frequency bands. These regions include: (i) Uniform region (A) (very low frequency) (ii) Smooth edge region (B) (mid range frequency), and (iii) High frequency region (C) having sharp variation of texture. Spatial locations of these regions are shown in Figure 5.8(a). Spectral reflectances of different SR methods for the selected regions A, B and C in Figure 5.8(a) are shown in Figures 5.8(b), (c) and (d), respectively. Here the spectral reflectance for each region is computed by using a 3×3 patch in every region and computing the average reflectance. So, we have 3 vectors of 9×1 for each band corresponding to three regions and there are total of 196 bands. The plots showing the bands Vs spectral reflectance for various approaches including the original are shown in Figures 5.8(b), (c) and (d). Separate plots are shown for each region. The average computed over the bands is given in Table 3 for quantitative comparison. In Figure 5.8(a) region A represents smooth region having no significant reflectance variations. One can see from Figure 5.8(b) that the plots of all SR algorithms closely match that of the groundtruth as far as the region A is concerned. This is also evident in Table 5.3 where we can see that average spectral reflectances of bicubic interpolation, IBP, Yang et al., and proposed approach are closer to groundtruth. This indicates that low frequency regions are better preserved by most of the approaches. Region B has mid frequency content and in this case IBP as well as Yang et al. approaches perform better in addition to the proposed method. But, there is obvious deviation in bicubic interpolation as evident from Table 5.3 entries for region B. We can see that average spectral reflectance of bicubic interpolation is significantly deviating from the groundtruth when compared to IBP and Yang et al. methods. Region C has sharp variations in texture, and in this case the spectral reflectance plot of the proposed method is closer to groundtruth as evident from Figure 5.8(d). From Table 5.3, one can see that although the approach by Yang et al. [79] performs better than other two approaches, the proposed approach performs even better.

Before we conclude we would like to compare performance of this method with the previous method discussed in chapter 4. Since ERGAS represents the overall error in the super-resolved image irrespective of resolution factor, we use the same for comparison. From chapter 4 Table 4.1, one can see that the value of ERGAS is 7.065 in proposed

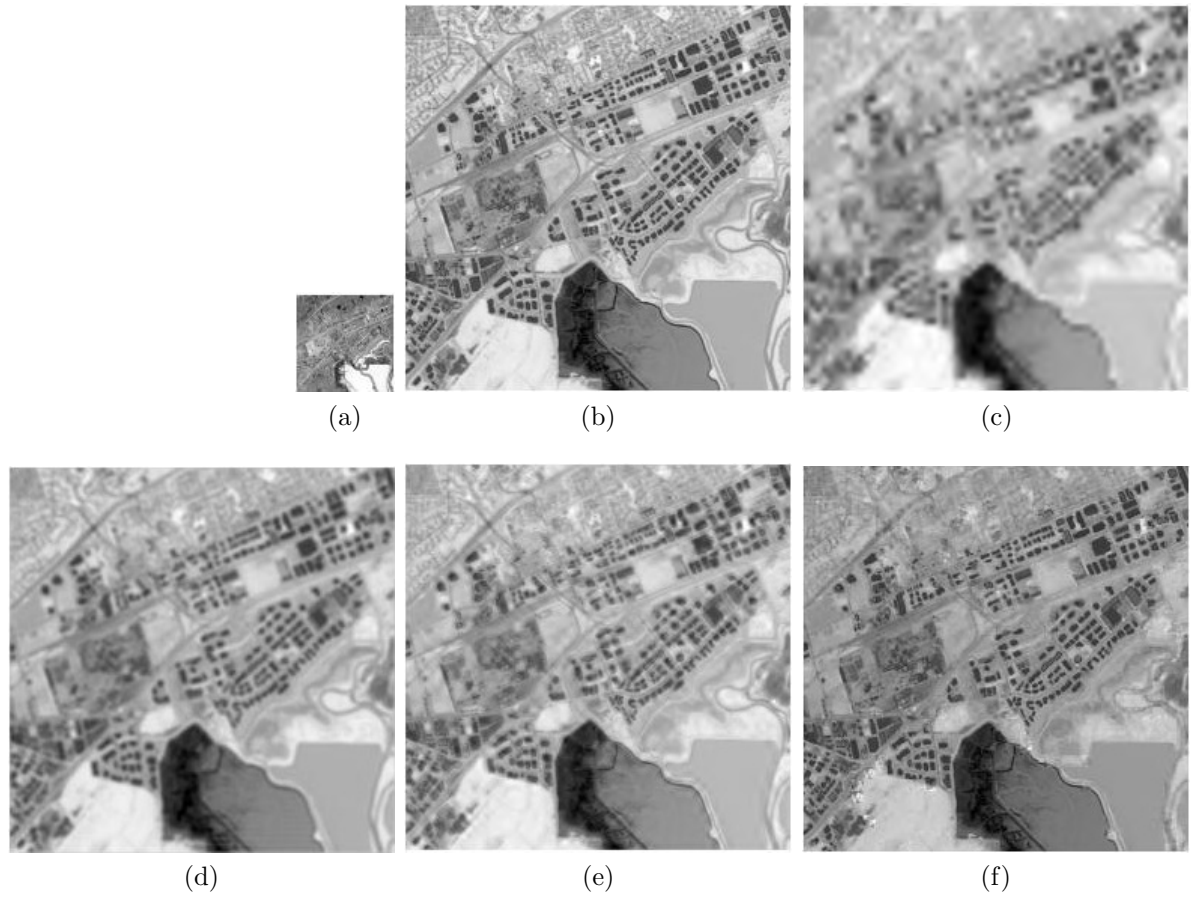


Figure 5.5: Experimental results on PCA-II band of AVIRIS data for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach.

approach when we considered a resolution factor of 2, while it is 2.9725 (see Table 5.2) by the approach proposed in this chapter, even though the resolution difference between LR and SR is 4. This indicates reduced global distortion for the approach proposed in this chapter. The use of learned dictionaries in compressive sensing based approach on all significant components improves initial estimate in the present method when compared to the use of raw dictionaries in chapter 4. Bicubic interpolation of remaining significant components (except first one) cannot preserve high frequency details in the initial SR estimate discussed in the previous chapter. The performance is further improved due to the use of estimated entries of decimation matrix here, as well as use of Gabor prior. In this approach, the trained dictionaries have the tendency to adapt to local structures of the images, and the regularization based on Gabor prior preserves the spectral as well as spatial information better.

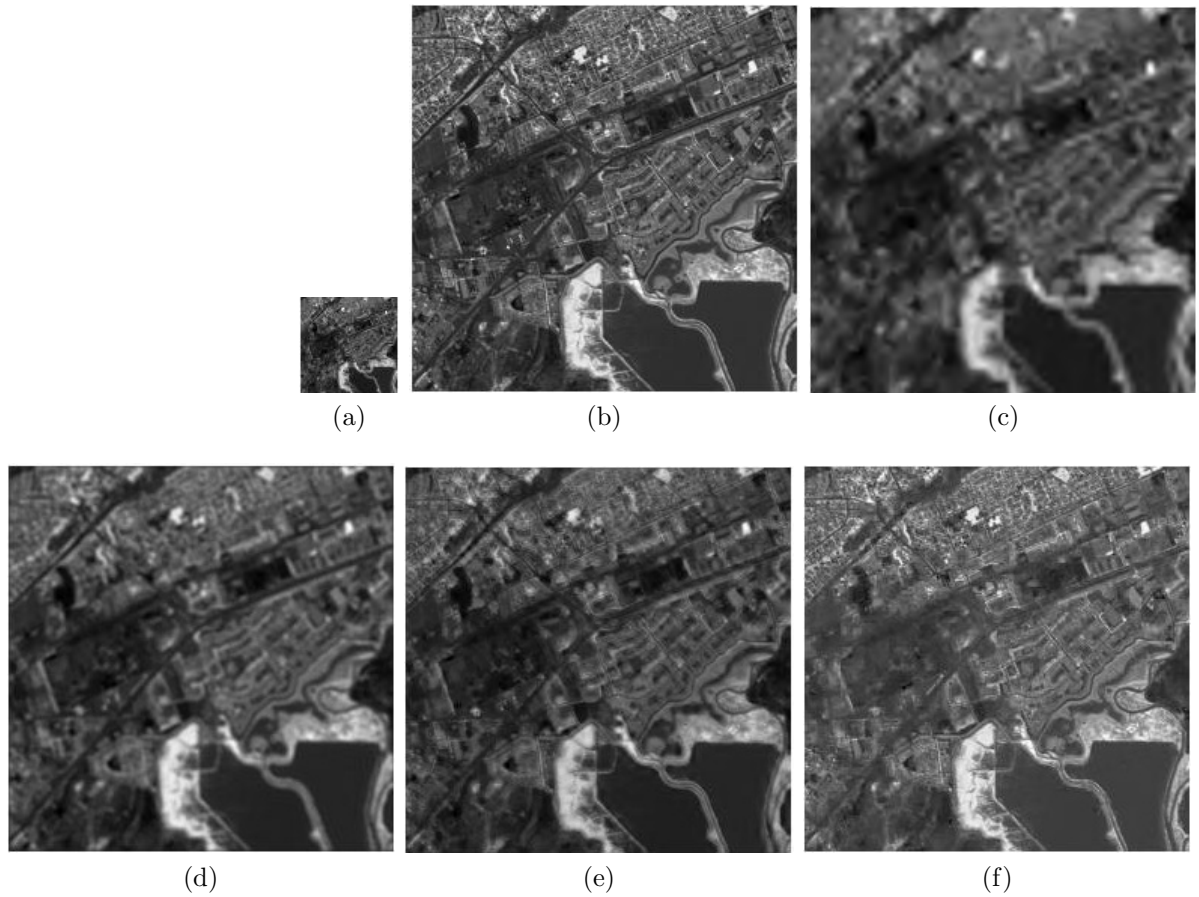


Figure 5.6: Experimental results on PCA-III band of AVIRIS data for $r = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) Iterative backprojection method [152], (e) SR image using Yang et al. method [79], and (f) SR image using the proposed approach.

5.6 Conclusion

We have presented SR algorithm for HSIs based on the compressed sensing theory in which jointly learned dictionaries are used to obtain SR images in reduced dimension space. We construct the raw dictionaries of LR and HR from a training database and used K-SVD algorithm to obtain learned dictionaries for all primary PCA components. Using learned dictionaries in CS based approach we obtain initial estimates of SR for each primary PCA component in the PCA domain. High spatial resolution initial SR images and the corresponding low-resolution observed images in PCA domain were used to estimate the decimation. Subsequently, a regularization scheme is employed using Gabor priors considering varying degradation or PSF in the spectral space. Gabor prior was considered on downsampled as well as HR versions of initial estimates.

Table 5.2: Quantitative evaluation metrics of AVIRIS SR for $r = 4$

	CC_{avg}	ERGAS	$Q2^n$
Bicubic interpolation [148]	0.9163	5.2020	0.9074
IBP [152]	0.9509	3.9862	0.9451
Yang et al. [79]	0.9635	3.4362	0.9513
Proposed Approach	0.9807	2.9725	0.9681
Reference	1.0000	0.0000	1.0000

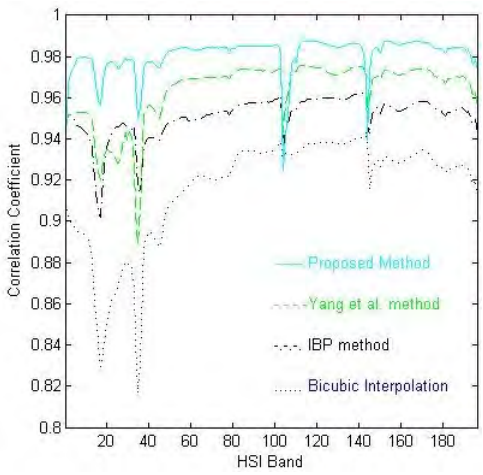


Figure 5.7: Correlation coefficient Vs band number for SR on AVIRIS data for $r = 4$

Table 5.3: Spectral reflectance (%) of 3×3 pixels averaged over 196 bands at different region locations shown in Figure 5.8(a) of AVIRIS SR for $r = 4$

	Region A	Region B	Region C
Bicubic interpolation [148]	81.2244	67.7482	58.9277
IBP [152]	82.3980	76.5466	66.7565
Yang et al. [79]	82.9710	86.1222	68.2510
Proposed Approach	85.2926	81.9088	72.9670
Ground truth	85.9464	82.9792	74.2430

The advantage of the proposed technique is that there is no need of auxiliary registered HR image or multiple LR observations of the same scene with subpixel shifts while super-resolving. Use of Gabor prior in regularization preserves features at bandpass spatial frequencies. Use of estimated entries of degradation matrices for all significant PCA components represent the optimum PSF in regularization that aids in obtaining better solution. Super-resolution results obtained using proposed method show better preservation of spatial details over those obtained using raw dictionaries and averaged PSF. Quantitative comparison of score indices show that our method enhances spatial information without introducing significant spectral distortion.

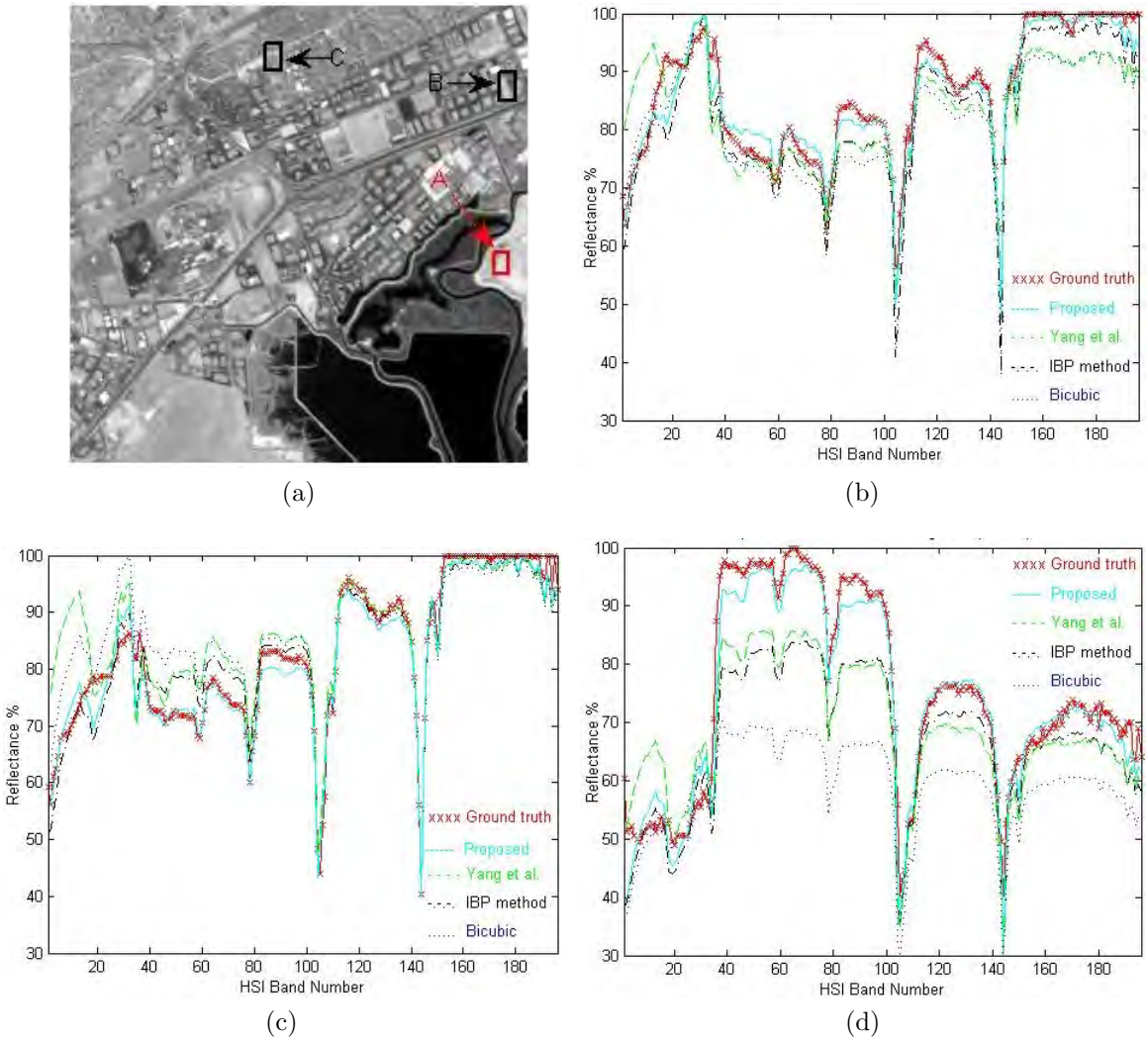


Figure 5.8: Spectral reflectance plots averaged over 3×3 pixels, of groundtruth and SR image obtained using different algorithms on AVIRIS data for $r = 4$. (a) Selected regions shown using the band 100, (b) Spectral reflectance at region A, (c) Spectral reflectance at region B, (d) Spectral reflectance at region C.

The proposed super-resolution technique uses learned dictionaries of LR and HR images to produce the initial super-resolved images in reduced dimensional space. It may be noted here that we estimate degradation using the initial estimate of SR. It is to be noted that degradation matrix entries considered here are non-overlapping and used while obtaining final results of SR. In the next, chapter we consider the estimation of degradation in the form of PSF representing the low-pass filtering by estimating filter coefficients. The estimated wavelet filter coefficients are used to define the degradation matrix with overlapping entries while obtaining initial as well as final super-resolved HSI.

Chapter 6

Super-Resolution using Optimum Wavelet Filter Coefficients and Sparsity Regularization

In the previous chapter we considered non overlapping degradation matix entries while regularizing the solution. That means, the considered degradation matrix (PSF) of sensor is depending only on the position of the detectors. In practice this is not true since the aperture of the optical device is limited and it introduces diffraction. In this chapter a new learning based approach for SR using the wavelet transform is proposed. The SR algorithm is applied on the reduced dimension obtained by using the principal component analysis (PCA). The novelty of our approach in this chapter lies in designing application specific wavelet basis. We use low and high spatial resolution image pairs, consisting of materials of interest, to estimate the wavelet filter coefficients (basis). An initial estimate of SR is obtained by using these filter coefficients and by learning the high frequency details in the wavelet domain. The final solution is obtained using a sparsity based regularization framework in which image degradation and the sparseness of SR are estimated using the estimated low pass filter coefficients and the initial SR estimate, respectively. The advantage of the proposed algorithm lies in (1) the use of estimated filter coefficients to represent an optimal PSF to model image degradation process, (2) use of sparsity prior to preserve neighborhood dependencies in SR image, and (3) avoiding the use of registered images while learning the initial estimate. Experiments are conducted

on HSI of natural scene as well as on real HSI collected by Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Visual inspection and quantitative comparison confirm that our method enhances spatial resolution without introducing considerable spectral distortion.

6.1 Related Work

The problem of SR has been attempted by many researchers since early 1980. Tsai and Huang [20] were the first to suggest resolution improvement of an image using several downsampled noise free images of the same scene. There are many ways to improve spatial resolution of hyperspectral images (HSIs) such as SR reconstruction using a few low-resolution images [33, 34, 5], pan-sharpening fusion using coincident HR image (e.g. PAN) followed by super-resolution mapping (SRM) [159, 55], SRM after spectral unmixing [24, 26, 60], wavelet based methods [22, 87, 27], etc. Depending on the number of LR images involved, the SR method is called multi-frame [33, 34, 5] or single-frame SR [159, 55, 24, 26, 60]. Multi-frame based SR methods use subpixel shifted LR observations of the same scene to obtain SR results, while single-frame approaches learn the detail information from image database that has large number of HR (high-resolution) or LR (Low resolution) and HR training images. An accurate registration of the low-resolution images is critical in multi-frame SR since the method is based on exploiting the non-redundancy available in the subpixel shifted LR observations. When working using remotely sensed images, many times it is difficult to obtain subpixel shifted LR observations of the same scene, specifically for highly dynamic scenes. Therefore, in remote sensing single-frame SR image mapping has become a popular area of research.

Several SR techniques have been proposed based on single-frame super-resolution mapping (SRM) techniques [159, 55, 24, 26, 60]. SRM based techniques exploit the spatial information by making use of coincident HR image (e.g. PAN image) [159, 55] or unmixing model that describes the spatial distribution of the contents of mixed pixels [24, 26, 60]. An algorithm proposed by Foody in [159] uses a simple regression based approach to enhance the spatial resolution of LR HSI using coincident HR image. Improved result is obtained by using an SR mapping technique in which location of landcover classes are predicted by fitting class membership contours that results in reducing the blockiness

in final SR output. The main limitation of this algorithm is the need of secondary HR image coincident with LR test image. Nguyen et al. [55] used fused image as an additional source of information for SRM using a Hopfield neural network (HNN). Need of secondary HR coincident image is the limitation of this algorithm as well. Besides this, algorithms based on HNN require higher computational time. In a different approach Gu et al. [24] proposed an SR algorithm that uses an indirect approach based on spectral mixture analysis (SMA) and the learning based SRM is performed by using backpropagation neural network (BPNN). A set of HR training images unassociated with test image are used for training the BPNN. An advantage of this method is that no supplementary source of information associated with LR test image is required. In a similar work, Mianji et al. [26] used LR and its downsampled version to train the BPNN. Then learning based SRM is performed after SMA by considering spatial correlation of different materials present in the HSI. Villa et al. [60] used spatial regularization by simulated annealing to perform SRM which is performed after coarse classification using support vector machine and SMA steps.

There are considerable number of techniques in which wavelet decomposition is used to increase the spatial resolution of remote sensing images [22, 87, 27]. These methods are based on the decomposition of the image into multiple levels based on their local frequency contents. The wavelet transform decomposes images into a number of new images each having different spatial resolution. Need of the coincident HR auxiliary information is the main limitation of these methods. Besides, all these methods use fixed wavelet basis like Db4 in their implementation and they require accurate co-registration to achieve acceptable results. Mertens et al. [76] proposed use of predicted wavelet coefficients to obtain SR image. They learn relation between approximate and detail coefficients using training data in neural network, without making any assumption about data distribution. In a recent approach, Li et al. [88] characterized the wavelet coefficients by a mixed Gaussian distribution and the dependencies between the coarser and the finer scale wavelet coefficients were modeled as prior by using the universal hidden Markov model and the problem was solved as an maximum a posteriori (MAP) framework. Recently, learning based SR approaches for single wideband and multiband images have been explored by the researchers to solve the super-resolution problem [89, 90, 91, 92, 93, 94, 95, 96, 26, 79, 97] in which high frequency details are obtained using the training data. These methods

use a database of HR images or LR-HR image pairs in order to learn the high frequency details for SR. Use of sparsity as a prior for regularization of ill-posed problems has been validated by many researchers [80, 81, 82, 84, 79, 61].

Most of the earlier research on SR of HSI assumes implicitly or explicitly that each LR pixel of individual spectral band is obtained as a equally weighted sum of pixels of corresponding HR spectral band, and they are perfectly aligned with HR pixels. This means, the point spread function (PSF) of sensor is same over the entire spatial and spectral region, depending only on the position of the detectors. But, in practice, PSF depends on various factors of hyperspectral imager such as fill factor of CCD array, camera gain, zoom factor, imaging wavelength etc. [17]. The effect of diffraction is significant at higher wavelength in a hyperspectral imager. This results in spatially and spectrally varying PSF of the degradation function.

In this chapter, we address the problem of single-frame image super-resolution using learning based approach in wavelet domain, where we obtain high frequency contents from HR training images unassociated with test image. This eliminates the need of registration while obtaining these frequencies. Novelty of our approach lies in estimating the wavelet filter coefficients that takes care of spectrally varying PSF. Here we are not considering spatially varying PSF, which is quite involved as this requires the estimation of PSF at every pixel. The estimated filter coefficients are then used to learn high frequency details in a given band in wavelet domain, obtaining an initial estimate of SR image. The final SR image is obtained using the sparsity based regularization that has the observation model constructed using the estimated filter coefficients.

For the estimation of optimum wavelet filter coefficients, LR-HR pairs of HSIs referred to as the training images can be created in two different ways:

1. by changing the configurations of the hyperspectral imager. For example, optically varying the width of the observed target strip projected onto the sensor's slit facilitates manipulation of the spatial resolution of the system independent of the spectral resolution of the system [160].

2. by changing the height of the platform since the spatial resolution of HSI depends on the platform height. For example, a typical mission, mounting AVIRIS on a NASA aircraft (ER-2), produces a spatial resolution of about 20 meters, but it can

be improved to 5 meters by flying at lower altitudes.

It has to be noted here that acquiring the HR data and LR data is a one time and offline operation. Once a database is created, the LR images captured by hyperspectral imager can be super-resolved using our approach. This is greatly beneficial, as one can capture low spatial resolution HSIs (even though it is capable of capturing HR HSIs) and reduce the memory, transmission bandwidth and power requirements. One may transmit the LR HSIs from the satellites and aircrafts and obtain super-resolved HSIs at the receiver end by using the available database as training images. In order to reduce the computational complexity due to the use of large data base, we use PCA and work with first few principal components only. In our work, filter coefficients are estimated for each of the reduced set of PCA bands. Use of estimated filter coefficients in learning the initial SR as well as in the degradation model incorporates wavelength dependent i.e. spectrally varying PSF while estimating the SR image. Efficacy of the proposed method is tested by conducting experiments on three different data sets, namely single band natural images, a 31-band hyperspectral image of a natural scene captured under controlled illumination and a 224 band AVIRIS remotely sensed data. The results of proposed approach are compared with bicubic interpolation technique [148], learning based SR method of Jiji et al. [161] that uses fixed basis wavelet coefficients Db4, and two recently published SR methods that are based on sparse representation [79, 61].

6.2 Block Diagram Description of the Proposed Approach

The proposed technique of learning based super-resolution of HSI is illustrated in Figure 6.1. Given LR test image and a database of LR-HR HSIs, the proposed technique is implemented using the following steps

1. Form a training database of registered low-resolution and high-resolution HSIs.
2. Reduce dimensionality of the HSIs using PCA. We now have PCA transformed LR-HR images.
3. Estimate the optimum wavelet filter coefficients using a database created in step 2.

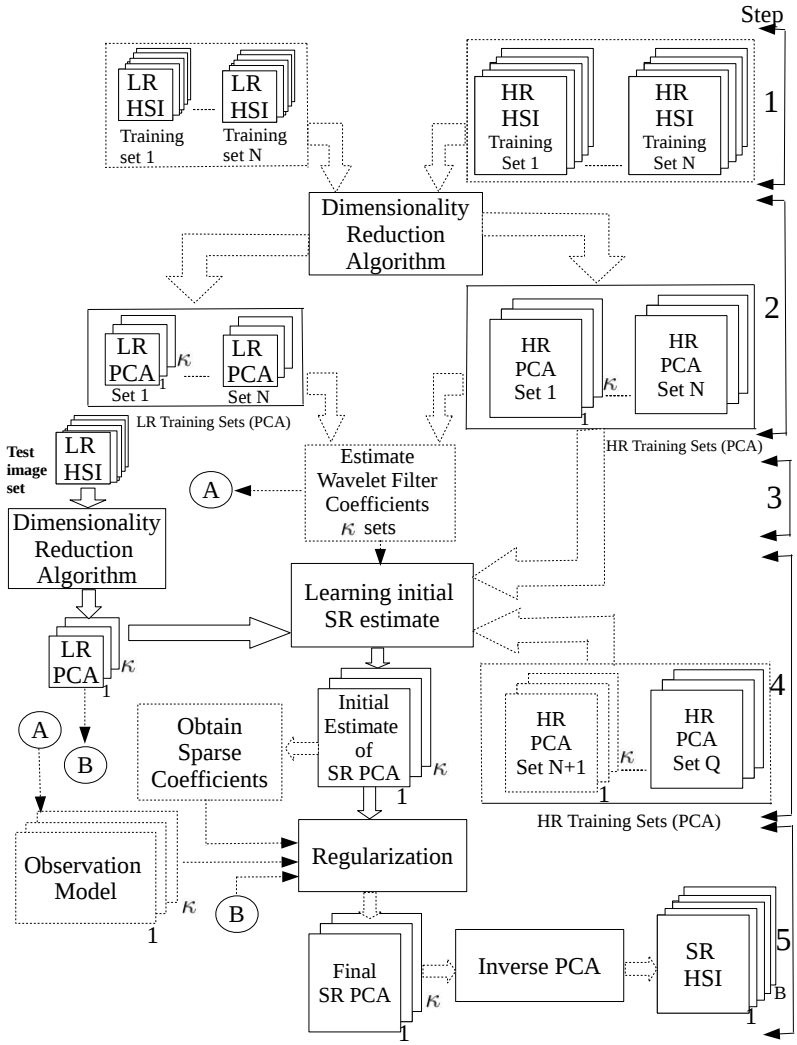


Figure 6.1: Detailed block diagram of proposed approach for HSI super-resolution

4. Use the estimated filter coefficients and obtain the initial estimate of SR using the wavelet transform based learning. Additional HR HSIs are used while learning the initial SR estimate.

5. Use regularization to obtain the final super-resolved image in PCA domain and subsequently the super-resolved HSI is obtained after inverse PCA.

To estimate wavelet filter coefficients to be used in learning high frequency details in step 4, we first need to create database of registered LR-HR pairs of HSIs (which contains materials of interest in sufficient amount) using any one of the approach described in section I. In practice if it is not feasible to capture LR-HR pairs from the imager, one may use only the HR training images and LR images are obtained by simulation. Here we assume that we have access to HR training images to learn the detail information. In case

of nonavailability of LR-HR pairs one can choose an indirect approach to estimate the optimum wavelet filter coefficients from the available training images at one resolution. For example, one can obtain point spread function of the imager, using one of the methods described in [162]. The estimated PSF may then be used to generate the LR images and the resulting LR-HR pairs can be used to estimate wavelet filter coefficients.

Hyperspectral images are composed of large number of spectral bands (e.g., AVIRIS acquires 224 bands). Applying super-resolution technique to each band separately is prohibitive because of high time complexity. Since the spectral content of HSIs are inherently low dimensional, it can be exploited by using PCA, a standard tool for analysis of multivariate data [8]. Hence, we first apply dimensionality reduction on LR and HR training images having B bands in each LR-HR training set using PCA. This transformation incorporates most of the spectral variability of HSI data in first few principal components. We retain only the first κ significant eigen vectors of spectral covariance matrix, corresponding to significant eigen values. Since the number of eigen vectors retained is much less compared to the total number of HSI bands (B), one cannot reconstruct original hyperspectral image exactly by inverse PCA, thus causing information loss. However, it is reasonable to assume that the spectral signature of the materials/objects of interest is present in sufficient amount in reasonable number of spectral bands. Note that the number of PCA components retained (κ) is application dependent and it can be increased at the cost of computational speed, and information loss and the reconstruction error may be made arbitrarily small in order to take care of classification accuracy.

The database of LR-HR images in PCA domain can now be used in wavelet based learning to obtain an initial estimate of SR HSI. However, while using discrete wavelet transform (DWT) it is not a good idea to use the conventional basis such as Haar, Daubechies or Coiflet as they are not optimized over the class of images. This motivates us to estimate the wavelet filter coefficients before learning the initial SR estimate. Using the registered LR-HR PCA data sets of training images, optimum filter coefficients are estimated for each PCA band individually. Thus we obtain a total of κ sets of wavelet filter coefficients for PCA bands 1 to κ. These coefficients are then used in wavelet transform based learning to obtain the initial SR estimate for the κ test (LR) HSIs which are also in the PCA domain. These filter coefficients are also used to define the PSF/degradation in the observation model that is used in regularization to obtain the final SR image.

To this end our method for SR uses adaptive wavelet basis which is optimized for a group of HSIs. One may note that there is no need of registration while learning initial SR estimate as we use only the HR database while obtaining the initial SR estimate. This gives us freedom to include additional HR training HSIs in the database (i.e., $N+1$ to Q) after the dimensionality reduction as shown in Figure 6.1. This inclusion would enhance the accuracy of initial SR estimate. Applying inverse DWT gives us SR initial estimates of LR HSIs in PCA domain. To avoid confusion, we use the word "SR" for algorithmic output only (e.g. initial SR, final SR in Figure 6.1), elsewhere we use the word "HR".

Our method of obtaining initial SR estimate do not consider the contextual dependencies among pixels as it is patch based. This results in artifacts in initial SR image around the patch boundaries. Hence prior knowledge about HSI is utilized in order to obtain better solution. Regularization based on sparsity as prior is performed in order to obtain final solution for each PCA components. As shown in Figure 6.1, observation model and the sparse coefficients are used in regularization. Observation/degradation model is constructed for each LR image using the already estimated κ sets of estimated filter coefficients. A patch based approach is obtaining the sparse coefficients using the initially estimated SR PCA components and they represent the dependence of an SR patch on its nearby patches. Note that though the individual PCA bands are uncorrelated, spatial dependency exists within the pixels of PCA image [163]. Our final cost function being differentiable, a simple optimization technique like gradient descent is used to minimize the same. This results in final SR HSI in the PCA domain. Applying inverse PCA transformation results in super-resolved HSI.

6.3 Estimation of Wavelet Filter Coefficients

For the last two decades discrete wavelet transform has become one of the most important tools in the field of image processing. Many researchers have attempted to increase spatial resolution of natural [161, 164, 165, 166] as well as remotely sensed images [87, 22, 124, 27] using DWT . The limitation of these algorithms is the use of a specific type of wavelet transform where the basis is fixed. That is, they use known filter coefficients and hence do not guarantee the optimum performance. In contrast, in the proposed approach, we derive optimal filter coefficients which are then used in learning the initial SR estimate.

A number of general conditions and unbounded degrees of freedom can be used to adapt wavelets to a desired signal processing application. Here DWT basis coefficients are not explicitly specified, instead they are computed from the signal (images in our work) itself by computing the impulse response coefficients of a particular wavelet filter. As already mentioned this is done on the reduced set, i.e. after performing PCA.

To start with, a database of N pairs of LR and HR HSIs is created, generating κ LR-HR pairs of principal components. We choose LR-HR database of HSIs to represent materials and objects of interest. The estimation of wavelet filter coefficients is carried out for each PCA component separately. We first take the wavelet transform of HR PCA images of training pairs and perform one level and two level wavelet decomposition for the magnification factor of 2 and 4, respectively. To estimate wavelet filter coefficients, we use the fact that the coarser part of wavelet transformed HR PCA image should be close to the LR PCA image in the mean squared sense. This has to be true for all LR-HR pairs in the database as shown in Figure 6.1. If these coefficients are used for initial estimate learning, it represents a better approximation to the SR.

Although numerous wavelet basis are available, there exist difficulty in finding optimal length wavelet filter basis for image super-resolution application. The discrete Haar wavelet transform having filter length of 2 for low pass (LP) and high pass (HP) has the advantage of having smaller filter length and simpler to compute, but it is not continuous, resulting in introduction of blockiness in the learned SR image. Hence we need to use higher order filter with overlapping response to make it continuous, which helps to preserve continuity in SR image. But, increase in the filter length adds to computational burden with no significant improvement in the performance. As a compromise between computational burden and the performance, we have chosen a design length of 4 for the wavelet filter to suite our requirement in SR algorithm and use the necessary conditions to obtain the optimal coefficients. In [167] Daubechies described a family of filters for wavelet transform (WT) computation. For details of WT one may refer to [167, 168]. Here we describe the procedure to estimate the wavelet filter coefficients for one of the PCA bands of HSI. The same procedure is repeated for all κ bands. For example, if there are 3 significant PCA bands corresponding to $\kappa = 3$, then we will have 3 sets of LP as well as HP filter coefficients.

We now explain the mathematical theory for finding our filter coefficients in terms

of a single coefficient and describe how this coefficient can be estimated from the data. Consider an LR PCA image having size of $M \times M$ and corresponding HR PCA image be of size $2M \times 2M$ giving a resolution difference of $q = 2$. We write a system of equations that must be solved to find low pass wavelet filter coefficients $l = (l_0, l_1, l_2, l_3)$. The high pass filter coefficients $h_0, ..., h_3$ can then be determined from $l_0, ..., l_3$. Considering 1-D case, the wavelet transformation of vector of length $2M$ is given by

$$\text{Transformed vector} = W_{2M} \times \text{Input vector}, \tag{6.1}$$

where W_{2M} is the wavelet transformation matrix given by

$$W_{2M} = \begin{bmatrix} l_3 & l_2 & l_1 & l_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & l_3 & l_2 & l_1 & l_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ l_1 & l_0 & 0 & 0 & 0 & 0 & \cdots & l_3 & l_2 \\ h_3 & h_2 & h_1 & h_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & h_3 & h_2 & h_1 & h_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ h_1 & h_0 & 0 & 0 & 0 & 0 & \cdots & h_3 & h_2 \end{bmatrix}_{2M \times 2M.}$$

The upper half of rows in W_{2M} matrix perform the low pass filtering operation and generate coarser part of the signal and the lower half performs high pass filtering operation and generates finer details of the signal. Now consider

$$W_{2M} = \begin{bmatrix} L \\ \overline{H} \end{bmatrix}, \tag{6.2}$$

where L and H are low pass and high pass filter matrices, each of size $M \times 2M$. Considering the fact that the two-dimensional DWT is separable, 2-D transform W_t can be performed in two steps, each of which involves a one-dimensional transform operation.

We can find wavelet transform of 2-D image I of size $2M \times 2M$ as

$$\begin{aligned} W_t &= W_{2M} I W_{2M}^T \\ &= \left[\begin{array}{c|c} LIL^T & LIH^T \\ \hline HIL^T & HIH^T \end{array}\right]. \end{aligned} \tag{6.3}$$

Considering left to right multiplication in equation (6.3), it first computes one dimensional transform along columns of image I (i.e. $W_{2M}I$) and then computes one dimensional transform along rows of image $W_{2M}I$ to obtain W_t. This results in the coarser version of image I as block LIL^T, of size $M \times M$ and the vertical, horizontal and diagonal details of the image I are contained in blocks LIH^T, HIL^T, and HIH^T respectively, each of size $M \times M$. To derive the filter coefficients, we consider LIL^T and design a transformation matrix W_{2M}, which has the desired filter coefficients. W_{2M} is designed as follows. Considering the orthonormality constraints we have

$$\sum_{n=0}^{3} l_n^2 = 1, \tag{6.4}$$

$$\text{where } l_0 l_2 + l_1 l_3 = 0. \tag{6.5}$$

Taking the discrete time Fourier transform (DTFT) of the sequence $l_n, n = 0, 1, 2, 3$ gives us

$$H(\omega) = \sum_{n=0}^{3} l_n e^{jn\omega}. \tag{6.6}$$

Now for low pass response, $H(\omega) = 1$ at $\omega = 0$ and it is 0 at $\omega = \pi$. Using this in equation (6.6) we obtain,

$$H(0) = \sum_{n=0}^{3} l_n = 1 \text{ and} \tag{6.7}$$

$$H(\pi) = \sum_{n=0}^{3} (-1)^n l_n = 0. \tag{6.8}$$

But equation (6.7) violets distance preserving property of orthogonal matrices [169]. To

satisfy orthonormality conditions and the low pass condition we must have,

$$\sum_{n=0}^{3} l_n = \pm\sqrt{2} \text{ and} \tag{6.9}$$

to satisfy equation (6.5) we must have

$$[l_2, l_3]^T = c[-l_1, l_0]^T \text{ for } c \neq 0. \tag{6.10}$$

Equation (6.10) in conjunction with equation (6.8) leads to

$$l_1 = \frac{1-c}{1+c} l_0 \text{ for } c \neq -1. \tag{6.11}$$

Using equations (6.4), (6.10) and (6.11) we obtain the following set of equations.

$$\begin{aligned} l_0 &= \frac{1+c}{\sqrt{2}(1+c^2)} \qquad l_1 = \frac{1-c}{\sqrt{2}(1+c^2)} \\ l_2 &= -\frac{c(1-c)}{\sqrt{2}(1+c^2)} \qquad l_3 = \frac{c(1+c)}{\sqrt{2}(1+c^2)}. \end{aligned} \tag{6.12}$$

Since c can be any real value except -1, we endup with infinite number of solutions. A unique solution for the coefficient c is obtained by solving the following optimization problem. Let $Y_{m1,n1}$ be the LR PCA image in the training set and $B_{m1,n1}$ be the coarser resolution version of the HR PCA images (i.e. LIL^T in equation (6.3)), where $m1, n1 = 1, 2, ..., M$ indicate spatial locations. There are N such LR - HR pairs. We formulate the minimization problem as

$$\epsilon = \arg\min_{\forall c} \sum_{P=1}^{N} \sum_{m1=1}^{M} \sum_{n1=1}^{M} (Y_{m1,n1}^{(P)} - B_{m1,n1}^{(P)})^2, \tag{6.13}$$

where $Y^{(P)}$ is the P^{th} low-resolution PCA training image and $B^{(P)}$ is the coarser part of the wavelet transformed HR PCA training image number P. Here $Y_{m1,n1}^{(P)}$ is known and $B_{m1,n1}^{(P)}$ can be expressed in terms of c. Equation (6.13) is convex, hence simple optimization technique like gradient descent can be used to find optimum value of c, which in turn can be used to determine optimum $l_0, ..., l_3$ (see equation (6.12)). Note that in equation (6.13) $B_{m1,n1}$ is obtained as 2D convolution of each 4×4 block of HR

image with $L = [l_0, l_1, l_2, l_3]$ i.e.,

$$\begin{aligned} B_{m1,n1} &= l_3^2 I_{m,n} + l_2 l_3 I_{m,n+1} + l_1 l_3 I_{m,n+2} + l_0 l_3 I_{m,n+3} + l_3 l_2 I_{m+1,n} + l_2^2 I_{m+1,n+1} \\ &+ l_1 l_2 I_{m+1,n+2} + l_0 l_2 I_{m+1,n+3} + l_1 l_3 I_{m+2,n} + l_1 l_2 I_{m+2,n+1} + l_1^2 I_{m+2,n+2} \\ &+ l_0 l_1 I_{m+2,n+3} + l_0 l_3 I_{m+3,n} + l_0 l_2 I_{m+3,n+1} + l_0 l_1 I_{m+3,n+2} + l_0^2 I_{m+3,n+3}, \end{aligned} \tag{6.14}$$

where $m = 2 \times (m1 - 1) + 1$ and $n = 2 \times (n1 - 1) + 1$.

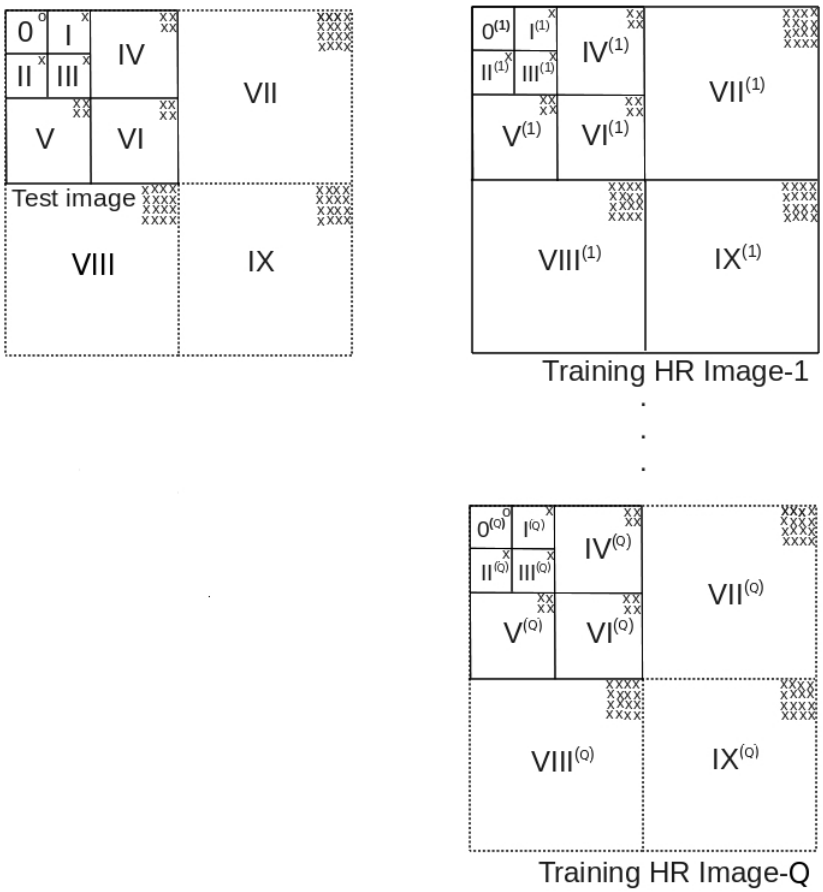
Here $I_{m,n}$ represents the intensity of the HR PCA image pixel at (m, n). For a resolution factor of 4, the above equation can equivalently represented as a convolution of low pass filter coefficients with HR image block of size 16×16, resulting in a total of 256 terms in the right hand side (RHS) of equation (6.14). The high pass filter coefficients can now be obtained, to satisfy orthonormality condition of matrix W_{2M}, as

$$h_0 = l_3; \quad h_1 = -l_2; \quad h_2 = l_1; \quad h_3 = -l_0. \tag{6.15}$$

It may be noted that the derived LP filter coefficients from equation (6.13) are optimal in the mean squared sense i.e. they minimize the square of the error between the LR and coarser part of the wavelet transformed HR PCA images. Therefore, the corresponding high pass filter should yield better edge details in the SR image. These estimated filter coefficients are used in learning the initial SR PCA estimate as well as in constructing the degradation model.

6.4 Learning Initial SR Estimate

The richness of the texture in the real-world images is difficult to derive analytically. Hence learning based approaches work well while obtaining the missing high frequency details. Use of learning based approaches for super-resolving natural as well as remotely sensed images is considered by many of the researchers [89, 90, 91, 92, 93, 94, 95, 96, 26, 137, 97]. In our approach, once the filter coefficients are estimated, our next task is to learn the high frequency details in terms of detail coefficients. The estimated filter coefficients are used in obtaining the detail wavelet coefficients, by taking the wavelet transform of the test image and the training images using these coefficients. This would



(a) (b)

Figure 6.2: Illustration of learning of detail wavelet coefficients for $q = 2$ using a database of HR PCA images. (a) Two level wavelet decomposition of test PCA image (LR observation). Dotted lines show wavelet coefficients to be learned. (b) Three level wavelet decomposition of HR PCA training images.

minimize the error while learning that may result when using the known basis. Our learning uses only HR HSI in training database as opposed to LR-HR image pairs used by few researchers [161, 165]. For learning purposes, the detail wavelet coefficients are learned for a decimation factor of $q = 2$ and $q = 4$, respectively. Considering a decimation factor 2, we use two level and three level wavelet decomposition for the test and training images, respectively. Figure 6.2 shows the block schematic for learning of wavelet coefficients for one of the test image for $q = 2$. Note that the test and training images are in the PCA domain. Figure 6.2(a) shows the subbands 0 to VI of the LR test image, while the dotted lines show subbands $VII - IX$ that have to be learned. Subband 0 represents the coarser part of the DWT transformed image and subbands $I - III$ and $IV - VI$ represent the vertical, horizontal, and diagonal details. Figure 6.2(b) shows three level wavelet

decomposition of HR training images having subbands $0^{(r)} - IX^{(r)}$, $r = 1, 2, ..., Q$. Here subbands $I - III$, $IV - VI$, and $VII - IX$ represent the vertical, horizontal, and diagonal details in level 3, 2 and 1, respectively. To learn the wavelet coefficients for subbands VII of LR test image, we compare the coefficients in subbands I and IV of the LR test image with that in subbands $I^{(r)}$ and $IV^{(r)}$, $r = 1, 2, ..., Q$ of the HR training images and obtain best match coefficient from subbands $VII^{(r)}$ of a training image. Similarly, we learn wavelet coefficients for subbands $VIII$ and IX of LR test image. Learning procedure is described in detail as below.

Consider an LR test image of size $M \times M$ pixels. The corresponding HR image is of size $2M \times 2M$ pixels giving a resolution difference of 2. We have a total of Q HR training images. In order to learn the wavelet coefficients we exploit the idea of zero tree concept, i.e. in a multiresolution system, every coefficient at a given scale can be related to a set of coefficients at the next coarser scale of similar orientation [170]. Using this idea we follow the minimum mean squared error between the known DWT coefficients of test and training images to learn the unknown detail wavelet coefficients. Suppose $\phi(i, j)$ is the wavelet coefficients at location (i, j) in subband 0, where $0 \leq i, j < M/4$ of the LR test image. Corresponding detail coefficients in subbands I, II, and III are at locations $\phi(i, j + M/4)$, $\phi(i + M/4, j)$, and $\phi(i + M/4, j + M/4)$, respectively and the coefficients in subband IV, V, and VI correspond to blocks of size 2×2 $\{\phi(p, q + M/2)_{p=2i,q=2j}^{p=2i+1,q=2j+1}\}$, $\{\phi(p + M/2, q)_{p=2i,q=2j}^{p=2i+1,q=2j+1}\}$, and $\{\phi(p + M/2, q + M/2)_{p=2i,q=2j}^{p=2i+1,q=2j+1}\}$, respectively. These coefficients of $I - VI$ are used in learning the missing 4×4 blocks in subbands $VII - IX$. For a pixel at (i, j) in the test image at subband 0 following minimization is carried out to pick the missing block of size 4×4 in subband VII that gives us the horizontal edge details.

$$\epsilon = \arg\min_{\forall l,m,r} \left[\phi_I(i, j + M/4) - \phi_I^{(r)}(l, m + M/4)\right]^2 + \left[\sum_{p=2i}^{2i+1}\sum_{q=2j}^{2j+1} \phi_{IV}(p, q + M/2) - \sum_{l1=2l}^{2l+1}\sum_{m1=2m}^{2m+1} \phi_{IV}^{(r)}(l1, m1 + M/2)\right]^2, \quad (6.16)$$

where $r = 1, ..., Q$, and $0 \leq l, m < M/4$, and $\phi_I^{(r)}$ and $\phi_{IV}^{(r)}$ denote the wavelet coefficients for the r^{th} training image at I^{st} and IV^{th} subbands. The corresponding wavelet detail coefficients from the subband VII of HR training image are copied into the subband VII

of the test image. This is repeated for every location in the subbands. In this way we obtain

$$\phi_{VII}(s,t1)_{s=4i,t=4j}^{s=i1,t=j1} := \phi_{VII}^{(r)}(s_n,t_n)_{s_n=4l,t_n=4m+M}^{s_n=4l+3,t_n=4m+3+M} \tag{6.17}$$

where $t1 = t+M$, $i1 = 4i+3$, and $j1 = 4j+3$. This way we endup learning the unknown (missing) wavelet coefficients in subbands VII of LR test image. To find the vertical and diagonal details, subscripts of ϕ are changed to (II and V) and (III and VI), respectively, in place of (I and IV), in addition to appropriate displacement of pixel indices by $M/4$ or $M/2$ in equation (6.16). Similarly, subscripts of ϕ are changed to $VIII$ and IX with appropriate displacement of pixel indices by M in equation (6.17). Essentially we are searching for the best matching horizontal, vertical and diagonal detail coefficient blocks separately. By applying inverse wavelet transformation to this learned image gives the initial SR estimate. Similar procedure is used on all PCA images to obtain the initial SR estimate for every test image in PCA domain. If the error (ϵ) is quite large, it signifies that the 4×4 patch does not have its corresponding HR representation in database. To avoid such spurious learning, we consider the DWT coefficients only when the error (ϵ) is less than a chosen threshold. The goodness of the learning depends on how extensive is the training data set. Our database consists of sufficiently large data set in order to avoid large errors.

6.5 Final Solution using Regularization

In the DWT based learning process, the detail coefficients are learned from the training set using block based approach. Thus, spatial correlation is not considered while learning these coefficients. Since we choose the high frequency components of each 4×4 region independently as per the best fit, corresponding SR image lacks any spatial context dependency which may cause an unwanted abrupt variation across the 8×8 blocks when we consider a resolution factor $q = 2$. This necessitates further refinement of the initial estimate to obtain a better solution by using the prior information about the solution. A better solution can be obtained by formulating the problem in regularization framework. For regularization purpose one needs to have a data fitting term and a regularization

term. Hence we first model the image formation in order to obtain the data fitting term.

6.5.1 Observation Model

Let $Y_\beta, (\beta = 1, 2, ..., \kappa)$, be the PCA transformed LR HSI test image of size $M \times M$ and and Z_β be the corresponding SR PCA image of size $qM \times qM$. Assuming a linear model, for image formation, the LR observation $\mathbf{y}_\beta$ can be expressed as

$$\mathbf{y}_\beta = A_\beta \mathbf{z}_\beta + \mathbf{n}_\beta, \tag{6.18}$$

where $\mathbf{y}_\beta$ and $\mathbf{z}_\beta$ represent the lexicographically ordered vectors of size $M^2 \times 1$ and $q^2M^2 \times 1$ respectively, with $\mathbf{z}_\beta$ representing the SR vector. A_β is the degradation matrix of size $M^2 \times q^2M^2$ that takes care of degradation that includes aliasing caused as a result of downsampling. Generally, the degradation matrix used to obtain the aliased pixel intensities from the HR pixels has the form as mentioned in [126] that has q^2 non-zero entries in every row having values of $1/q^2$. Before we move on, we would like to point out that in the earlier research works on SR, matrix with fixed entries [126] was used as a degradation model for all bands of the multispectral and HSIs [23, 13, 25, 61]. This means, an LR pixel is the average of light intensity that falls on the HR pixels assuming that the entire area of a pixel is acting as the light sensing area and fill factor for the CCD array is unity for all spectral bands. In practice this is not true, and incorporation of improved degradation model leads to better solution.

In our approach, we use the estimated LP wavelet filter coefficients $l_0, ..., l_3$ to construct the degradation matrix A_β. Instead of considering LR pixel as the averaging of HR pixels, we represent it as a linear combination of HR pixels i.e., $\mathbf{z}_\beta$ weighted appropriately by using the estimated low pass filter coefficients. In this case for an integer factor of q, matrix A_β consists of q^4 non-zero elements along each row at appropriate locations. For a resolution factor of $q = 2$ the values and locations of each element are determined from equation (6.14), where $I_{m,n}$ and $B_{m1,n1}$ in equation(6.14) correspond to Z_β and Y_β, respectively in equation (6.18). Thus, the LR intensity represents the weighted average of the HR intensities over a neighborhood of q^4 pixels corrupted with additive noise. Here the noise $\mathbf{n}_\beta$ is the independent and identically distributed (i.i.d.) vector with zero mean and variance σ_n^2 and has same size as that of $\mathbf{y}_\beta$. It is important to note that

we estimate LP wavelet filter coefficients for each PCA image separately and hence we are using degradation operator optimized for each band. Doing so, makes the estimated entries of matrix A_β closer to the true values for the chosen model. Since the obtained filter coefficients represent the close approximation to SR. Besides, the model has overlap of HR pixels, horizontally and vertically as seen from equation (6.14). This relaxes the assumption that LR HSI is strictly defined by the specific detector area only.

6.5.2 Sparsity as a Prior

Super-resolution is an ill-posed inverse problem. There are infinite number of solutions to equation (6.18). Hence selection of appropriate model as the prior information and use of regularization helps to obtain better solution. In the field of image processing and computer vision Markov random field (MRF) is the most general model for including the prior information. But the use of MRF often tend to make the solution smooth. This is because MRF is defined on the basis of local dependencies. In recent years, sparse representations of signals have attracted a great deal of attention in signal and image processing researchers. Olshausen and Field [77] proved that a natural image can be represented with a relatively small number of basis functions chosen from over-complete descriptor sets. Compared to methods based on orthonormal transforms or direct time domain processing, sparse representation usually offers better alternate for efficient signal modeling [78]. This motivates us to use sparsity based regularization. The use of sparsity as a prior for SR is explored by many researchers [79, 80, 81, 82]. They use trained dictionaries of HR and LR patches and assume same sparsity for LR and HR patches in order to obtain SR image. Here, sparse coefficients for each LR patch are found using trained LR patch dictionary, and these coefficients are used in generating the HR output. One may use sparsity constraint by directly imposing the condition that the solution should be sparse [80]. In such cases we need to know the sparseness of SR or it has to be learned using dictionary training. In our work we do not require any kind of dictionary training since we are using the dictionary constructed from the initial estimate itself. The sparsity in our work is imposed in a different way as follows.

In our work the sparseness is represented by the weights of the SR patches when a particular SR patch is represented as a linear combination of other patches. Since our

objective is to preserve spatial correlation, we consider that a patch in SR can be represented as a sparse linear combination of the other patches, mostly nearby. By imposing the condition that final solution should have the same sparseness as the groundtruth, we obtain an SR solution that preserves the spatial dependencies. But to know the true sparse coefficients we need the groundtruth which is not available since it has to be estimated. However, we do have the close approximation to SR in the form of initial estimate and we have made use of the same in obtaining the necessary sparse coefficients.

Suppose, $\hat{D}_{H_p} \in R^{n \times K}$represents an over-complete dictionary of K atoms ($K \gg n$) formed by considering the patches/blocks in the initially estimated SR PCA image represented as vectors. Let $\mathbf{x_p} \in R^K$ be the sparse approximation over this dictionary. The atoms represent the lexicographically ordered patches in the initial estimate. Then a measurement vector $\hat{\mathbf{z}}_\mathbf{p} \in R^n$, a patch of initial estimate, can be represented as a linear combination of a few atoms from the dictionary $\hat{D}_{H_p}$ i.e., sparse linear combination of other patches in the image. Thus, $\hat{\mathbf{z}}_\mathbf{p}$ can be written as $\hat{\mathbf{z}}_\mathbf{p} = \hat{D}_{H_p}\mathbf{x_p}$, where $\mathbf{x_p}$ has very few nonzero entries i.e. $\mathbf{x}_p$ is sparse. Note that $\hat{D_{H_p}}$ has column vectors excluding the patch under consideration i.e. $\hat{\mathbf{z_p}}$. Given $\hat{\mathbf{z_p}}$ and $\hat{D_{H_p}}$, $\mathbf{x_p}$ can be obtained by solving the l_1 minimization using standard optimization tool such as linear programming, by posing the problem as

$$\min_{\mathbf{x_p} \in R^K} ||\mathbf{x_p}||_{l_1} \text{subject to } \hat{\mathbf{z}}_\mathbf{p} = \hat{D}_{H_p}\mathbf{x_p}, \tag{6.19}$$

$$\text{where } ||\mathbf{x_p}||_{l_1} = \sum_{i=1}^{K} |x_{p_i}|.$$

Using the above formulation we find the sparse coefficients for every patch of initial estimate in terms of other patches in the image, for the considered PCA band. To do this, the initial estimate is divided into patches of size 10×10 and sparse coefficients are found for each patch, in terms of dictionary atoms (100×1) formed from the same image excluding the patch under consideration. For example if our initial estimate has a size of 128×128 then each atom is of size 100×1 and there are $169 - 1 = 168$ atoms in dictionary. For simplicity, the example is given for non-overlapped patches, though the overlap is considered while estimating the sparseness and hence final SR. We exploit the fact that the pixel intensities do not vary much within a local neighborhood i.e., neighboring patches

are correlated, exploiting the contextual dependency. In order to remove unwanted abrupt variations across the patches of SR image, we considered patches overlapped horizontally and vertically by 2 pixel rows and columns, respectively while forming the dictionary. Thus all but boundary patches of the image are overlapped with 8 neighboring patches while the boundary patches are overlapped by 5 or 3 neighboring patches depending on whether they belong to border or corner. This results in maintaining spatial dependencies among patches. As an example, for an image sizing of 128×128, we obtain a total of $K+1 = 256$ patches each of size 100×1. Border pixels of last two rows and columns are reflected in order to meet the size requirement of patches.

6.5.3 Regularization with Sparsity Coefficients

Considering one of the PCA bands, regularization is carried out as follows. Sparse coefficients obtained for every patch from the initial SR image serve as weights for the final SR image which is now estimated as the unknown atoms of a dictionary. The regularization is obtained using a patch based approach which provides the constraint of sparsity in the final solution. Using a data-fitting term, and sparsity prior term, the cost function for a single patch of PCA band β can be written as,

$$\epsilon_\beta = \underset{\forall \mathbf{z_p}}{\operatorname{argmin}} ||\mathbf{y_p} - A_p \mathbf{z_p}||_2^{\,2} + \lambda ||\mathbf{y_p} - A_p D_{H_p} \mathbf{x_p}||_2^2, \tag{6.20}$$

where $\mathbf{y_p}$ is the LR test patch, $\mathbf{x_p}$ is sparse coefficient vector which is already estimated, $\mathbf{z_p}$ is the SR patch to be estimated, A_p is degradation matrix taking care of aliasing. Here D_{H_p} is the dictionary of SR atoms that has to be estimated and λ represents the weightage given to the sparsity term, chosen empirically. Considering overlapped SR patches, equation (6.20) is constructed for each patch and the final cost consists of sum of these. Note that for an LR test patch of size 4×4 (i.e. 16×1) we consider HR patch size of 10×10 (i.e. 100×1), instead of 8×8. This is because of the consideration of overlapped patches while constructing dictionary. This results in maintaining correlatedness among SR patches while avoiding abrupt variations at patch boundaries as mentioned in section 6.5.2.

For an image size of 128×128, with a patch size of 10×10 (vector of 100×1) having overlapping of 2 pixels in horizontal and vertical directions, we have 16 patches

in each direction, giving us a total of 256 patches. Then the dictionary D_{H_p} will be of size 100×255 and the degradation matrix A_p is of 16×100. Here A_p consists of q^4 non-zero elements along each row, whose values and locations are determined using the estimated low pass wavelet filter coefficients for each PCA band separately as per the equation (6.14), where $I_{m,n}$ represents Z_β and $B_{m1,n1}$ corresponds to Y_β. Note that all the patches are processed simultaneously in order to obtain the final solution.

The above cost function is convex. Hence it can be minimized using a simple optimization technique like gradient descent. In order to provide good initial guess and to speedup the convergence, the learned initial estimate is used as initialization. It may be noted that since we are regularizing the PCA transformed images, we expect a better spectral consistency in final solution. Inverse PCA gives us the SR image in spatial domain.

Before we proceed to the discussion on experimental results, we briefly explain the number of comparisons required in learning the initial estimate in our approach as this step adds to overall computational complexity of our approach. Proposed approach is divided into three steps: (i) estimation of optimum wavelet filter coefficients (ii) learning high frequency wavelet coefficients in order to obtain the initial estimate, and (iii) regularization. Estimation of wavelet filter coefficients is one time offline procedure, hence we have not considered it in complexity analysis. Regularization was carried out by using the available code [171], so we do not include it in this discussion. However, we discuss in detail the number of computations required for learning the high frequency wavelet coefficients, since it involves significant number of comparisons. It involves finding the number of comparisons required to obtain the detail wavelet coefficients when using a database of Q HR training images. For a test image of size $M \times M$ we decompose the same into W levels of wavelet transform, resulting in $3W$ subbands corresponding to horizontal (H), vertical (V), and diagonal (D) details and one subband having the coarser information. We need to learn detail coefficients for each of the coefficients in the coarse subband. For each of the coefficients in this subband, the best matching coefficients at finer level in the training database can be searched by comparing $\sum_{u=0}^{u=W-1} 2^{2u}$ coefficients within each of the detail subbands. Considering subband corresponding to horizontal details, we search for the best matching wavelet coefficients in the entire horizontal subband of all the HR training images. The size of the subband at W^{th} level is $M/2^W \times M/2^W$ and

it consists of $(M/2^W)^2$ coefficients. These coefficients are compared with corresponding horizontal coefficients at all locations $((M/2^W)^2)$ in each of the HR training images (Q) in the database. Thus, the number of comparisons required for learning horizontal details is $Q(M/2^W)^4 \sum_{u=0}^{u=W-1} 2^{2u}$. Similar comparisons are used to find detail coefficients of vertical and diagonal subbands. Hence the total number of comparisons required for learning all detail coefficients amounts to $3Q(M/2^W)^4 \sum_{u=0}^{u=W-1} 2^{2u}$. In our experiment of the single band image, we use two level decomposition (W=2) of the test image of size 64×64 and learn detail coefficients using the database of 100 training images (Q=100). In this case, the number of comparisons required are $3 \times 100 \times (64/2^2)^4 \times (2^0 + 2^2)$. We would like to mention here that, although this involves significant number of comparisons of 98.304×10^4, it will not cause computational burden in our case due to the use of high performance computer and because of the process being non iterative.

6.6 Experiments and Result Analysis

In this section we show the effectiveness of the proposed method by conducting experiments on different data sets. Experiments are carried out on : (1) Single band natural images (2) Natural hyperspectral images, and (3) Remotely sensed HSIs (AVIRIS). Due to the lack of availability of the true LR-HR pairs of hyperspectral images, as a simple sanity check, the proposed SR approach is first tested on a single but wide band natural images. Data for this experiment include three sets of size 64×64, 128×128 and 256×256, captured by computer controlled camera. These data sets are used to test the effectiveness of our method in estimating the wavelet filter coefficients and learning of the high frequency details. The data sets used in the experiments on hyperspectral data constitute images with 31 and 224 spectral bands, respectively. Detailed analysis of the results is performed on 224 band AVIRIS HSI. We performed experimentations for $q = 2$ as well as for $q = 4$. Due to the space constraint, we are demonstrating results only for the case of $q = 4$. For all our experiments, the step size for gradient descent algorithm was chosen as 0.01. The weightage to sparsity term was set as $\lambda = 0.23$ in the regularization equation (6.20) through a trial-and-error procedure for all experiments. Here, one can use a generalized cross-validation technique [156] to identify the optimum value of λ, but it is computationally expensive and it is specific to a given image only.

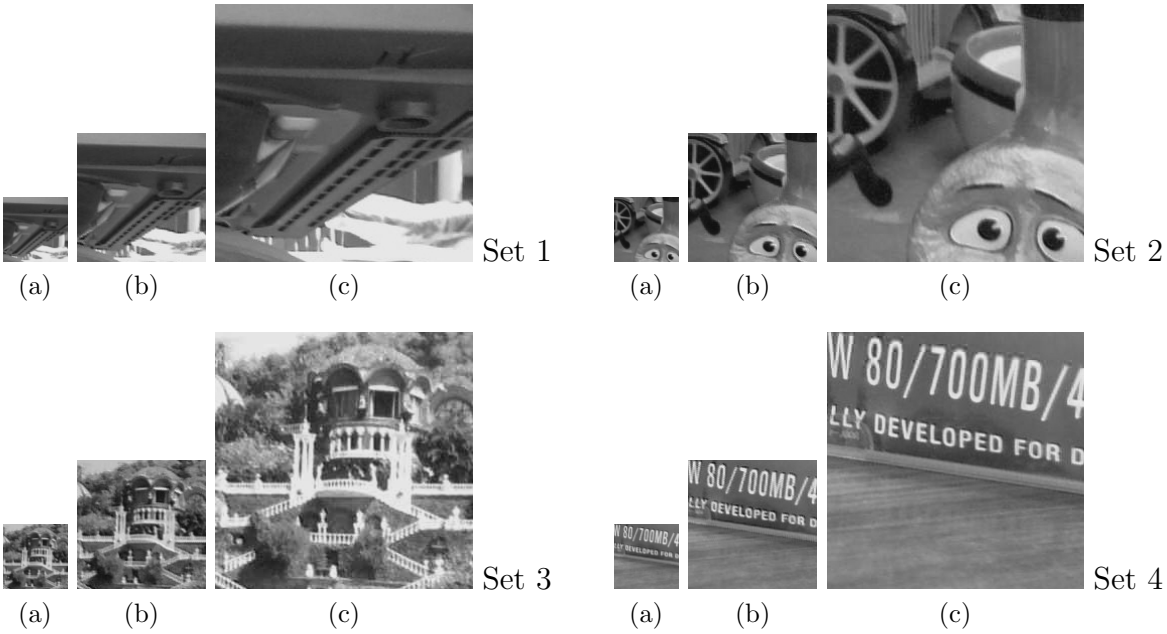


Figure 6.3: Randomly selected training image sets from the database. (a) LR images 64×64, (b) HR images 128×128 (for $q = 2$), and (c) HR images 256×256 (for $q = 4$)

We show the visual as well as quantitative comparison for experiments on all the three data sets. For quantifying the results on single band and 31 band natural HSI we used mean squared error (MSE) as a preliminary evaluation index. Definition of this measure is given in chapter 3 section 3.3.1 (see equation 3.15). For remotely sensed hyperspectral data, detailed quantitative evaluation of spatial and spectral fidelity is performed using different measures such as erreur relative globale adimensionnelle de synthese (ERGAS), spectral angle mapper (SAM), and $Q2^n$. Brief review of ERGAS and SAM is given in chapter 4 section 4.4.1 (see equations 4.9 and 4.8), while that of $Q2^n$ is given in chapter 5 section 5.5.1 (see equation 5.7 and 5.8).

Table 6.1: Performance comparison showing importance of initial estimate on single band "Ganapati" image for $q = 4$ in terms of MSE between groundtruth and initial SR, and groundtruth and final SR

Estimate	Bicubic interpolation as initial SR	Learned SR as initial SR
Initial SR	0.0063	0.0046
Final SR	0.0055	0.0032
Reference	0.0000	0.0000

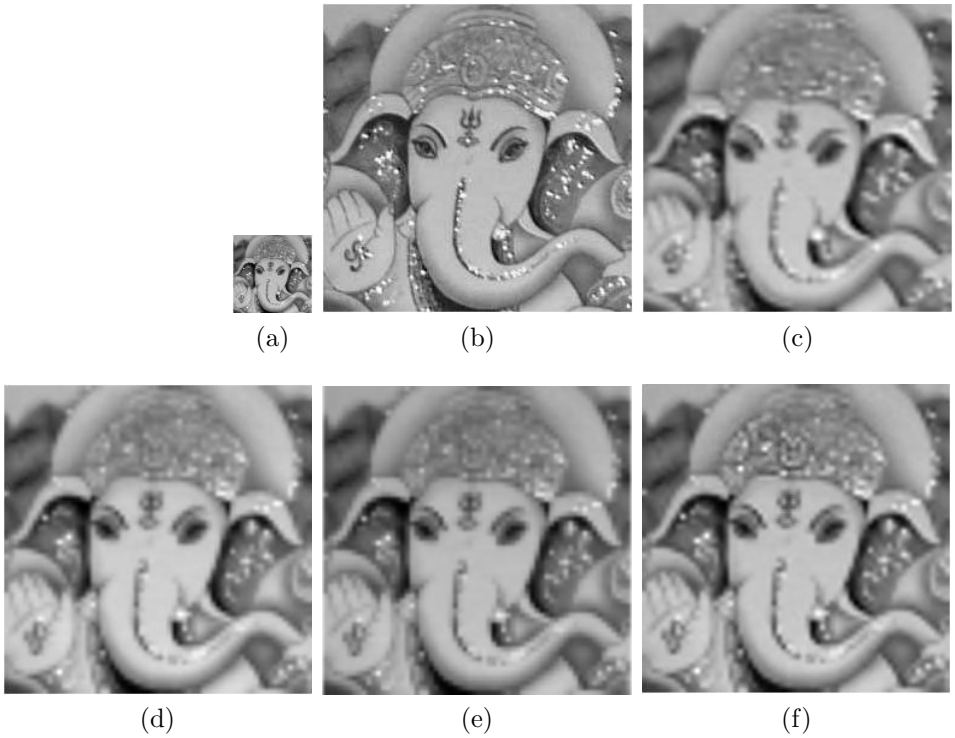


Figure 6.4: SR results for $q = 4$ showing importance of initial estimate. (a) LR image of size 64×64, (b) Ground truth image of size 256×256, (c) Bicubic interpolated image, (d) Learned SR image using estimated wavelet filter, (e) Regularization result when using bicubic interpolated image as initial SR estimate, and (f) Regularization result when using learned SR image as initial SR estimate (Proposed approach)

6.6.1 Experiments on Single band Natural Images

Here we used LR-HR grayscale image pairs captured by varying the optical zoom setting of a simple low cost camera. PCA is not used in this experiment as all the images correspond to single band only. Our database consist of LR images of size 64×64 and the HR images of size 128×128 and 256×256, respectively. An LR image of size 64×64 available in the database is used as a test image and the HR images in the database are used to obtain SR for $q = 2$ and $q = 4$, respectively. Note that the true HR of test image is not used in the experiments. They are used for computing the quantitative measure and for visual comparison only. Our database consist of 100 images of different scenes each having three different resolutions, that include indoor as well as outdoor scenes captured at different times. This results in a total of $100 \times 3 = 300$ images in the database. All the scenes are real world images captured by a computer controlled camera. To capture these images, a stable and isolated physical setup was used. Images were captured by triggering the camera using MATLAB program. The time difference of less than one

Table 6.2: Estimation of wavelet filter coefficients and comparison of MSE between true LR and reconstructed coarser images for single band images

Image	q	Estimated c	MSE using Db4 coefficients [161]	MSE using the derived filter coefficients
Car	2	0.13	0.0099	0.0077
Text	2	0.22	0.0087	0.0081
Car	4	−0.01	0.0023	0.0015
Text	4	0.02	0.0039	0.0024

Table 6.3: Mean squared error comparison for SR results on single band images for $q = 4$

Image	Bicubic interpolation [148]	Yang et al. [79]	Initial SR (Db4) Jiji et al. [161]	Initial SR (Proposed)
Car	0.0079	0.0064	0.0067	0.0058
Text	0.0097	0.0078	0.0088	0.0074

millisecond was kept between two successive triggers. Mean correction was applied on the captured images to compensate for the illumination variations.

Figure 6.3 shows randomly selected training images in the database. First of all, to demonstrate the effectiveness of initial SR estimate in obtaining the final SR, we show the regularization results obtained by using two different images as initial SR estimates. Figure 6.4(a) and (b) show LR and groundtruth images of Lord "Ganapati". Figure 6.4(c) shows the upsampled image obtained using the bicubic interpolation while Figure 6.4(d) displays learned initial SR estimate obtained using the proposed approach. When we use the bicubically interpolated image as initial SR estimate to learn the sparsity coefficients and perform regularization we obtain the image shown in Figure 6.4(e). Similarly use of Figure 6.4(d) as the initial SR estimate (which is more closer to groundtruth image) in regularization, we obtain the result shown in Figure 6.4(f). Comparing the images in Figure 6.4(e) and (f), it is clearly observed that the SR image of the proposed method has better details when compared to SR with bicubic interpolated image used as initial estimate. We can see that the ψ shape on the forehead of "Ganapati" is clearly visible in Figure 6.4(f) when compared to that shown in Figure 6.4(e). This may be because of better sparseness obtained using the proposed approach for initial estimate. The benefit of using the learning in SR is also evident from the MSE values listed in Table 6.1. Observe that MSE is closer to reference value of 0, when we use the learned SR image as

initial estimate when compared to the bicubic interpolation.

Now, we analyze the results on the usefulness of filter coefficients estimation. For estimating the filter coefficients we make use of images shown in Figure 6.3 (a) and (b) for $q = 2$ and use images in Figure 6.3 (a) and (c) for $q = 4$. We selected 10 LR-HR image pairs having similar kinds of edge and texture details to estimate the wavelet filter coefficients. Table 6.2 shows the estimated values of c for Car and Text images for $q = 2$ and 4, respectively. Note that c gives us the optimum values of LP and HP filter coefficients as in equations (6.12) and (6.15), respectively. The MSE between the true LR and the coarser part reconstructed using the estimated wavelet filter coefficients and the standard Db4 wavelet for $q = 2$ and $q = 4$ are also given in Table 6.2. From this table we can see that as the resolution factor is changed the values of estimated filter coefficients are also changing and they are also dependent on the image. The MSE for the estimated filter coefficients is less when compared to using the fixed filter coefficients, indicating that estimating the filter coefficients has the advantage.

Finally, we discuss the results on SR for natural images. Figure 6.5 shows the SR results for the experimentation on Car and Text images, respectively for a resolution factor of 4. Figures 6.5(a) and (b) show the LR test images, and groundtruth images, respectively. Images displayed in Figure in 6.5(c) correspond to those expanded using bicubic interpolation. Figure 6.5(d) shows SR images obtained using Yang et al. method [79]. To show the effectiveness of estimated wavelet filter coefficients when compared to fixed Db4 wavelet, we show the initial SR obtained, using Jiji et al. [161] in Figure 6.5(e) and that obtained using the proposed method in Figure 6.5(f). We mention here that learning based method proposed by Jiji et al. [161] uses initial SR obtained using the fixed wavelet basis.

It can be observed that the fading appears in spokes of the wheel in image of Figure 6.5(c) and the blockiness is also visible in the spokes in Figures 6.5(d) and (e). The initial SR obtained using estimated wavelet filter coefficients compares well with the groundtruth as can be observed from Figure 6.5(f). Spreading and blockiness of characters are reduced considerably in the initial SR estimate of proposed approach. This is expected since we learn the wavelet filter coefficients as well as the high frequency details. Note that the artifacts are significantly reduced in the SR images of Figure 6.5(f) which shows the effectiveness of estimated wavelets. The quantitative comparison of these results is

shown in Table 6.3. It is clearly observed that the MSE between the true and the SR using the estimated wavelet in proposed method is significantly less when compared to bicubic interpolation, Yang et al. method [79], and initial SR obtained using fixed basis Db4 wavelet in Jiji et al. method [161].

Table 6.4: Effect of different PSFs on estimation of wavelet filter coefficients for 31-band Natural hyperspectral image for $q = 4$ (Here $\kappa = 3$)

Band	Filtering kernel used	Estimated c	MSE using Db4 coefficients [161]	MSE using estimated coefficients
PCA-1	NN	−0.22	0.0014	0.0012
PCA-2	NN	−0.17	0.0053	0.0044
PCA-3	NN	−0.18	0.0254	0.0230
PCA-1	Gaussian*	−0.37	0.0020	0.0012
PCA-2	Gaussian*	−0.29	0.0033	0.0031
PCA-3	Gaussian*	−0.32	0.0124	0.0117
PCA-1	Gaussian**	−0.19	0.0025	0.0017
PCA-2	Gaussian**	−0.15	0.0222	0.0199
PCA-3	Gaussian**	−0.17	0.0351	0.0295
PCA-1	Motion[1]	−0.20	0.0034	0.0030
PCA-2	Motion[1]	−0.23	0.0912	0.0901
PCA-3	Motion[1]	−0.19	0.0773	0.0741
PCA-1	Motion[2]	−0.26	0.0021	0.0016
PCA-2	Motion[2]	−0.24	0.0686	0.0674
PCA-3	Motion[2]	−0.29	0.0595	0.0561
	* Std. deviation 0.5	**Std. deviation 0.8	[1]No rotation	[2]30 Degree rotation

6.6.2 Experiments on Hyperspectral Data Sets

For hyperspectral images, the experiments are conducted on two different data sets. The first set consists of 31-band reflectance images of natural scene, having spectral range of $400nm - 700nm$ all acquired under the direct sunlight in clear or almost clear sky [149]. HSI cubes available in [172] and [173] are used in this experiment. Cropped region of "Scene 5" of hyperspectral images of natural scenes 2002 [172] is used as test data. Our second HSI data set is comprised of 224-bands of AVIRIS HSI cubes available in [174]. In this case cropped region of an urban area in Moffett Field is used as test data. After discarding few bands having low signal to noise ratio (SNR), 196 bands were used for super-resolving by a factor of 2 and 4, respectively. The band removal was based on

Table 6.5: Quantitative evaluation measures for SR of 31-band Natural hyperspectral image using different techniques for $q = 4$

Quantitative Measures	Bicubic interpolation [148]	Jiji et al. [161]	Zhao et al. [61]	Yang et al. [79]	Proposed approach
MSE-PCA-1	0.0085	0.0067	0.0071	0.0069	0.0019
MSE-PCA-2	0.0116	0.0098	0.0126	0.0128	0.0029
MSE-PCA-3	0.0429	0.0102	0.0122	0.0196	0.0055
ERGAS	3.7358	2.7877	3.4723	4.3172	2.4568
SAM(Deg)	6.3577	4.9012	4.8340	6.0791	4.6200
$Q2^n$	0.9610	0.9774	0.9662	0.9679	0.9810

visual inspection of the images. Care was taken to include bands having spectral range in accordance to that of the test image while creating training data.

The above data sets have high spatial dimensions and hence specific regions are cropped from them and experiments are carried out on the cropped regions. Here, we do not have the true LR-HR pairs of HSIs. Hence the low spatial resolution (LR) images were created from these cropped images by using filtering and downsampling operations. Remaining regions of original as well as other HSI cubes are used to generate training data sets. We used 5 sets of generated LR-HR training pairs cropped from the same HSI cube excluding the test image cube for estimating the filter coefficients. Use of same cube to create training LR-HR pairs ensures inclusion of large number of materials and objects of interest. Additional 15 HR training HSIs were included while estimating the initial HR for both the experiments on HSIs. These images are different for natural HSI and AVIRIS HSI. In order to evaluate the performance of our approach using quantitative measures, we need the groundtruth images. Since these images are not available, we consider original cropped HSIs of size 256×256 as ground truths and generated the LR HSIs of size 128×128 and 64×64 by applying downsampling operation by a factor of $q = 2$ and $q = 4$, respectively. The SR algorithm was then applied on these LR HSIs.

Each set in our training database has one LR image, an HR image with $q = 2$ and an HR image with $q = 4$. In order to restrict the maximum spatial frequency in the image we use low pass filtering operation before downsampling. The low pass filtering operation was performed and tested using five different kernels namely, nearest neighbor (NN), Gaussian filter with standard deviations of 0.5 and 0.8, horizontal motion blur (5 pixels) with no rotation, and with 30 degree rotation. For this purpose we used filter

masks of size 5×5. Table 6.4 shows the estimated values of c and LR reconstruction error (i.e., error between true LR and LR obtained using the filter coefficients derived from c) for PCA bands 1 to 3 for 5 different masks for the 31-band natural HSI. We can see that as we change the filter mask, the estimated filter coefficients also change, thus making it clear that the degradation operation plays significant role while super-resolving. One can see that the values of MSE are significantly reduced when using the estimated filter coefficients. Low value of MSE shows that the use of estimated filter coefficients for SR purposes is better since in general the imaging operation may be modeled in different ways.

Table 6.6: Quantitative evaluation metrics of AVIRIS SR for $q = 4$

	ERGAS	SAM (Deg)	$Q2^n$
Bicubic interpolation [148]	4.9550	8.3634	0.9367
Jiji et al. [161]	2.8108	6.9554	0.9734
Zhao et al. [61]	4.8694	7.2628	0.9438
Yang et al. [79]	3.1952	7.8997	0.9665
Proposed approach	2.5349	6.5849	0.9773
Reference	0.0000	0.0000	1.0000

We next consider the visual and quantitative assessment of SR using the hyperspectral images. Since Gaussian filter was performing better in terms of MSE, we used it prior to downsampling operation in order to generate LR hyperspectral images. Results of the proposed algorithm on 31-band natural HSI is presented in Figure 6.6 for $q = 4$. After taking the PCA we retained 3 images corresponding to the principal components with highest variance and applied the SR algorithm on them. Figure 6.6 displays the results for the PCA-I. Figure 6.6(a) and (b) display the LR test image and the ground truth image, respectively. The SR results obtained using different methods are shown in Figures 6.6(c-g). SR obtained using sparsity based regularization of the proposed approach shown in Figure 6.6(g) has sharper borders of white colored table on the top left corner. Visual inspection of images in Figures 6.6(c-f) indicate that these borders appear blurred in the bicubic interpolated image, SR image using Jiji et al. method [161], Zhao et al. method [61] and the method proposed by Yang et al. [79]. One can see that the SR image obtained using the proposed method displayed in Figure 6.6(g) compares well with the groundtruth. The letters in the text written on the ball are more clear in Figure 6.6(g). The method proposed by Zhao et al.[61] does not use dimensionality reduction

algorithm such as PCA. Hence to display the results, PCA is applied on the SR result and the first three PCA components are retained for comparison purpose only. Quantitative comparisons for this experiment are given in Table 6.5. From the table we can see that the MSEs between the true and the estimated SR PCA components are significantly less for the proposed method. The use of estimated filter coefficients and regularization improves the results in our approach as evident from quantitative evaluation measures such as ERGAS, SAM and $Q2^n$. These measures show that proposed approach better preserves spatial and spectral fidelity in the super-resolved images.

In Figure 6.7 the SR results on remotely sensed data acquired using AVIRIS hyperspectral imager is shown for a specific band 100. Results are listed for the first 3 PCA bands that include 99.3% of spectral variability of HSI. Figure 6.7(a) shows the LR test image and the original HSI band is displayed in Figure 6.7(b). From Figure 6.7(c) we can see that bicubic interpolation blurs the image when upsampled and the high frequency spatial details are lost. One may notice that bicubic interpolation in Figure 6.6(c) appears better than the result in Figure 6.7(c). This is because there is significant high frequency content in AVIRIS data when compared to natural HSI displayed in Figure 6.6 and the interpolation fails to preserve the high frequency details indicating that the interpolation techniques are not suitable for solving the SR problem. Quantitative comparisons given in Table 6.5 and Table 6.6 further proves this observation. Result of Zhao et al.[61] shown in Figure 6.7(e) is less blurred compared to bicubic interpolated image but it shows artifacts and has loss of high frequency details. Jiji et al. [161] approach results in improved visual quality than the interpolation and the method proposed by Zhao et al.[61], but the overall contrast of the image is not preserved as seen from Figure 6.7(d). Sparsity based SR result of Yang et al. [79] method shown in Figure 6.7(f) is visually better than SR obtained using Zhao et al. [61] (see Figure 6.7(e)), but it fails to preserve high frequency details as evident from vertical lines in the middle region of the image.

As seen from Figure 6.7(g), the use of estimated filter coefficients and sparsity regularization results in reduced artifacts and also takes care of preservation of high frequency details. It gives better visual quality closely resembling the groundtruth. We can see that the white patches visible in LR observation appear grayish in Jiji et al. method (See Figure 6.7(d)), but the result is improved in the proposed approach. One can clearly

discriminate the vertical lines appearing in the mid region of the image in Figure 6.7(g) indicating that edge details are very well preserved in the proposed approach.

As far as the quantitative comparison is concerned, it is clear from Table 6.6 that the proposed method provides scores that are closer to the reference values when compared to bicubic interpolation [148], Jiji et al. method [161], Zhao et al. approach [61] and Yang et al. approach [79]. Lower value of ERGAS in the proposed method indicates lesser global distortion in super-resolved HSI. Generally a value of ERGAS below 3 is believed to be an image with good quality [124]. We see that when compared to other approaches the remaining measures such as SAM, $Q2^n$ are better for the proposed method. Lower value of $Q2^n$ in Table 6.6 indicates minimum spatial as well as spectral distortions by the proposed approach. To further support the performance improvement in terms of spectral fidelity using our approach, we show the SAM error plot in Figure 6.8. The plot depicts the total count of pixels in super-resolved image having specific degree of spectral angle error. We can see that when compared to other methods, the proposed approach has maximum number of pixels having spectral distortion less than 5 degrees. Also the number of pixels having spectral error more than 27 degrees is quite less. Lower values for SAM indicate that the proposed method provides better spectral fidelity.

We would like to point out that compared to the SR method discussed in previous chapters, here the use of estimated wavelet filter coefficients in initial estimate as well as in the decimation matrix entries improves the performance of proposed method. It is to be noted that in the present method we use point spread function (PSF)/degradation with overlapping entries, while in previous method non-overlapping entries were considered to define PSF. Improvement in the performance of this approach is evident from the quantitative measure ERGAS given in Table 6.6 and Table 5.2. The value of ERGAS is 2.5349 for the approach proposed in this chapter while it is 2.9725 for the approach based on learned dictionaries and Gabor priors as discussed in chapter 5. We also found that as far as other quantitative measures are concerned the proposed approach performs better. Thus this method outperforms the other previous methods of super-resolution discussed in this thesis.

Finally, we discuss about the timing complexity of our method considering a resolution factor of 4 for SR of HSIs. All algorithms were implemented using MATLAB 7.0 on Intel(R) Core(TM) i3 CPU M380 having operating frequency of 2.53 GHz. Comparisons

of the running time of all the methods for different input images is given in Tabel 6.7 for each experiment. First two rows correspond to computation times of single band images, while third and fourth rows correspond to 31-band and 196-band HSIs. It is to be noted that in Zhao et al. [61] approach spectral regularization step is suitable only for HSI, hence experiments are not performed on Car and Text images and corresponding entities in Table 6.7 are left blank. For the proposed method, the time required for estimating the filter coefficients was 15 seconds which is a one time offline procedure. The time taken for detail coefficients learning i.e., initial SR estimate was only 4 seconds. Note that as already explained in section 6.5, this step involved significant number of comparisons and the use of high speed processor reduced the computing burden. Finally, regularization and inverse PCA was done in 27 seconds. Thus the total time required was about 46 seconds to obtain the final SR image as indicated in the table. From the table we see that time taken for SR of 31-band and 196-band HSIs is significantly higher than that of car and text images. This is because here the total time corresponds to super-resolution of more than one image. The difference in computation time for 31-band and 196-band HSIs in the proposed approach is due to the execution time difference in PCA and inverse PCA. We see that the execution time of bicubic interpolation is much less when compared to other methods. Unfortunately, interpolation fails to retain high frequency details, hence it is not considered as SR technique. Time complexity to estimate SR image in Jiji et al. [161] approach is several hours. This is because the cost function used in their regularization is non-differentiable and hence the use of simulated annealing for cost minimization increases the time complexity. But the time for obtaining the initial estimate in their approach is less when compared to our approach because they do not estimate the filter coefficients. The method proposed by Zhao et al. [61] and Yang et al. [79] are quite expensive in terms of time requirements since the dictionary learning employed in both these approaches takes few hours. Besides, the method of Zhao et al. [61] super-resolve each band separately without using any dimensionality reduction. Hence, time complexity increases as the number of bands increases. For AVIRIS HSI their algorithm takes several hours. This is because the use of sparsity based regularization which is computationally expensive adds to their computational complexity. Another reason for increase in time is due to regularization done separately on each band. However, in the proposed approach computation of sparse coefficients is one time operation and

Table 6.7: Computational time in seconds of different algorithms for LR image size 64×64 and $q = 4$

Image	Bicubic interpolation [148]	Jiji et al. [161]	Yang et al. [79]	Zhao et al. [61]	Proposed approach
Car	0.4087	3813.7965	199.2996	−−	19.6112
Text	0.4048	3913.8527	204.1095	−−	18.5480
Natural HSI (31-band)	1.3288	11440.1020	628.0139	30130.5742	40.2350
AVIRIS HSI (196-band)	8.1160	11647.9102	750.3732	191487.6880	46.3458

training of dictionary is also not required. In conclusion, for the proposed approach, spatial and spectral fidelity is better preserved with reduced time complexity. Further improvement in the speed can be achieved by implementing the algorithm using C or C++ using optimized code.

6.7 Conclusion

We have presented an SR approach for hyperspectral images based on the design of an adaptive wavelet basis. We estimate the wavelet filter coefficients using a database and use them in our learning based super-resolution. The decimation matrix entries for each significant spectral band is represented in terms of estimated low-pass wavelet filter coefficients. Use of sparsity based regularization and the use of optimum degradation matix results in improved quality of super-resolved image when compared to other approaches. The advantage of the proposed technique is that there is no need of registration while learning. Quantitative comparison of score indices indicates that our method enhances spatial information without introducing significant spectral distortions. Experimental results show that the proposed approach outperforms other methods in terms reconstruction quality and computational complexity.

Adaptive wavelet basis will have positive impact on the subsequent hyperspectral image processing applications where high spatial and spectral resolution is desirable. We conclude that the wavelet basis can be tailored to take care of the variability in sensor characteristics. Our future work involves the incorporation of spectral mixing models in order to improve the estimation of filter coefficients and hence the SR.

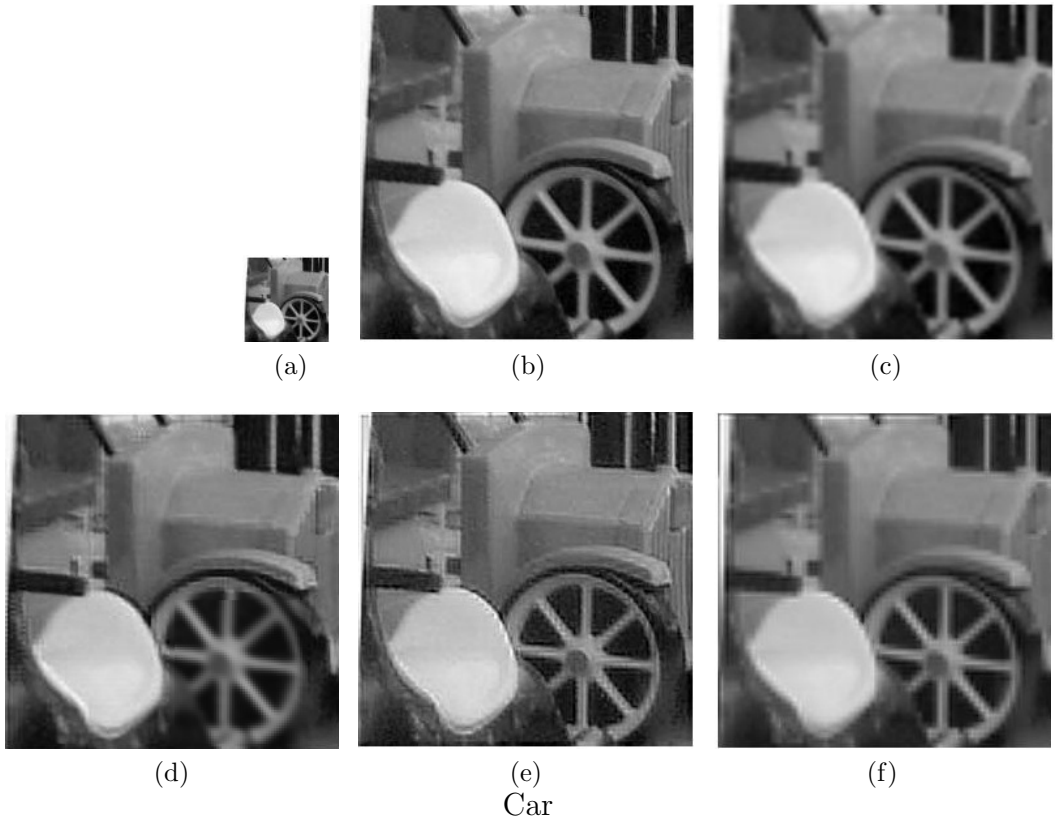


Figure 6.5: Experimental results on single band images for $q = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) SR image using Yang et al. method [79], (e) Initial SR image using Db4 wavelet [161], and (f) Initial SR estimate using proposed approach

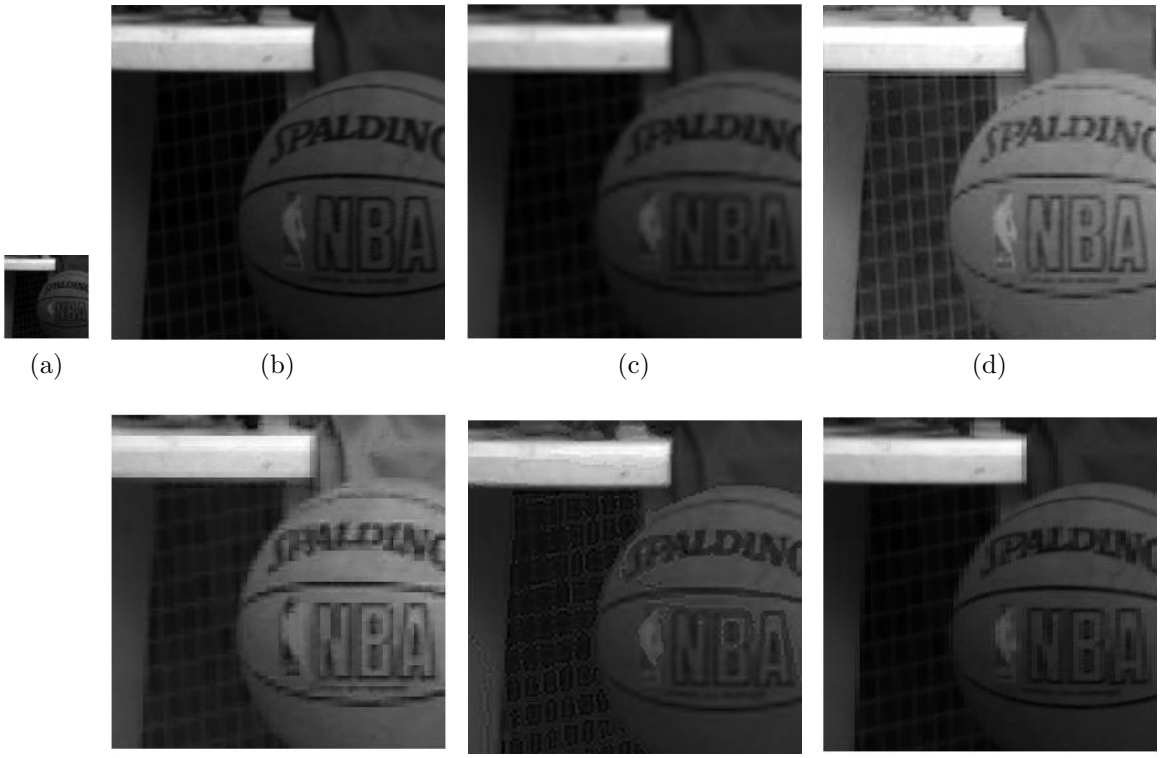
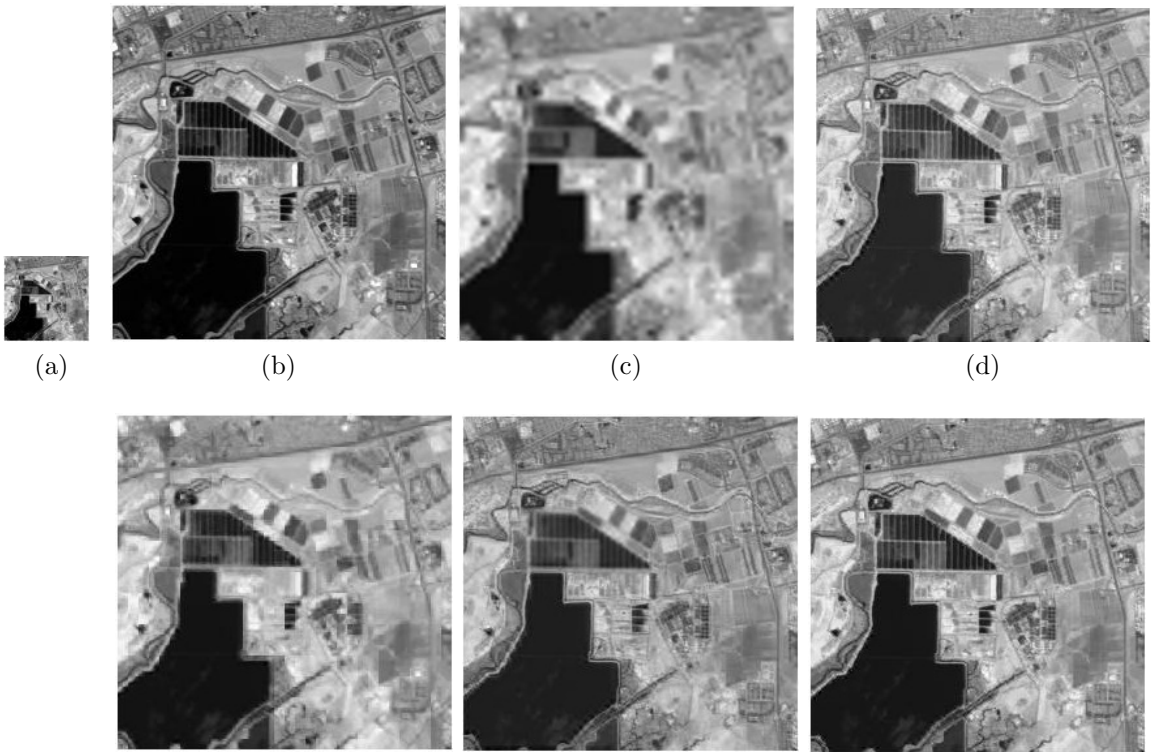


Figure 6.6: Experimental results on PCA-1 of 31-band natural HSI for $q = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) SR image using Jiji et al. method [161], (e) SR image using Zhao et al. method [61], (f) SR image using Yang et al. method [79], and (g) SR image using the proposed approach.



(a) (b) (c) (d)

(e) (f) (g)

Figure 6.7: Experimental results on 100^{th} band of AVIRIS data for $q = 4$. (a) LR test image of size 64×64, (b) Ground truth of size 256×256, (c) Bicubic interpolation [148], (d) SR image using Jiji et al. method [161], (e) SR image using Zhao et al. method [61], (f) SR image using Yang et al. method [79], and (g) SR image using the proposed approach.

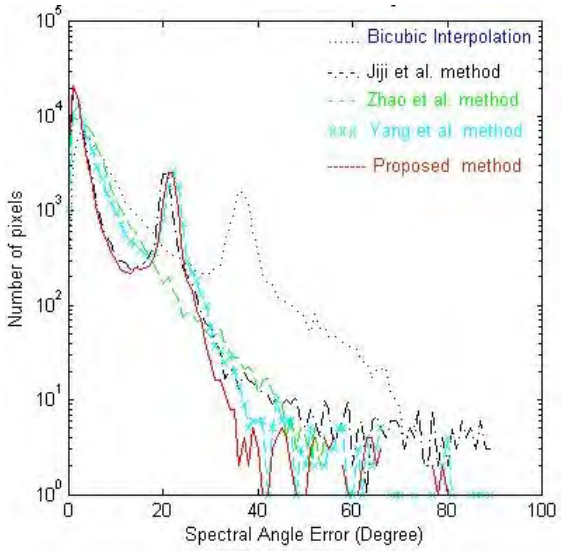


Figure 6.8: Spectral angle error for AVIRIS HSI for $q = 4$.

Chapter 7

Conclusion and Future Research Work

7.1 Conclusion

In this thesis, we have addressed the problems of multiresolution fusion and super-resolution (SR) to increase the spatial resolution of multispectral and hyperspectral images, respectively. Approach based on multiresolution fusion needs auxiliary HR panchromatic (PAN) image in registered form, while SR based approaches do not need auxiliary information. We first enhanced spatial resolution of multispectral images using multiresolution fusion in compressive sensing framework. This approach uses registered HR PAN image of the same geographic area as the auxiliary information. We next consider the super-resolution techniques based on single observation without using auxiliary information and registration. However, it is not possible to extract additional non-redundant information for resolution enhancement by using a single observation. Hence, explore different options for obtaining the additional information. We know that it is very difficult to capture the richness of textures in hyperspectral images analytically. This has motivated us to use empirical data rather than theoretical model in order to obtain non-redundant information for the super-resolution of hyperspectral image. Since hyperspectral images have high dimensions, we first perform dimensionality reduction using principal component analysis and then apply super-resolution algorithms on the reduced set. After performing SR in transform domain, inverse transform gives us SR for the entire hyper-

spectral cube. Our approaches are based on CS, in which sparse representation of the image plays significant role to obtain the super-resolution. In compressive sensing based approaches, we obtain sparse representation of LR observations in terms of LR dictionary entries, and applying the same sparseness on HR we obtain the SR. In our work we used dictionaries derived using different approaches. We first considered raw LR-HR dictionaries of most informative PCA component having large number of atoms. Then we used jointly optimized LR-HR dictionaries of PCA components representing most of the spectral variance of the hyperspectral images having very few atoms. We also solved the problem of super-resolution using wavelet based learning in which adaptive wavelets are designed from the given HS image itself. The inrverse problem of obtaining SR was solved using regularization approach in which the forward process of image formation is represented using a linear model. Different kinds of point spread functions to represent the degradation were considered in the image formation model and priors such as autoregressive, sparsity and Gabor were considered for regularizing the solution.

To begin with, we first considered the spatial resolution enhancement of the multispectral (MS) images using the available panchromatic image. This is referred to as multiresolution fusion or pansharpening and the fusion of the MS and PAN images is obtained in two steps. In the first step, we used wavelet to decompose panchromatic image to make the size of its coarser part same as that of MS image. A CS based method was used to obtain the missing detail coefficients. Assuming same sparseness for LR and HR MS images, the HR details in the fused image are obtained as the linear combination of the HR atoms of Pan image. The inverse-discrete wavelet transform (IDWT) of the learned image is considered as an initial estimate to fused image. In the second step, regularization using maximum a posteriori - Markov random field (MAP-MRF) approach is used to obtain final fused image. Here MRF model parameter was estimated from initial estimate of fusion. Experiments conducted on simulated and real data indicate superior performance of our approach.

Subsequently in this thesis, we present problems of increasing spatial resolution of hyperspectral images without using auxiliary high-resolution image. Hyperspectral images have very large number of spectral bands, hence we first reduced dimensionality of HSIs using principal component analysis (PCA) and worked in reduced dimensional space i.e. we used PCA components for SR. We create LR-HR raw dictionaries by randomly

choosing large number of small blocks (patches) from PCA-I component i.e. first PCA component of training images. Applying CS based approach to first PCA component results in initial SR estimate. Assuming that the pixel intensity in the image is a linear combinations of the intensities of its neighboring pixels, a better solution to SR is obtained by using the autoregressive and sparsity prior models on the high-resolution hyperspectral image. Remaining principal components were bicubically interpolated and further improved using autoregressive prior model. Use of autoregressiveness as a prior model improves the solution as well as reduces computational burden. Here we considered the degradation matrix as the average of non-overlapping blocks by considering each LR pixel as sum of equally weightaged corresponding HR pixels in a block. Applying inverse PCA results in SR HSI in spatial domain. Experiments were conducted on real hyperspectral images. Visual inspections and quantitative comparison confirm the effectiveness of proposed method.

Compressive sensing (CS) based approach was further extended to include all significant PCA components to cover most of the spectral information of hyperspectral image. Here LR and HR raw dictionaries of significant principal components (PCA) having large variance and having large number of atoms were jointly trained using K-SVD algorithm to reduce the dictionary size considerably. Initial SR of principal components in PCA domain are obtained by using the optimized dictionaries using the CS framework. In this case we used a new prior while regularizing the solution. This prior termed as "Gabor prior" was used to obtain solution for each of the SR PCA components. Instead of considering the ideal square point spread function (PSF) (i.e. averaging) we estimate the degradation matrix separately for each of the initial SR PCA images that gives us better model to represent the degradation between LR and HR images. Visual inspections and quantitative comparison with various other methods confirmed enhancement of spatial information without introducing significant spectral distortion.

Finally, we proposed a learning based approach for SR using wavelet transform based method. Both the low pass and high pass wavelet filter coefficients of the chosen wavelet are estimated from the data itself. Tailored wavelets were then used in learning the missing high frequency details i.e., edges in horizontal, vertical, and diagonal by making use of a database of HR images. Use of tailored wavelets while learning improved the initial SR estimate over fixed basis. These estimated wavelet filter coefficients were also used in

obtaining the degradation in the image formation model which was then used in getting the final SR image. The refinement of the initial estimate is done by using the sparsity based prior in regularization framework which takes care of contextual dependencies. The final cost function being convex was minimized using gradient descent optimization technique. Experiments were conducted on three different kinds of datasets including HSIs captured in controlled laboratory environment as well as on remotely captured real HSIs. The results show that the proposed approach offers significant performance as well as computational advantages over the other methods.

Before we end the conclusion it may be of interest to mention the following points regarding the reasons for use of initial estimate in our works. We assume that the initial estimate represents close approximation to the final fused/super-resolution (SR) image. Note that though, it can be considered as closer to the final solution one cannot accept it as the final solution. This is because the problem of fusion/SR is ill-posed and a solution space of initial estimate without regularization is larger when compared to that obtained after regularization. However, one may still consider the initial estimate to represent the global characteristics such as edge details, spatial dependencies etc. of the final solution. In addition to this, we often require an accurate estimate of input-output relationship where the output is not known. In such situations it is reasonable to derive an initial estimate of the output from the given data and use the same for finding the transformation. Due to these reasons we derive an initial estimate and use it for obtaining an improved solution.

7.2 Future Research Work

Improving spatial resolution of hyperspectral images requires additional information in the form of multiple observations, HR auxiliary image, or a large database that has high-resolution images. Unfortunately, there do not exist an universal solution to improve spatial resolution of hyperspectral images. Most of the spatial resolution enhancement techniques for HSIs suffer from the limitations such as higher dimensionality, need of auxiliary image of the same geographic area, need for registration, high computational complexity, need for large training database, etc. Although significant amount of work has been carried out in the area of SR of hyperspectral images, there is still enough scope

for improvement. In this section, we discuss the directions for future research work in obtaining high spatial resolution hyperspectral images.

- We have proposed the super-resolution using wavelet transform based learning in which we have derived wavelet filter coefficients using the LR-HR hyperspectral training database. However limitations of wavelets is that they represent limited directions. In addition 2-D wavelets do not perform well when there are edges along counters. The hyperspectral images consists of intrinsic geometric structures with edges located along smooth curves as well. The contourlet transform is capable of capturing high frequency information oriented at various directions in multiple scales with flexible aspect ratios. Hence, one can design adaptive contourlet transform from the database, and can use it for learning the fine details from the database of HR images.

- In wavelet learning based approach as a compromise between performance and computational load we have chosen a filter length of 4. We can estimate optimum length of the filter to design wavelets using training database and can use them while learning the high frequency details.

- We considered space invariant blur while constructing image observation model. In practice, point spread function of hyperspectral imager depends on various factors such as fill factor of CCD array, camera gain, zoom factor, imaging wavelength etc. The effect of diffraction is significant at higher wavelength in a hyperspectral imager. This results in spatially and spectrally varying point spread function (PSF). Thus one may consider PSF which varies spatially as well as spectrally. In order to incorporate spatially varying PSF one has to estimate the PSF at every pixel.

- In our work we considered a linear image formation model. One may consider more general non-linear model and also relax the assumption on independent and identically distributed noise.

- Most of the work on SR uses iterative methods to arrive at the final solution. Hence real time implementation becomes a difficult task. One may consider future works that result in closed form solutions with reduced computational cost.

Bibliography

[1] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Proc. Mag.*, pp. 12–16, 2002.

[2] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, pp. 17–28, 2002.

[3] G. A. Shaw and H. hua K. Burke, "Spectral imaging for remote sensing," *Lincoln Laboratory Journal*, vol. 14(1), pp. 3–28, 2003.

[4] P. K. Varshney and M. K. Arora, *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*. Springer, 2004.

[5] H. Zhang, L. Zhang, and H. Shen, "A super-resolution reconstruction algorithm for hyperspectral images," *Signal Processing*, vol. 92, no. 9, pp. 2082–2096, 2012.

[6] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 466–479, March 2005.

[7] P. Mitra, C. A. Murthy, and S. Pal, "Unsupervised feature selection using feature similarity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 301–312, Mar 2002.

[8] I. T. Jolliffe, *Principal Component Analysis, Second Edition*. Springer, 1986.

[9] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[10] S. Panasyuk, J. Freeman, and A. Panasyuk, "Medical hyperspectral imaging for evaluation of tissue and tumor," Nov. 27 2012, US Patent 8,320,996.

[11] J. Freeman, S. Panasyuk, M. Hopmeier, K. Schomacker, and D. Brand, "Hyperspectral imaging in diabetes and peripheral vascular disease," Jul. 17 2012, US Patent 8,224,425.

[12] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[13] M. V. Joshi, L. Bruzzone, and S. Chaudhuri, "A model-based approach to multiresolution fusion in remotely sensed images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2549–2562, 2006.

[14] K. P. Upla, P. P. Gajjar, and M. V. Joshi, "Multiresolution fusion using contourlet transform based edge learning," in *Proc. IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2011, pp. 523–526.

[15] M. E. Winter and E. M. Winter, "Resolution enhancement of hyperspectral data," in *Proceedings IEEE Aerospace Conference*, 2002.

[16] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 13(9), pp. 1174–1184, 2004.

[17] A. P. Cracknell, "Synergy in remote sensing - what's in a pixel?" *International Journal of Remote Sensing*, vol. 19, no. 11, pp. 2025–2047, 1998.

[18] G. Vane, H. T. E. Robert O Green, Thomas G Chrien, E. G. Hansen, and W. M. Porter, "The airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environ.*, vol. 44(2-3), pp. 127–143, 1993.

[19] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, Supplement 1, no. 0, pp. S110 – S122, 2009, imaging Spectroscopy Special Issue.

[20] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Advances in Computer Vision and Image Processing,JAI Press Inc., Greenwich, CT*, vol. 1(2), 1984.

[21] T. A. Wilson, S. K. Rogers, and M. Kabrisky, "Perceptual-based image fusion for hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35(4), pp. 1007–1017, 1997.

[22] R. B. Gomez, A. Jazaeri, and M. Kafatos, "Wavelet-based hyperspectral and multispectral image fusion," in *Proceedings of the SPIE*, vol. 4383, 2001, pp. 36–42.

[23] M. T. Eismann and R. C. Hardie, "Application of the stochastic mixing model to hyperspectral resolution enhancement," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 9, pp. 1924–1933, 2004.

[24] Y. Gu, Y. Zhang, and J. Zhang, "Integration of spatial spectral information for resolution enhancement in hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46(5), pp. 1347–1358, 2008.

[25] H. Shen, M. K. Ng, P. Li, and L. Zhang, "Super-resolution reconstruction algorithm to MODIS remote sensing images." *Computer Journal*, vol. 52, no. 1, pp. 90–100, 2009.

[26] F. A. Mianji, Y. Zhang, and Y. Gu, "Resolution enhancement of hyperspectral images using a learning-based super-resolution mapping technique," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS (3)*, 2009, pp. 813–816.

[27] Y. Zhang, S. D. Backer, and P. Scheunders, "Noise-resistant wavelet-based bayesian fusion of multispectral and hyperspectral images." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3834–3843, 2009.

[28] A. Garzelli, B. Aiazzi, S. Baronti, M. Selva, and L. Alparone, "Hyperspectral image fusion," in *Proceedings Hyperspectral 2010 Workshop, Frascati, Italy*, 2010.

[29] S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences: A comprehensive review with directions for future research," in *Technical Report*, 1998.

[30] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, pp. 21 – 36, 2003.

[31] J. Mateos, M. Vega, R. Molina, and A. Katsaggelos, "Pansharpening of multispectral images using a TV-based super-resolution algorithm," in *7th International Workshop on Information Optics (WIO-08), Journal of Physics: Conference Series*, vol. 139. Annecy (France), June 2008, p. 012022.

[32] M. Gonzlez-Audicana, J. L. Saleta, R. G. Cataln, and R. Garcia, "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 6, pp. 1291–1299, 2004.

[33] T. Akgun, Y. Altunbasak, and R. M. Mersereau, "Super-resolution reconstruction of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 14, no. 01, pp. 1860–1875, 2005.

[34] J. C. W. Chan, J. Ma, P. Kempeneers, and F. Canters, "Super-resolution enhancement of hyperspectral CHRIS/Proba images with a thin-plate spline nonrigid transform model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48(6), pp. 2569–2579, 2010.

[35] X. Xu, Y. Zhong, L. Zhang, and H. Zhang, "Sub-pixel mapping based on a MAP model with multiple shifted hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 580–593, April 2013.

[36] R. Nishii, S. Kusanobu, and S. Tanaka, "Enhancement of low spatial resolution image based on high resolution-bands," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 5, pp. 1151–1158, Sep 1996.

[37] M. T. Eismann and R. C. Hardie, "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43(3), pp. 455–465, 2005.

[38] M. Simões, J. M. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "Hyperspectral image superresolution: An edge-preserving convex formulation," *International Conference on Image Processing (ICIP)*, vol. arXiv:1403.8098v2, 2014.

[39] J. C. Price, "Combining panchromatic and multispectral imagery dual resolution satellite instruments," *Remote Sensing of Environment*, vol. 21, no. 2, pp. 119–128, Mar. 1987.

[40] L. Capobianco, A. Garzelli, L. A. Filippo Nencini, and S. Baronti, "Spatial enhancement of hyperion hyperspectral data through ALI panchromatic image," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2007.

[41] D. E. Bar, K. Wolowelsky, Y. Swirski, Z. Figov, A. Michaeli, Y. Vaynzof, Y. Abramovitz, A. Ben-Dov, O. Yaron, L. Weizman, and R. Adar, "Target detection and verification via airborne hyperspectral and high-resolution imagery processing and fusion," *IEEE Sensors Journal*, vol. 10(3), pp. 707–711, 2010.

[42] Y. Zhao, J. Yang, and J.-W. Chan, "Hyperspectral imagery super-resolution by spatial−spectral joint nonlocal similarity," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2671–2679, June 2014.

[43] H. Ur and D. Gross, "Improved resolution from sub-pixel shifted pictures," *CVGIP: Graphical models and Image Processing*, vol. 54, pp. 181 – 186, 1992.

[44] M. S. Alam, J. G. Bognar, R. C. Hardie, and B. J. Yasuda, "Infrared registration and high resolution reconstruction using multiple translationally shifted aliased video frames," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, pp. 915 – 923, 2000.

[45] Q. Tian and M. N. Huhns, "Algorithm for subpixel registration," *Computer Vision, Graphics, Image Processing*, vol. 35, pp. 220 – 233, 1986.

[46] L. G. Brown, "A survey of image registration techniques," *ACM Computer Surveys*, vol. 24, no. 4, pp. 325 – 376, 1992.

[47] R. Li, B. Zeng, and M.-L. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 4, no. 4, pp. 438–442, Aug 1994.

[48] J. Chalidabhongse and C.-C. Kuo, "Fast motion vector estimation using multiresolution-spatio-temporal correlations," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 3, pp. 477–488, Jun 1997.

[49] Q. Wang, X. Tang, and H. Shum, "Patch based blind image super resolution," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1, Oct 2005, pp. 709–716 Vol. 1.

[50] Y. He, K.-H. Yap, L. Chen, and L.-P. Chau, "A soft MAP framework for blind super-resolution image reconstruction," *Image and Vision Computing*, vol. 27, no. 4, pp. 364 – 373, 2009.

[51] D. Zhang, H. Li, and M. Du, "Fast MAP-based multiframe super-resolution image reconstruction," *Image and Vision Computing*, vol. 23, no. 7, pp. 671 – 679, 2005.

[52] A. Zomet and S. Peleg, "Multi-sensor super-resolution," *Applications of Computer Vision, IEEE Workshop on*, vol. 0, pp. 27–31, 2002.

[53] M. E. Winter and E. M. Winter, "Physics-based resolution enhancement of hyperspectral data," *Proceedings of the SPIE*, vol. 4725, pp. 580–587, 2002.

[54] F. J. Gaecia-Haro, M. Gilabert, and J. Melia, "Linear spectral mixture modelling to estimate vegetation amount from optical spectral data," *International Journal of Remote Sensing*, vol. 17, no. 17, pp. 3373–3400, 1996.

[55] M. Nguyen, P. Atkinson, and H. Lewis, "Superresolution mapping using a hopfield neural network with fused images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 3, pp. 736–749, 2006.

[56] P. Atkinson, M. Cutler, and H. Lewis, "Mapping sub-pixel proportional land cover with AVHRR imagery," *International Journal of Remote Sensing*, vol. 18, no. 4, pp. 917–935, 1997.

[57] R. A. Schowengerdt, *Remote Sensing, Second Edition: Models and Methods for Image Processing*. San Diego, USA: CA Academic, 1997.

[58] M. Brown, S. R. Gunn, and H. G. Lewis, "Support vector machines for optimal classification and spectral unmixing," *Ecological Modelling*, vol. 120, no. 23, pp. 167 – 179, 1999.

[59] A. B. Ruescas, J. A. Sobrino, Y. Julien, J. C. Jimnez-Muoz, G. Sria, V. Hidalgo, M. Atitar, B. Franch, J. Cuenca, and C. Mattar, "Mapping sub-pixel burnt percentage using AVHRR data. application to the alcalaten area in spain," *International Journal of Remote Sensing*, vol. 31, no. 20, pp. 5315–5330, 2010.

[60] A. Villa, J. Chanussot, J. A. Benediktsson, and C. Jutten, "Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 521–533, 2011.

[61] Y. Zhao, J. Yang, Q. Zhang, L. Song, Y. Cheng, and Q. Pan, "Hyperspectral imagery super-resolution by sparse representation and spectral regularization," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, p. 87, 2011.

[62] A. Villa, J. Chanussot, J. A. Benediktsson, C. Jutten, and R. Dambreville, "Unsupervised methods for the classification of hyperspectral images with low spatial resolution," *Pattern Recognition*, vol. 46, no. 6, pp. 1556–1568, 2013.

[63] M. K. Arora and K. Tiwari, "Subpixel target enhancement in hyperspectral images," *Defence Science Journal*, vol. 63, no. 1, pp. 63–68, January 2013.

[64] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," *ECCV, LNCS, Springer International Publishing Switzerland*, vol. VII, no. 8695, pp. 63–78, 2014.

[65] A. J. Tatem, P. Atkinson, H. G. Lewis, and M. S. Nixon, "Super-resolution target identification from remotely sensed images using a hopfield neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 4, pp. 781–796, April 2001.

[66] F. A. Mianji, Y. Gu, Y. Zhang, and J. Zhang, "Enhanced self-training superresolution mapping technique for hyperspectral imagery," *IEEE Geoscience Remote Sensing Letters*, vol. 8, no. 4, pp. 671–675, 2011.

[67] Y. Zhang, Y. Du, F. Ling, S. Fang, and X. Li, "Example-based super-resolution land cover mapping using support vector regression," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1271–1283, April 2014.

[68] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," pp. 11–14, 1993.

[69] ——, "Analysis, understanding, and visualization of hyperspectral data as convex sets in n space," *Proceedings SPIE*, vol. 2480, pp. 14–22, 1995.

[70] J. M. Grossmann, J. H. Bowles, D. Haas, J. A. Antoniades, M. R. Grunes, P. J. Palmadesso, D. Gillis, K. Y. Tsang, M. M. Baumback, M. Daniel, J. Fisher, and I. A. Triandaf, "Hyperspectral analysis and target detection system for the adaptive spectral reconnaissance program (ASRP)," vol. 3372, 1998, pp. 2–13.

[71] M. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 3, pp. 542–552, May 1994.

[72] *N−FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data*, vol. 3753, 1999.

[73] R. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," in *Fourth International Airborne Remote Sensing Conference and Exhibition /21st Canadian Symposium on Remote Sensing, Ottawa, Ontario, Canada*, 1999.

[74] A. Plaza, P. Martnez, R. M. Prez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2025–2041, 2002.

[75] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles," *Pattern Recognition*, vol. 37, no. 6, pp. 1097 – 1116, 2004.

[76] K. Mertens, L. Verbeke, T. Westra, and R. De Wulf, "Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients," *Remote Sensing of Environment*, vol. 91, no. 2, pp. 225–236, 2004.

[77] D. J. F. B A. Olshausen, "Natural image statistics and efficient coding," *Network : Computation in Neural Systems*, vol. 7, pp. 333–339, 1996.

[78] K. Huang and S. Aviyente, "Sparse representation for signal classification," in *Advances in Neural Information Processing Systems, NIPS*, 2006, pp. 609–616.

[79] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[80] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15(12), pp. 3736 –3745, 2006.

[81] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53–69, 2008.

[82] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[83] P. Sen and S. Darabi, "Compressive image super-resolution," in *Proceedings of the 43rd Asilomar Conference on Signals, Systems and Computers*, ser. Asilomar'09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 1235–1242.

[84] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 972–982, 2010.

[85] Z. Guo, T. Wittman, and S. Osher, "L1 unmixing and its application to hyperspectral image enhancement," vol. 7334, 2009, pp. 73 341M–73 341M–9.

[86] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49(6), pp. 2014–2039, 2011.

[87] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pal, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1204–1211, 1999.

[88] F. Li, X. Jia, D. Fraser, and A. J. Lambert, "Super resolution for remote sensing images based on a universal hidden markov tree model." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3-1, pp. 1270–1278, 2010.

[89] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.

[90] C. V. Jiji and S. Chaudhuri, "Single-frame image super-resolution through contourlet learning," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, 2006.

[91] A. Yoshikawa, S. Suzuki, T. Goto, S. Hirano, and M. Sakurai, "Super resolution image reconstruction using total variation regularization and learning-based method," in *Proceedings of the International Conference on Image Processing, ICIP 2010, September 26-29, Hong Kong, China*. IEEE, 2010, pp. 1993–1996.

[92] W. Wu, Z. Liu, and X. He, "Learning-based super resolution using kernel partial least squares," *Image and Vision Computing*, vol. 29, no. 6, pp. 394–406, May 2011.

[93] D. P. Capel and A. Zisserman, "Super-resolution from multiple views using learnt image models," in *CVPR (2)*, 2001.

[94] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image analogies," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '01, 2001, pp. 327–340.

[95] K. I. Kim and Y. Kwon, "Example-based learning for single-image super-resolution," in *Proceedings of the 30th DAGM symposium on Pattern Recognition*. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 456–465.

[96] C. Su, Y. Zhuang, L. Huang, and F. Wu, "Steerable pyramid-based face hallucination," *Pattern Recognition*, vol. 38, no. 6, pp. 813–824, 2005.

[97] G. Chen, S.-E. Qian, J.-P. Ardouin, and W. Xie, "Super-resolution of hyperspectral imagery using complex ridgelet transform," *International Journal of Wavelets, Multiresolution and Information Processing, IJWMIP*, vol. 10, no. 3, 2012.

[98] M. Ng, T. Chan, M. G. Kang, and P. Milanfar, Eds., *EURASIP Journal on Advances in Signal Processing: Special Issue on Super-resolution Imaging: Analysis, Algorithms, and Applications*, 2006, vol. 2006.

[99] R. C. Hardie, R. R. Schultz, and K. E. Barner, Eds., *EURASIP Journal on Advances in Signal Processing: Special Issue on Super-resolution Enhancement of Digital Video*, 2006, vol. 2007.

[100] N. K. Bose, R. H. Chan, and M. K. Ng, Eds., *International Journal of Imaging Systems and Technology: Special Issue on High-Resolution Image Reconstruction-I*, 2004, vol. 14.

[101] M. G. Kang and S. Chaudhuri, Eds., *IEEE Signal Processing Magazine: Special Issue on Super-Resolution Image Reconstruction*, 2003, vol. 20, no. 3.

[102] J. J, *Image Fusion Techniques for Remote Sensed Data.* LAP Lambert Academic Publishing, 2011.

[103] H. Nguyen, *Spatial statistical data fusion for remote sensing applications.* Proquest, Umi Dissertation Publishing, 2011.

[104] S. Chaudhuri and K. Kotwal, Eds., *Hyperspectral Image Fusion.* Springer, 2013.

[105] G. B. A. Alparone, *Remote Sensing Image Fusion.* CRC Press, 2014.

[106] S. B. A. G. Luciano Alparone, Bruno Aiazzi, *Remote Sensing Image Fusion.* CRC Press, Taylor Francis Group, 2015.

[107] S. Chaudhuri, Ed., *Super-resolution Imaging.* Kluwer, 2001.

[108] D. Capel, *Image Mosaicing and Super-Resolution.* Springer, 2004.

[109] S. Chaudhuri and M. V. Joshi, *Motion-free Super-resolution*. Springer, 2005.

[110] P. Milanfar, *Super-Resolution Imaging (Digital Imaging and Computer Vision)*. CRC Press, 2010.

[111] R. Hayden, G. W. J. Henkel, and J. E. Bare, "Application of the IHS color transform to the processing of multisensor data and image enhancement," in *Proceedings of International Symposium on Remote sensing of arid and semi-arid lands*, Cairo, Egypt, 1982, pp. 599–616.

[112] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1672–1682, June 2006.

[113] P. S. J. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 3, pp. 295–303, 1991.

[114] V. P. Shah, N. H. Younan, and R. L. King, "An efficient Pan-sharpening method via a combined adaptive PCA approach and contourlets." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1323–1335, 2008.

[115] Y. Zhang, "Problems in the fusion of commercial highresolution satellite images as well as landsat-7 images and initial solutions," *International Archives Of Photogrammetry Remote Sensing And Spatial Information Sciences*, vol. 34, no. 4, pp. 587–592, 2002.

[116] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 27, no. 3, pp. 235–244, 1995.

[117] D. Fasbender, J. Radoux, and P. Bogaert, "Bayesian data fusion for adaptable image pansharpening," *IEEE Transactions on Geoscience and remote sensing*, vol. 46, no. 6, pp. 1847–1857, June 2008.

[118] P. S. Pradhan, R. L. King, N. H. Younan, and D. W. Holcomb, "Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion," *IEEE Transactions on Geoscience and remote sensing*, vol. 44, no. 12, pp. 3674 – 3686, 2006.

[119] X. X. Zhu, X. Wang, and R. Bamler, "Compressive sensing for image fusion with application to pan-sharpening," in *Proc. IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2011, pp. 2793–2796.

[120] E. J. Cand'es, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 910, pp. 589–592, 2008.

[121] E. Candes and T. Tao, "The dantzig selector: Statistical estimation when p is much larger than n," *The Annals of Statistics*, vol. 35(6), pp. 2313–2351, 2007.

[122] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *Society for Industrial and Applied Mathematics*, vol. 51(1), pp. 34–81, 2009.

[123] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, no. 3, pp. 233–242, May 1994.

[124] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, "Image fusion − the ARSIS concept and some successful implementation schemes," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, pp. 4–18, 2003.

[125] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr 2004.

[126] R. Schultz and R. Stevenson, "A bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, no. 3, pp. 233–242, 1994.

[127] M. V. Joshi, L. Bruzzonne, and S. Chaudhuri, "A model based approach to multiresolution fusion in remotely sensed images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2549–2562, 2006.

[128] S. Z. Li, *Markov Random Field Modeling in Computer Vision.* London, UK, UK: Springer-Verlag, 1995.

[129] Y. Qian, F. Yao, and S. Jia, "Band selection for hyperspectral imagery using affinity propagation," *IET Computer Vision*, vol. 3(4), p. 213 222, 2009.

[130] D. M. Tralli, V. Z. Ronald G. Blom, A. Donnellan, and D. L. Evans, "Satellite remote sensing of earthquake, volcano, flood, landslide and coastal inundation hazards," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 59, p. 185 198, 2005.

[131] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14(1), 2003.

[132] H. K. Jusoff and H. M. H. M. Yusoff, "Airborne hyperspectral imagery for agricultural businesses in Malaysia," *International Business Research*, vol. 1(3), pp. 54–62, 2008.

[133] D. G. Goodenough, A. D. Jay Pearlman, Hao Chen, J. M. Tian Han, Jingyang Li, and K. O. Niemann, "Forest information from hyperspectral sensing," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2004.

[134] U. Heiden, K. S. Wieke Heldens, Sigrid Roessner, T. Esch, and A. Mueller, "Urban structure type characterization using hyperspectral remote sensing and height information," *Landscape and Urban Planning*, vol. 105, p. 361375, 2012.

[135] J. C.-W. Chan, J. V. B. T. Spanhove, J. Ma, D. Paelinckx, and F. Canters, "Natura 2000 habitat identification and conservation status assessment with superresolution enhanced hyperspectral (Chris/Proba) imagery," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII-4/C7, 2010.

[136] J. Yang, T. H. John Wright, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[137] S. Yang, F. Sun, M. Wang, Z. Liu, and L. Jiao, "Novel super resolution restoration of remote sensing images based on compressive sensing and example patches-aided dictionary learning," in *Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM), 2011 International Workshop on*, 2011, pp. 1–6.

[138] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 884–892, April 2010.

[139] J.-L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Transactions on Image Processing*, vol. 16(2), pp. 297–309, 2007.

[140] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transaction on Image Processing*, vol. 14(12), pp. 2091–2106, 2005.

[141] M. Elad, P. Q. J. L. Starck, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morph. compo. analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, pp. 340–358, 2005.

[142] A. Buades, J.-M. M. Triet M. Le, and L. A. Vese, "Fast cartoon + texture image filters," *IEEE Transactions on Image Processing*, vol. 19(8), pp. 1978 –1986, 2010.

[143] F. G. Meyer, A. Z. Averbuch, and J.-O. Strmberg, "Fast adaptive wavelet packet image compression," *IEEE Transactions on Image Processing*, vol. 9(5), pp. 792–800, 2000.

[144] E. L. Pennec and S. Mallat, "Bandelet image approximation and compression," *Multiscale Model. Simul.*, vol. 4(3), pp. 992–1039, 2005.

[145] M. Aharon, M. Elad, and A. Bruckstein, "K−SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54(11), pp. 4311–4322, 2006.

[146] A. Divekar and O. Ersoy, "Image fusion by compressive sensing," *17th International Conference on Geoinformatics*, vol. 2, no. 1, pp. 1–6, 2009.

[147] R. L. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Transactions on Information Theory*, vol. IT-29(1), pp. 60–72, 1983.

[148] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[149] . M. C. Nascimento, F. P. Ferreira, and D. H. Foster, "Statistics of spatial cone-excitation ratios in natural scenes," *Journ. of the Optical Soc. of America A*, vol. 19(8), pp. 1484–1490, 2002.

[150] F. A. Kruse, A. B. Lefkoff, A. T. S. J. W. Boardman, K. B. Heidebrecht, J. P. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, pp. 145–163., 1993.

[151] L. Wald, *Data Fusion : Definitions and Architectures - Fusion of Images of Different Spatial Resolutions*. Les Presses, Ecole des Mines de Paris, Paris, France, 2002.

[152] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical models and image processing*, vol. 53, no. 3, pp. 231–239, 1991.

[153] F. A. Mianji, Y. Zhang, H. K. Sulehria, A. Babakhani, and M. R. Kardan, "Super-resolution challenges in hyperspectral imagery," *Information Technology Journal*, vol. 7, pp. 1030–1036, 2008.

[154] B. Buttingsrud and B. K. Alsberg, "Superresolution of hyperspectral images," *Chemometrics and Intelligent Laboratory Systems*, vol. 84, pp. 62–68, 2006.

[155] N. Petkov, "Biologically motivated computationally intensive approaches to image pattern recognition," *Future Generation Computer Systems*, vol. 11, no. 4-5, pp. 451–465, 1995.

[156] B. Shahraray and D. J. Anderson, "Optimal estimation of contour properties by cross-validated regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 600–610, Jun. 1989.

[157] A. Garzelli, F. Nencini, L. Alparone, and S. Baronti, "A new method for quality assessment of hyperspectral images," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2007, pp. 5138–5141.

[158] W. D. Smith, "Quaternions, octonions, and now , 16-ons and 2^n -ons; new kinds of numbers." *Power*, pp. 1–68, 2004.

[159] G. Foody, "Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution," *International Journal of Remote Sensing*, vol. 19, pp. 2593–2599, 1998.

[160] L. Kleiman, "Hyperspectral imaging: the colorimetric high ground," *SPIE Proceedings, Color Imaging VIII: Processing, Hardcopy, and Applications*, vol. 5008, pp. 249–259, 2003.

[161] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, "Single-frame image super-resolution using learned wavelet coefficients," *International Journal of Imaging Systems and Technology*, vol. 14(3), pp. 105–112, 2004.

[162] B. Chalmond, "PSF estimation for image deblurring," *CVGIP: Graphical Model and Image Processing*, vol. 53, no. 4, pp. 364–372, 1991.

[163] R. C. Patel and M. V. Joshi, "Super-resolution of hyperspectral images using compressive sensing based approach," in *ISPRS Annals of Photogrammetry and Remote sensing*, 2012.

[164] H. Ji and C. Fermüller, "Robust wavelet-based super-resolution reconstruction: Theory and algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 649–660, Apr. 2009.

[165] P. P. Gajjar and M. V. Joshi, "New learning based super-resolution: Use of DWT and IGMRF prior," *IEEE Transactions on Image Processing*, vol. 5, pp. 1201–1213, 2010.

[166] Z. Hui and K.-M. Lam, "Eigentransformation-based face super-resolution in the wavelet domain," *Pattern Recogn. Lett.*, vol. 33, no. 6, pp. 718–727, Apr. 2012.

[167] I. Daubechies, "Orthonormal bases of compactly supported wavelets II: variations on a theme," *SIAM Journal on Mathematical Analysis*, vol. 24, no. 2, pp. 499–519, Mar. 1993.

[168] P. J. V. Fleet, *Discrete Wavelet Transformations: An Elementary Approach with Applications*. Wiley Interscience, 2008.

[169] G. Strang, *Linear Algebra and its Applications, 4e*. Brooks/Cole Cengage Learninig.

[170] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445 –3462, dec 1993.

[171] l1-magic. Available [Online]:, "http://www-stat.stanford.edu/ candes/l1magic/."

[172] Available [Online]:, "http://personalpages.manchester.ac.uk/staff/david.foster/hyperspectral_images_of_natural_scenes_02.html."

[173] Available [Online]: D. H. Brainard, "http://color.psych.ucsb.edu//hyperspectral/."

[174] Available [Online]: Aviris Free Data, Jet Propulsion Lab., California Inst. Tech., Pasadena, "http://aviris.jpl.nasa.gov/html/aviris.freedata.html."

List of Publications

Journal Publication

- Rakesh C. Patel, and Manjunath V. Joshi, "Super-Resolution of Hyperspectral Images: Use of Optimum Wavelet Filter Coefficients and Sparsity Regularization", IEEE Transactions on Geoscience and Remote sensing, Vol. 53(4), pp. 1728-1736, August 2014.

Conference Publications

- Rakesh C. Patel, Prakash P. Gajjar, Manjunath V. Joshi and Kishor P. Upla , "A Model Based Approach To Multiresolution Fusion: Use of CS Theory", IEEE International Geoscience and Remote Sensing Symposium (IGARSS-2012), July 22-27, 2012, Munich, Germany.

- Rakesh C. Patel, and M. V. Joshi, "Super-resolution of Hyperspectral Images using Compressive Sensing based Approach", ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume I-7, 2012, XXII ISPRS Congress, 25 August - 01 September 2012, Melbourne, Australia.

Publication Reward

- Travel Grants have been awarded to deserving researchers from ten (10) countries for their participation in XXII ISPRS Congress. I received a travel grant of 3100 A$ (AUD) (1,75,000.00 Rs) from The ISPRS Foundation (TIF).

Journal Submitted

- Rakesh C. Patel, and M. V. Joshi, "Super-resolution of Hyperspectral Images Using Sparse Representation and Gabor Prior", SPIE, Journal of Applied Remote Sensing, 2015.

Grant Awardees 2012

Honor Roll of 2012 Grantees

With thanks to our Donors, event sponsors and grant evaluation committees we are pleased to announce that The ISPRS Foundation (TIF) funding for 2010 has been provided in the four TIF grant categories of:

"Research Initiatives" " Travel Grants" " Awards"

Grants for Research Initiatives

Congratulations to the following successful applicants whose projects have been deemed worthy of financial grants to pursue their proposed projects.

CANADA –

Dr. Derek Lichti – University of Calgary "Compilation of an authoritative, on-line bibliography on Terrestrial Laser Scanning"
(Finalizing the project, see http://www.tlsdatabase.ucalgary.ca/)

top

Travel Grants

We are pleased to announce that eleven (11) Travel Grants have been awarded to deserving grantees from ten (10) countries for their participation in the events listed below.

Congratulations to the following Grantees:

ISPRS Congress 2012 in Melbourne
25 Aug - 1 Sep 2012 – Melbourne, Australia

BANGLADESH -
Mr. Abdul Nurunnabi

CHINA -
Ms. Xun Li

COLOMBIA -
Ms. Olga Piedad Rudas

HUNGARY -
Ms. Andrea Kovacs

INDIA -
Mr. Rakeshkumar Chandulal

IRAN -
Mr. Milad Niroumand Jadidi

IRAQ -
Mr. Bashar Alsadik
Mr. Ali Maliki

ITALY -
Ms. Cristina Re

PAKISTAN -
Mr. Amjad Ali

RUSSIA -
Mr. Alexander Velizhev

top

The ISPRS Foundation Brochure