Efficient Methodologies for
Real-time Image Restoration

Pradeepa D Samarasinghe
B.Sc.(Eng.), M.B.A.(IT) (University of Moratuwa, Sri Lanka)

October 2011

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Applied Signal Processing Group
Research School of Engineering
College of Engineering and Computer Science
The Australian National University
DEDICATED TO

MY LATE FATHER, WHO ALWAYS INSPIRED ME WITH HIS DETERMINATION

AND

MY MOTHER, WHOSE COMMITMENT TO MY CHILDREN MADE THIS THESIS A REALITY.
Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution. The research represented in this thesis has been performed jointly with Professor Rodney A. Kennedy and Dr. Hongdong Li.

Much of the work in this thesis has been published or has been submitted for publication as journal papers or conference proceedings. In some cases, the conference papers contain material overlapping with the journal publications. The following is a list of those publications.

Published


**Accepted for publication**


**Submission pending**


---

Pradeepa D. Samarasinghe  
College of Engineering and Computer Science,  
The Australian National University,  
Canberra,  
ACT 0200,  
Australia.
Acknowledgements

There are many who deserve my sincere thanks. First and foremost, it is with immense gratitude that I acknowledge the valuable guidance and encouragement of my supervisor, Professor Rodney Kennedy. His insight and enthusiasm for the research continuously inspired me. His patience and willingness to help went as far as spending many hours with me in explaining and guiding. Professor Kennedy persistently and convincingly conveyed the spirit and excitement in research, which resulted my exploration on different avenues of research. I couldn’t write enough to express my gratitude, but I am greatly indebted to Professor Kennedy for the continuous support and understanding throughout my research and I share the credit of my work with him.

I consider it an honor to work with Dr. Hongdong Li and Dr. Ramtin Shams and wish to thank them for their stimulating discussions and valuable suggestions. It gave me a great pleasure in doing my research with the applied signal processing group at the Research School of Information Sciences and Engineering. I would like to thank all the members of the group for the friendly and supportive research environment. In addition, special thanks go to Ms. Lesley Goldburg and Ms. Elspeth Davies, for their assistance in administrative work and Mr. James Ashton for IT support.

I would like to reference my sources of financial support, which made it possible to deliver this thesis on time. Firstly, I would like to acknowledge the financial support provided by the Australian National University. My sincere thanks are expressed to CSIRO for the stipend provided through their ICT Scholarship Fund Scheme. Prof. Kennedy and Dr. Thushara Abhayapala are also thanked for helping me on arranging the scholarship.

This thesis would have remained a dream had it not been for my mother who
cared for my three children in spite of her other interests. Thus, I owe my deepest gratitude to my mother for the commitment of her time. My three children deserve special thanks for bearing with me the time I spent on research in their playful age. Finally, thanks must also be extended to my husband Prasad for his constant encouragement and support, specially, the moral support in completing the thesis was invaluable.
Abstract

In this thesis we investigate the problem of image restoration. The main focus of our research is to come up with novel algorithms and enhance existing techniques in order to deliver efficient and effective methodologies, applicable in real-time image restoration scenarios.

Our research starts with a literature review, which identifies the gaps in existing techniques and helps us to come up with a novel classification on image restoration, which integrates and discusses more recent developments in the area of image restoration. With this novel classification, we identified three major areas which need our attention.

The first developments relate to non-blind image restoration. The two mostly used techniques, namely deterministic linear algorithms and stochastic nonlinear algorithms are compared and contrasted. Under deterministic linear algorithms, we develop a class of more effective novel quadratic linear regularization models, which outperform the existing linear regularization models. In addition, by looking in a new perspective, we evaluate and compare the performance of deterministic and stochastic restoration algorithms and explore the validity of the performance claims made so far on those algorithms. Further, we critically challenge the necessity of some complex mechanisms in Maximum A Posteriori (MAP) technique under stochastic image deconvolution algorithms.

The next developments are focussed in blind image restoration, which is claimed to be more challenging. Constant Modulus Algorithm (CMA) is one of the most popular, computationally simple, tested and best performing blind equalization algorithms in the signal processing domain. In our research, we extend the use of CMA in image restoration and develop a broad class of blind image deconvolution algorithms, in particular algorithms for blurring kernels with a separable prop-
erty. These algorithms show significantly faster convergence than conventional algorithms.

Although CMA method has a proven record in signal processing applications related to data communications systems, no research has been carried out to the investigation of the applicability of CMA for image restoration in practice. In filling this gap and taking into account the differences of signal processing in image processing and data communications contexts, we extend our research on the applicability of CMA deconvolution under the assumptions on the ground truth image properties. Through analyzing the main assumptions of ground truth image properties being zero-mean, independent and uniformly distributed, which characterize the convergence of CMA deconvolution, we develop a novel technique to overcome the effects of image source correlation based on segmentation and higher order moments of the source.

Multichannel image restoration techniques recently gained much attention over the single channel image restoration due to the benefits of diversity and redundancy of the information between the channels. Exploiting these benefits in real time applications is often restricted due to the unavailability of multiple copies of the same image. In order to overcome this limitation, as the last area of our research, we develop a novel multichannel blind restoration model with a single image, which eliminates the constraint of the necessity of multiple copies of the blurred image. We consider this as a major contribution which could be extended to wider areas of research integrated with multiple disciplines such as demosaicing.
List of Acronyms

1D       One Dimensional
2D       Two Dimensional
BCCB     Block Circulant Circulant Block
BTB      Block-Toeplitz-Block
CFA      Color Filter Array
CLS      Constrained Least Squares
CMA      Constant Modulus Algorithm
DFT      Discrete Fourier Transform
EM       Expectation Maximization
FIR      Finite Impulse Response
FSE      Fractionally Spaced Equalizer
FOPDO    First Order Partial Derivative Operators
FSOPDO   First and Second Order Partial Derivative Operators
GTI      Ground Truth Image
ICA      Independent Component Analysis
i.i.d    independently and identically distributed
IRLS     Iterative Re-weighted Least Squares
ISI      Inter Symbol Interference
MAP      Maximum A Posteriori
MCR      Multi Channel Restoration
MIMO     Multiple-Input Multiple-Output
ML       Maximum Likelihood
MRE      Mutually Referenced Equalizer
MSE      Mean Square Error
MSSIM    Mean SSIM
PAM      Pulse Amplitude Modulation
PCA      Principal Component Analysis
PDO      Partial Derivative Operators
PSF      Point Spread Function
PSR      Perfect Source Recovery
rgb      red, green and blue components of a color image
SIMO     Single-Input Multiple-Output
SOPDO    Second Order Partial Derivative Operators
In laying out this thesis, some symbols have been reserved in their meaning such as $\sigma$ and others have been chosen to allow for easy mnemonic reference such as $n$ for noise, $g$ for ground truth image, as listed below. In chapters where the number of symbols proliferate uncomfortably, we have tried to reuse symbols, within a similar context as its earlier usage such as $i$ and $j$, used for indexes.

The lowercase letters represent two-dimensional matrices while the bold lowercase represent a vector, where a two-dimensional matrix can be converted into a vector by lexicographically ordering the elements if necessary. Capital letters represent the Fourier transforms of their lower case counterparts with the exception of $N$, which is used to represent the normal distribution. Capital letters in square brackets represent special matrices such as Block-Toeplitz-Block matrix or Block-Circulant-Block Matrix.
\( g \)  Original or ground truth image
\( b \)  Degraded image
\( n \)  Noise corruption
\( k \)  point spread function
\( \hat{g} \)  Estimate of ground truth image
\( \mathbf{g} \)  Vector of \( g \): stacks the columns of \( g \)
\( g(i,j) \)  A point represented by \( i \)th row and \( j \)th column in matrix \( g \)
\( \sigma_g \)  Standard deviation of \( g \)
\( \mu_g \)  Mean of \( g \)
\( G \)  Fourier Transform of \( g \)
\( [G] \)  Special matrices, such as the Block-Toeplitz-Block matrix of \( g \)
\( L_1 \times L_2 \)  Support of a matrix
Mathematical Operators

\[ \| \cdot \| \quad \text{Frobenius norm} \]
\[ \| \cdot \|_1 \quad \text{L1 norm} \]
\[ E \{ \cdot \} \quad \text{mathematical expectation} \]
\[ \mathcal{F}(\cdot) \quad \text{discrete Fourier transform} \]
\[ \partial_x \quad \text{first order derivative in } x \text{ direction} \]
\[ \partial_y \quad \text{first order derivative in } y \text{ direction} \]
\[ \overline{\cdot} \quad \text{complex conjugate} \]
\[ * \quad \text{element-wise product} \]
\[ \sum \quad \text{Sum over all elements} \]
\[ \prod \quad \text{Product of all elements} \]
\[ H^T \quad \text{Transpose of matrix } H \]
\[ H^{-1} \quad \text{Inverse of matrix } H \]
\[ \otimes \quad \text{Convolution operator} \]
\[ \log(\cdot) \quad \text{Natural logarithm} \]
\[ \arg_x \quad \text{Argument } x \]
\[ \max \quad \text{Maximum} \]
\[ \min \quad \text{Minimum} \]
\[ N(\cdot) \quad \text{Gaussian distribution} \]
\[ \triangleq \quad \text{Defined as being equal to} \]
\[ p(\cdot) \quad \text{Probability} \]
Contents

Declaration i
Acknowledgements iii
Abstract v
List of Acronyms ix
Notations and Symbols xii
List of Figures xvii
List of Tables xxiii

1 Introduction 1
  1.1 Motivation and Background ................................ 1
  1.2 Aim and Scope ............................................... 3
  1.3 Thesis Structure ........................................... 4
  1.4 Thesis Contributions ....................................... 5

2 Image Restoration 11
  2.1 Overview ..................................................... 11
  2.2 Image formation ............................................ 12
    2.2.1 Image capturing ........................................ 13
    2.2.2 Image modeling ......................................... 14
  2.3 Degradation .................................................. 14
    2.3.1 Model .................................................. 14

xiii
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.2</td>
<td>Constraints</td>
<td>16</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Degradation Sources</td>
<td>16</td>
</tr>
<tr>
<td>2.4</td>
<td>Blind Image Restoration</td>
<td>18</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Characteristics</td>
<td>19</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Classification</td>
<td>20</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Direct PSF Identification</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>Adaptation Based Restoration</td>
<td>24</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Inverse PSF Estimation</td>
<td>24</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Simultaneous Estimation of Image and PSF</td>
<td>26</td>
</tr>
<tr>
<td>2.5.3</td>
<td>PSF Identification through Adaptation</td>
<td>26</td>
</tr>
<tr>
<td>2.6</td>
<td>Non-blind Deterministic Restoration</td>
<td>27</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Linear Restoration</td>
<td>27</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Non-linear Restoration</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>Non-blind Stochastic Restoration</td>
<td>32</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Maximum Likelihood Restoration</td>
<td>33</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Maximum a Posteriori Restoration</td>
<td>34</td>
</tr>
<tr>
<td>2.8</td>
<td>Multichannel Image Restoration</td>
<td>35</td>
</tr>
<tr>
<td>2.8.1</td>
<td>Channel Identification</td>
<td>36</td>
</tr>
<tr>
<td>2.8.2</td>
<td>Direct Restoration</td>
<td>36</td>
</tr>
<tr>
<td>2.9</td>
<td>Summary</td>
<td>37</td>
</tr>
</tbody>
</table>

3 Regularized Image Restoration

3.1 Overview | 39   |

3.2 Quadratic Regularization in Image Restoration | 40   |

3.2.1 Regularization | 40   |

3.2.2 Regularization Operators as Components in Quadratic Stabilizing Functionals | 44   |

3.2.3 Selection of Regularization Parameter | 46   |

3.2.4 Non-blind Image Restoration through SOPDO | 47   |

3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models? | 51   |

3.3.1 Problem Formulation | 51   |

3.3.2 Regularization - Bayesian Interpretation | 53   |
3.3.3 Image Restoration Evaluation .......................... 55
3.3.4 Performance of Quadratic Regularization Operators ....... 58
3.3.5 Regularization model comparison .......................... 61
3.4 Likelihood Model Analysis .................................. 66
  3.4.1 Likelihood Models in Image Restoration .................. 66
  3.4.2 Frequency Domain Deconvolution ......................... 69
  3.4.3 Likelihood Model Analysis ................................. 72
3.5 Contributions ................................................ 73
3.6 Conclusion ................................................... 75

4 Image Deconvolution through CMA .......................... 77
  4.1 Overview ................................................... 77
  4.2 Image Deconvolution and channel equalization .............. 79
  4.3 CMA on Image Deconvolution ................................ 80
    4.3.1 Adaptive System ...................................... 80
    4.3.2 Blind Cost Functions .................................. 81
    4.3.3 Blind Error Functions .................................. 82
    4.3.4 Algorithms ............................................. 83
  4.4 Dispersion Analysis ........................................ 84
    4.4.1 Scaling Properties of the Dispersion .................. 85
    4.4.2 Effect of Dispersion on Adaptation .................... 85
  4.5 Separable Linear Deconvolution ............................. 88
    4.5.1 Direct Parametrization ................................. 88
    4.5.2 Separable Parametrization .............................. 89
    4.5.3 Dispersion under Separability ......................... 90
    4.5.4 Simulations ............................................ 91
  4.6 Contributions ............................................... 94
  4.7 Conclusion ................................................ 95

5 Analysis of CMA Blind Image Deconvolution in Practice .... 97
  5.1 Overview ................................................... 97
  5.2 Preamble ................................................... 98
  5.3 Statistical Distribution of Ground Truth Image .............. 100
6.5.2 Demosaicing .................................................. 150
6.6 Image Restoration with the Proposed Model .................... 152
   6.6.1 Color Image Restoration .................................. 152
   6.6.2 Comparison ................................................ 155
6.7 Contributions .................................................. 156
6.8 Conclusions .................................................... 157

7 Conclusions and Future Research Directions 159
   7.1 Conclusions ................................................... 159
      7.1.1 Novel algorithms ......................................... 159
      7.1.2 Analysis of existing algorithms .......................... 160
   7.2 Future Research Directions .................................. 161

Bibliography 163
List of Figures

1.1 Novel thesis contribution areas indicated in bold .......................... 6
2.1 Digital imaging ........................................................................... 12
2.2 Classification of blind image restoration ........................................ 23
3.1 Image results for a highly textured image ................................. 51
3.2 Image results for a highly textured image .................................. 52
3.3 Image restoration model, where $g$ is the ground truth image, $b$ is the distorted image, $\mathcal{K}$ is the blur operator, $\mathcal{L}$ is the deblur process and $\hat{g}$ is the estimated image. .................................................. 57
3.4 Simulated image restoration model .............................................. 57
3.5 Image restoration results with prior models ................................. 58
3.6 Image restoration results with quadratic regularization models ... 60
3.7 Image restoration results for simulations in Table 3.5 ............. 60
3.8 Image restoration model for a naturally blurred image, where $\mathcal{K}$ is the blur process, $\mathcal{L}$ is the deblur process, $g$, $b$, $z$, $\hat{g}$ stand for ground truth image, blur image, deblurred image with artifacts, and the final estimated ground truth image respectively. The process $P_M(z)$ decouples “Model” and “Process” artifacts from the deblurred image .................................................................................. 61
3.9 Image restoration results for the system in Fig. 3.8 with FFT processing of Gaussian Prior and IRLS processing of Sparse and Laplacian prior models .......................................................... 62
3.10 Image restoration results for a noisy image under the same environment in Fig. 3.9 ................................................................. 63
3.11 Noise maps for Likelihood models ................. 74

4.1 The principle of blind adaptive equalization, where the equalizer parameters are adjusted based on the objective function output signal. 78

4.2 Adaptive blind deblur model: x is the unknown source, y is the known distorted source, and z is the deconvolved source. ........ 81

4.3 Scaling properties of Blind Cost Function, \( J_p(\cdot) \), given in (4.4), with respect to scaling the dispersion by \( \beta > 0 \). ............... 84

4.4 Scaling properties of Blind Error Function, \( \phi_p(\cdot) \), given in (4.6), with respect to scaling the dispersion by \( \beta > 0 \). ............... 84

4.5 Diagram showing the effect of dispersion on blind adaptation. The “Reference System” is the upper portion and the “Scaled System” is the lower portion. The two \( W \) adaptive filters are identical systems and may differ only in the specific values taken on by the adapted weights. ................................. 86

4.6 Difference between the actual and estimated deblur kernels with the dispersion change ................................. 92

4.7 Difference between the actual and estimated deblur kernels .......... 93

4.8 Difference between the actual and estimated deblur and blur kernel convolution ........................................... 94

4.9 Lena original image, blurred image and a comparison between the classical CMA with direct parametrization and our novel CMA with separable parametrization. ................................. 96

5.1 Adaptive blind deblur model. ............................. 98

5.2 Natural images used for testing distributional aspects. .............. 102

5.3 Adaptive blind deblur model with histogram equalization as a pre-process. ........................................... 103

5.3 CMA deconvolution results for the images in Fig. 5.2. ............... 106

5.4 Classical CMA deconvolution of a natural image and a randomized natural image ........................................... 111

5.5 Adaptive blind deblur model with whitening as a pre-process. ... 113

5.6 CMA deconvolution results on spatial correlation tests. .............. 114

5.7 CMA deconvolution results on whitening and histogram equalization. 115
5.8 Log histograms of natural ‘Lena’ image \((C_x = 2.1759)\) and whitened ‘Lena’ image \((C_x = 11.4199)\). .......................... 116

5.9 Image correlation model. .................................................. 117

5.10 Deconvolution based on the model in Fig. 5.9 via classical CMA adaptation and minimum kurtosis adaptation for a randomized source. 119

5.11 Comparison of classical and minimum kurtosis CMA Deconvolution based on the model in Fig. 5.9 for randomized and whitened images based on the same natural ground truth image. ............... 120

5.12 Difference between the actual and estimated deblur kernels for \(515 \times 515\) image deconvolved via classical CMA adaptation and minimum kurtosis adaptation based on the model in Fig. 5.9 .................. 122

5.13 Comparison between the classical CMA adaptation and our novel segmentation based CMA adaptation. .......................... 124

5.14 CMA deconvolution results on zero-mean tests. ..................... 126

6.1 Single channel system ......................................................... 131

6.2 Two channel system .......................................................... 133

6.3 Multichannel system with upsampling factor \(L\) .......................... 134

6.4 Equivalent interpolation systems ........................................... 136

6.5 Equivalent decimation systems ............................................ 136

6.6 Polyphase decomposition of an interpolation system .................. 137

6.7 Multirate system for \(L = 2\) .................................................. 137

6.8 Polyphase decomposition of \(L = 2\) multirate system .................. 138

6.9 Discrete multirate system equivalent to system in Fig. 6.1. ............ 138

6.10 Multichannel image restoration ............................................ 141

6.11 Proposed model for blur-deblur process ................................ 145

6.12 Ground truth image model; Upsampling by factor \(L\) and interpolation with filter \(f\). ......................................................... 145

6.13 Multichannel image restoration with a single image ................... 147

6.14 Spatial arrangement of \(p'\) for the \((i, j)^{th}\) pixel of \(p\). ............... 150

6.15 Bayer CFA arrangement for red channel. ................................ 151

6.16 Channel layout notation used in equation (6.16) ....................... 152

6.17 Results of image restoration with the proposed model .................. 154
6.18 Results of image restoration with a sparse kernel. . . . . . . . . . . . 155
6.19 Sparse blur kernels used for the results in Table 6.1. . . . . . . . . . . 155

7.1 Novel multichannel image restoration model . . . . . . . . . . . . . . . 162
# List of Tables

3.1 Comparison of stabilizing functional model. .......................... 48  
3.2 Efficiency in non-blind image deconvolution .......................... 49  
3.3 Efficiency results on scaling .............................................. 50  
3.4 Choice of regularization parameter ($\lambda$) values for different quadratic regularization operators used in the simulations of Table 3.5 .......................... 59  
3.5 Performance of quadratic regularization operators under varying regularization parameter values .............................. 59  
3.6 Efficiency of regularization operators. The times taken for restoration of grey-scale and colored images are given in seconds for each of the regularization operators. .......................... 64  
3.7 An example of set $\Theta$ in (3.28) with $R = 5$ .......................... 69  
3.8 Comparison of likelihood models .......................... 72  
6.1 Performance of non-blind restoration techniques .......................... 156
Chapter 1

Introduction

1.1 Motivation and Background

Images obtained through various imaging devices are known to be degraded representations of the real images due to imperfections in the imaging and capturing process. For example, such degradations can occur as a result of lens distortion (experienced in Hubble Space Telescope [1]), de-focus blur [2] and blur due to relative motion between the subject and the camera during the photographic exposure [3,4]. The problem of recovering the original image from a degraded observed image is usually known as the problem of image restoration. As the problem of recovering an input signal from a blurred output is ubiquitous in science and technology, image restoration has been an active research area for a period of close to a half a century. Initiated with astronomical imaging, image restoration has been extended to a wide variety of applications including medical imaging. These increasing range of applications require more accurate and efficient deblurring algorithms making image restoration an area for continuous research.

In most situations, the imaging system can be considered as linear and space invariant and the image restoration problem can be described as an additive linear degradation model given by

\[ b(l_1, l_2) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} k(m_1, m_2) g(l_1 - m_1, l_2 - m_2) + n, \]
where $b$, $k$, $n$ and $g$ respectively represent the degraded image, blur kernel (in some cases referred as PSF, standing for point spread function), additive noise and the true undistorted image that would have been recorded, had there been no blurring in the image recording process. The support of the blur kernel, $k$, is given by $M_1$ and $M_2$, while the spatial locations of the image and the kernel are given by $(l_1, l_2)$ and $(m_1, m_2)$ respectively. Image restoration could be mainly handled as a blind deconvolution problem (task of restoring an image when the blur kernel is not known) or a non-blind deconvolution problem (restoring an image when the PSF is known).

Blind deconvolution is a mature discipline within the domain of digital communication systems, having a strong theoretical basis with high performing deconvolution algorithms [5–11]. The Constant Modulus Algorithm (CMA) is referred to as the most widely tested and used blind equalization technique [10]. The qualitative and quantitative differences in multi-dimensional signal processing, such as image restoration, demand attention before applying any equalization algorithm to image restoration. Encouraged by this demand and the high performance of the equalization algorithms, we carry out a comprehensive analysis of applying and extending CMA in blind image restoration.

The challenge in blind image restoration with a single image has been often met in the literature with additional application dependent constraints such as type, direction or size of the blur kernel [1, 2, 12]. Even with this explicit knowledge, single channel image restoration could not achieve perfect blind image restoration. This prompted researchers to examine the applicability of multichannel restoration for blind image deconvolution, which has been successfully achieved with multiple images of the same scene [13,14]. The benefits of multichannel image restoration is mostly restricted by not being able to produce multiple copies of the same scene. This simulated our research on achieving the benefits of blind multichannel image restoration with a single image.

In recent years, a number of effective ways have emerged to estimate the PSF in image restoration [3, 15–18], claiming that the problem of blind deconvolution can be handled better with separate PSF estimation and non blind deconvolution [19]. Thus, we devote a separate section on this thesis in analyzing and enhancing non blind image restoration techniques. In reviewing the literature as
detailed in Chapter 2, we found out that most of the recent contributions under the maximum a posteriori (MAP) restoration technique and regularization techniques in the field of image restoration are subjected to applying various enhancements in terms of likelihood and prior models [2, 4]. The lack of proper theoretical expositions of these novel models motivated us to develop the theoretical aspects of these techniques, and thereby to produce a recommendation guide as to the necessity of these techniques in implementing highly effective and efficient image deconvolution algorithms.

In addition, performance evaluation of image restoration techniques in past research has been mostly carried out through the use of mean square error (MSE) metric [19]. Though MSE is a successful measurement in other contexts such as data communication, recently it is shown that MSE is not suitable for evaluations which include content dependent variations in image fidelity [20]. The introduction of SSIM (Structured SIMilarity) index [21] stressed the need for an enhanced evaluation of existing image restoration techniques. This has been addressed in this thesis, taking different prior models and comparing the efficient linear deterministic algorithms to the performance of stochastic non-linear algorithms.

1.2 Aim and Scope

The purpose of the research undertaken in this thesis is to develop effective and efficient image restoration techniques that are build on recent innovations and are suitable for use in a wide range of applications. The areas identified as having scope for development are the efficiency of the blind and non-blind image restoration algorithms and novel blind deconvolution algorithms.

In achieving these objectives, we first consider non-blind image restoration in line with the recent advances in blur kernel estimation [3, 15, 18, 22, 23], which indicate that blind deconvolution should be performed as two steps, namely,

1. Estimate the blur kernel and

2. Given the estimated blur kernel, restore the blurred image using a non-blind deconvolution algorithm.
In analyzing non-blind image restoration, our goal is to provide theoretical expositions to the aspects of non-blind image deconvolution, which have been overlooked in much research. In doing this, we also develop novel measures for the performance in image restoration, which are used for proving that certain stochastic non-blind image restoration techniques do not perform superior to the efficient deterministic algorithms.

In terms of blind deconvolution, a variety of deconvolution algorithms have been developed in the past and some of them have been performing remarkably well. As most of these deconvolution algorithms have originally catered for equalization in the context of data communication, we explore the potential of applying and extending these algorithms to image restoration.

In addition, we consider multichannel restoration as one of the areas, which has the potential for growth but with the constraints of having to use multiple images of the same scene and the registration of those multiple copies. In complementing this consideration, we focus on design and development of a novel multichannel restoration model and present new opportunities that have previously been unattempted. With this model, while enjoying the benefits of multichannel restoration, the restoration algorithms get the luxury of performing with a single blurred image.

### 1.3 Thesis Structure

This chapter, which presents an overview of the thesis is followed by more detailed chapters with novel contributions.

**Chapter 2** provides an extensive literature review starting with the fundamental theories behind image restoration. The fundamental theories are constantly and continuously emphasized as they form the basis of the novel contributions detailed later on. In addition to the analysis of the existing methodologies, in chapter 2 we build up our novel classification of the most prominent techniques in image restoration.

**Chapter 3** is dedicated to the analysis of regularization in image restoration. Although the concept of regularization has been widely applied in image
restoration for the last few decades, we show the necessity of a proper analysis, as some recently developed methodologies do not contribute to high performance in spite of their complex regularization models.

**Chapter 4** explores and extends the application of CMA in image deconvolution. A broad class of image deconvolution algorithms are developed in this chapter resulting in more efficient algorithms than the conventional CMA algorithm.

**Chapter 5** analyzes the performance difference of CMA in image deconvolution to the performance under well established theories in communication context. The qualitative differences of deconvolution in image processing context and communications context lead to a comprehensive study of the assumptions underlying CMA deconvolution and development of novel techniques to overcome the effects of those assumptions.

**Chapter 6** is devoted to the investigation of multi channel image restoration (MCR) algorithms. As the promising results in MCR are always achieved through the use of multiple images in the literature, we develop a novel image restoration model, which could be utilized to achieve the benefits of multichannel processing through a single image.

**Chapter 7** concludes the thesis elaborating on recommendations and future research directions.

### 1.4 Thesis Contributions

The novel contributions of this thesis are demonstrated through the classification illustrated in Fig. 1.1. This classification itself is one of our contributions, for which the characteristics and the development details are furnished in Chapter 2. Further, this classification together with the analysis, provided the basis for the identification of the major contribution areas. Each of the contribution areas are summarized below, followed by the list of journal and conference papers we have prepared under that area. The blocks with bold letters indicate our novel contribution areas.
Figure 1.1: Novel thesis contribution areas indicated in bold
• Non-blind Image Restoration - Our motivation for research in non-blind image restoration techniques is encouraged by the recent research [24] and emerging effective PSF estimation techniques as detailed in Chapter 3. Under deterministic restoration methods, we analyze existing linear restoration models and develop a class of novel models with better performance. Using regularization as the basis, we link linear deterministic models and stochastic restoration models. By introducing a novel visual metric for comparison of regularization models in image restoration, we demonstrate that stochastic prior models do not result in a superior restoration over linear deterministic prior models. In addition, we show that the high complexity derivative likelihood models under the maximum a posteriori (MAP) framework offer no advantage to a properly configured normal likelihood model.


• Inverse PSF Estimation - The characteristics shared by the channel equalization in data communication context and image restoration in image processing context led the research to estimate the inverse PSF in image restoration through the well established equalization algorithms. The proven record of being one of the best performed equalization algorithms [25] and the broad applicability, makes CMA a natural choice for image restoration. Taking CMA as the basis, in Chapter 4, we develop a broad class of blind image deconvolution algorithms, in particular algorithms for blurring kernels with a separable property, which show significantly faster convergence than conventional algorithms. The properties of the dispersion parameter, are also analyzed, as they play an important role in blind image deconvolution algorithm cost functions, and scaling properties and optimal values are derived. The extension of blind image deconvolution through CMA is further carried out in Chapter 5. An exhaustive analysis in this chapter shows that the
effects of the assumptions on the source, which were negligible in channel equalization, carry a significant weight on the performance of image restoration through CMA. In this analysis, we show the impact of each of the source assumptions, namely, source of uniform distribution, white and being zero mean on the performance of blind image restoration through CMA. In addition, we develop novel techniques based on higher order moments of the source distribution and other mechanisms to overcome or lessen the effects of source assumptions on the performance of blind image restoration through CMA.


- Multi Channel Image Restoration - Inverse PSF estimation in image deconvolution with a single image has always been challenging, as the support of the deblur kernel tends to be large for a small support PSF. In literature, this challenge has been addressed by the use of multiple images of the same scene acquired through different sensors or at different times. As producing
multiple images of the same scene is not realistic in most of the consumer imaging systems, we develop a novel image restoration model to process multi channel image restoration with a single image. With our novel model, we demonstrate that perfect image restoration could be achieved with a single image. Furthermore, we show the practical applicability of our new model with demosaicing.
Chapter 2

Image Restoration

2.1 Overview

As humans, the impact of information gathered through our eyes and reasoned through the visual cortex dominates the effect of any of our other senses. Thus, images and ‘imaging’ play a central role in the existence of mankind. The term ‘imaging’ here refers to acquiring and processing of images in the digital domain and ‘imaging’ generally introduces many kinds of distortions to the final result.

“Digital image restoration is a field of engineering that studies methods used to recover an original scene from degraded observations” [1]. With it ongoing in astronomical imaging, image restoration has been extensively studied for its practical importance, expanding into a number of areas such as medical imaging, surveillance, forensics, weather forecasting, and resource exploration. Based on the literature, image restoration began with the efforts of scientists from United States and the Soviet Union in the 1950s and 1960s [1]. In these programs, the scientists were taking photographs of the Earth and the solar system, but these images were subject to many degradations. Restoration at the time was highly expensive. Scientists began to look for less expensive methods to perform image restoration.

Since then, restoration has been an active area of research, where remarkable advances in techniques throughout the years have been made. Our goal in this chapter is not to provide an exhaustive survey on the techniques developed during
the last half a century period, but to stress the fundamental building blocks, which are essential in developing novel algorithms and analysis in the chapters to follow. As restoration techniques require some knowledge of imaging process and degradation, either in the form of analytic models or as priori information, a considerable emphasis in the first part of the chapter is placed in understanding imaging and degradation models. The latter part of the chapter covers a glimpse of the plethora of image restoration techniques with a novel classification, which integrates recently developed as well as the matured algorithms.

2.2 Image formation

Image formation, in other words, recording a visual scene, has come a long way and in this digital era, the process of capturing a scene as a digital image takes only few seconds. Understanding the processes and interpretations behind image formation is crucial for image restoration as improvements rely on the underlying processes. In addition, the knowledge of image formation and what constitutes a valid solution helps the restoration process by imposing various constraints on the form of the solution. Furthermore, the contents of this section become compelling as novel techniques such as multichannel Restoration can be developed with this understanding, as detailed in Chapter 5.
2.2 Image formation

2.2.1 Image capturing

As shown in Fig. 2.1, a natural image is a representation of a natural scene on a domain of continuous support. When discretized, a sampled version of this natural image is represented by the band limited image \( g \). Image capturing process can be thought as a neighboring process, where a point \((x, y)\) of the recorded image \( g \) is dependent not only on the point \((x, y)\) of the natural image, but the neighborhood points surrounding point \((x, y)\).

The principle aim of imaging is to record an accurate representation of an object or scene on a rectangular grid for analysis and interpretation. A digital image is typically formed by quantizing the spatial position and corresponding intensity value with sufficient resolution to provide an accurate representation. This can be achieved by digitizing an existing analogue recording or by more direct means such as a charge coupled device camera which samples the scene directly using an array of photosensitive solid-state devices. The intensity at each point is recorded by detecting particles normally either photons or electrons emitted by, reflected from or transmitted through an object. The object may emit light as a star or galaxy does or require external illumination by electromagnetic radiation such as visible light or X-rays or by an electron beam. Intensity resolution is limited by the quantization imposed by individual particles. The imaging sensor may also have a bandpass intensity response so that only a selected region of the frequency or energy spectrum is sampled.

Once the natural image is sampled and quantized, each sample point is known as a pixel (for picture element). Then the pixel at \(i\)th row and \(j\)th column in a matrix \( g \) is represented as

\[
g(i, j) = \mathcal{SQ}\{\text{im}_{\text{nat}}(x, y)\},
\]

where \(\mathcal{SQ}\{\cdot\}\) represent the sampler and quantizer for a natural scene \(\text{im}_{\text{nat}}\) and the pixel at \(i\)th row and \(j\)th column corresponds to a sample at the position of \(x\) horizontal and \(y\) vertical in the continuous space [26]. From this point onwards, we refer \(g\) as the natural image or the ground truth image (GTI), representing the natural scene \(\text{im}_{\text{nat}}\).
2.2.2 Image modeling

An image model is an abstraction of the structure or nature of a class of images. Although research use two image models, statistical and deterministic [26, pp.42-58], in this thesis, we follow the deterministic image model to represent the natural images, as they assume only the fundamental nature of the images. Deterministic models generally impose constraints upon the general class of sampled images such as

1. Non-negativity - Image pixel values being either zero or positive is one of the most powerful constraints imposed. The number of photons or electrons detected cannot be negative, so the image matrix $g$ is positive at all points.

2. Boundness - A digitized image is of finite support and carries a finite energy. Thus, the images we use in this thesis are represented by bounded finite matrices.

We will employ these general constraints in our image restoration formulations to follow.

2.3 Degradation

Out of the limitless possible causes of image degradation, some of the more common degradation mechanisms include diffraction effects, atmospheric turbulence, out-of-focus and non-stationary objects. Describing the processes and equations of image degradation is fundamentals before undertaking improvement by restoration. Hence, this section covers the degradation model we intend to apply in this thesis with the details of constraints related to the model.

2.3.1 Model

In the literature, a variety of degradation models are postulated, that contribute to the recording of a distorted image. Although there are non-linear degradation models, where the distortion becomes a function of the GTI itself, such as X-ray images, in this work, we will resort to linear models, which are reasonable to model
2.3 Degradation

many distortions in photography. The simplest model takes the form of a result of two phenomena, namely, degradation due to image acquisition or defects of the imaging system and distortion due to random noise. Mathematically, the model accounting for these degradation can be represented as,

\[ b = Kg + n, \]  

where \( b \) is the distorted image, \( K \) is a linear operator and \( n \) represents the additive random noise. In image restoration, \( K \) is mostly referred as the point spread function (PSF) or blur kernel, and is represented by \( k \) in this thesis. An ideal imaging system would have a delta function as the PSF, which however is not realistic.

The operator \( K \) can be mainly in two forms. When the linear operator changes the spatial position as well as the function, it is regarded as a space variant PSF, while if the operator only changes position, it is called the space invariant PSF. A space invariant model or sometimes referred as a stationary model for the blur, is commonly motivated due to difficulties encountered in handling a spatially varying blur [1]. With space invariant PSF model, the degradation could be represented as a convolution between the GTI and the PSF, mathematically denoted as

\[ b(l_1, l_2) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} k(m_1, m_2) g(l_1 - m_1, l_2 - m_2), \]  

where \( g \) and \( b \) are of finite support \([L_1, L_2]\), the PSF of support \([M_1, M_2]\) and \((l_1, l_2) \in [0, L_1 - 1] \times [0, L_2 - 1]\) [26, pp.65-67]. The work in this thesis is based on the space invariant PSF model and it is also assumed that the statistical properties (mean and correlation function) of the image and noise do not change spatially.

Although the most meaningful and realistic model for the captured image \( b \) in both physical and practical sense is the continuous-discrete system [26, pp.67], for ease of algorithmic evaluations, we use the discrete-discrete model, where the recorded image is represented by (2.3). The notation in (2.3) may be deceptive and sometimes misleading. From physical intuition, we could see that even though \( k \) has a potential of \( L_1 \times L_2 \) degrees of freedom in (2.3), there are much more points in a natural image in continuous space. The finiteness of \( g \) imposes image
capturing effects on $b$, which could be overcome by forcing several constraints such as making boundaries to be zero or the boundary values to be the reflection of the input image. A detailed discussion of this is presented in Chapter 3, where we investigate the effects of $g$ being finite on restoration techniques developed in current research.

### 2.3.2 Constraints

In developing the above model, the following factors are taken into consideration.

1. Non-negativity – In the linear degradation model we use, $b$, PSF and GTI are all taken as non-negative quantities. This is based on the fact that light intensity in imaging is always non-negative, leading PSF also to be componentwise non-negative [26, pp.68]

2. Bounded restoration - This constraint is the result of the boundness property in the imaging model, detailed in Section 2.2.2.

3. Unity gain imaging - With the concept that the total energy in GTI is preserved in $b$, this implies that imaging system does not absorb or generate optical energy. The ramification of this assumption results in

$$
\sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} k(m_1, m_2) = 1.
$$

### 2.3.3 Degradation Sources

Out of the various sources of degradation in imaging systems such as temporal degradations, chromatic degradations and spatial degradations, in this thesis, we consider only spatial degradations and a few commonly encountered spatial degradation types are briefly discussed below.

**Motion blur**

Motion blur is the result of camera panning [18] or fast object movements [27] and is the most common degradation type found in the consumer photography.
Camera shake is mostly experienced with the popularity of small high-resolution cameras, whose light weight make them difficult to be held sufficiently steady, or in the case of insufficient light, where a slow shutter speed is unavoidable. Traditional signal processing formulations were used to handle simple motion blurs, such as a translation, where the scene to be recorded translates relative to the camera at a constant velocity and could be modeled as a one dimensional blur [1].

In reality, camera motions could follow convoluted paths involving translation, rotation, a sudden change of scale, or to some combinations of these. In addition to the improved deconvolution methods to handle these complex blurs [3], the improvements of camera and sensor technologies in modern cameras have also contributed in motion blur restoration by addressing the camera panning and tilting with optically stabilized lenses [28]. Although the deconvolution algorithms developed in this work are not restricted to motion blur degradations, most of our simulations have been carried out on naturally or synthetically motion blurred images.

### Out-of-focus blur

Degradation due to defocus is the result of spreading a point of incoming light across a circle of confusion. In other words, on the image formed by an optical system such as a convex lens, objects at a particular depth from the lens will be focused whereas objects at other distances will be blurred by varying degrees depending on their distance. The focusing system of the camera depends on a range of parameters such as focal length, aperture size, object depth, wavelength of incoming light and the effects due to diffraction [29]. Due to the fact that out-of-focus blur is dependent on wavelength, the three color channels (red, green, blue) will have different blur functions [30]. As a parametric blur description, if the degree of defocusing is large relative to the wavelengths considered, the out-of-focus blur can be approximated by a uniform circular disk [1], which corresponds to the circle of confusion. Restoration of defocused images were mostly carried out in the literature by changing the imaging system [2,31] or using multiple images [29].
Atmospheric Turbulence Blur

Long-term exposure to atmospheric distortions caused by optical turbulence and scattering and absorption by particulate results in a degraded image. This type of degradation is commonly experienced in remote sensing and aerial imaging [1]. Although the blur introduced by atmospheric turbulence depends on a variety of factors (such as temperature, wind speed, exposure time), for long-term exposures the point-spread function can be described reasonably well by a Gaussian function given by

\[ k(i, j) = \beta \exp \left( -\frac{i^2 + j^2}{2\sigma^2} \right), \]

where \( \beta \) is a normalizing constant which satisfies lossless imaging and \( \sigma^2 \) is the variance that determines the severity of the blur. The separability of this type of blur helps in efficient restoration implementations.

2.4 Blind Image Restoration

As mentioned earlier, image restoration is the estimation of the unknown GTI, given the distorted image. The techniques used to achieve this are mainly twofold. Firstly, using hardware techniques, where the attempt to remove the distortion is made before the image is recorded on the sensor [18, 23, 27, 32]. Though these techniques show impressive results, they are often found to be expensive, requiring accurate knowledge of the distortion process and most of the time are specific to one type of degradation. On the other hand, software solutions are more flexible and over time a plethora of techniques have been introduced to handle image degradation. Few research have also been on combining hardware and software approaches together to achieve better results [2]. Unless otherwise stated, we mainly focus on the software approach and in this section, we stress on a few important points from the extensive literature.
2.4.1 Characteristics

Image restoration can be characterized as a problem with the following characteristics [26, pp.114-116].

1. singular – In most realistic distortions, the inverse transformation of the distortion does not exist. This is experienced when the blur filter is an ideal low pass filter or when the Fourier representation of the blur filter has zeros at selected frequencies [30],

2. ill-conditioned – As mentioned earlier in Section 2.2, in reality, image formation should be described in the continuous infinite dimensional space. In this case, the restoration problem becomes a Fredholm integral equation of the first kind, which is ill-posed [33]. The ill-posed nature implies that inherent bounded data perturbations would cause undesirable unbounded effects in the solution. Once discretized, the ill-posed problem results in an ill-conditioned matrix representing the PSF [1]. Due to this behavior, usually trial and error methods cannot be applied on image restoration as a small observation error would not ensure a small error on the solution.

3. non-unique – The solution is not unique. This situation arises when the discrete representation of $\mathcal{K}$ has a nonempty null-space (where the null-space represent the collection of all images that produce zero output) resulting in unobservable data. An example would be the set of DC or constant images when $\mathcal{K}$ is a high-pass filter.

As image restoration is ill-conditioned, it lacks a unique solution, leading to a diverse range of techniques. Thus, this section would not provide a comprehensive classification of image restoration, which is too ambitious, but would establish some common benchmarks or points of reference.
2.4.2 Classification

Existing Image Restoration Classifications

Before diving into the depth of image restoration classification literature, we would like to mention few general classification terms, which are commonly found in the image restoration community. Broadly, image restoration can be divided into blind and non-blind image restoration, where restoring an image with only the distorted image is termed as blind image restoration, while the non-blind case refers to the restoration, when an estimate of the blur kernel is available. Each of these restoration methods can be further classified as spatially variant and spatially invariant, where spatially variant techniques accommodate distortions due to spatial variant PSF, as discussed in Section 2.3.1.

Out of the many classifications in the literature, [1] contains a substantial amount of information related to image restoration algorithms that grow out of regularization methods. In terms of classification, they have mainly concentrated on non-blind image restoration with the following structure.

1. Direct Approach
   
   (a) Stochastic Approach
   
   (b) Deterministic Approach

2. Iterative Approach

3. Recursive Approach

The popular Wiener filter (discussed in Section 2.6.1) falls under the stochastic approach, while the constrained least squares filter (discussed in Section 2.6.1) comes under deterministic approach. The main drawback of this classification is that it covers only a segment of the image restoration field.

Kundur and Hatzinakos have addressed the classification of blind image restoration in their paper [34]. Under their classification, zero sheet separation has drawn much attention with the argument that multidimensional degraded signals are deconvolvable based on the analytic properties of the z-transform in multiple dimensions. Although conceptually the zero sheets are correct, later it is found out that
they have little practical application since the algorithm is highly sensitive to noise and prone to numerical inaccuracy for large image sizes [35]. Further in the image restoration classification in [34], blur identification methods supporting only simple parametric blur and restoration methods supporting deterministic constraints are discussed. Though [34] furnishes an extensive classification on blind image restoration, most of the techniques mentioned in that have evolved with time and some have been categorized as non-practical.

Thus, the literature justifies the necessity for a proper classification on image restoration, incorporating previous successful techniques and novel methods introduced in the recent past.

**Blind Image Restoration Classification**

There is a wide array of image restoration algorithms developed up to date and with the growing field of research, new methods are constantly being developed. Thus, an exhaustive treatment of the available algorithms is beyond the scope of this section, but we have tried to provide a summary of successful algorithms by classifying them on their deblurring techniques. We classify the deblurring techniques so far developed in the literature of image restoration under the assumption that only the blurred image is given.

Unlike previous classifications, our classification, shown in Fig. 2.2, attempts to incorporate the latest successful algorithms developed in the field of image restoration and we purposely omitted mentioning a number of algorithms from the past, which were not performing satisfactorily in the evaluations. In analyzing the literature, as we noticed some recent developments have incremental contributions to restoration, in the form of changing the cost function or introducing a combination of regularization options, those methods are also avoided in the discussion below.

Further, the emphasis in our classification is placed primarily on methods related to signal processing and computer vision compared to other approaches such as medical imaging, optics and astronomy.

Our classification of image restoration is based on the blind image restoration, as we believe that estimating the PSF is the first part of non-blind image restoration. The main advantage of the techniques in non-blind restoration class
is their low computational complexity. A large number of blind deconvolution algorithms fall under the adaptation based restoration, where the estimation of PSF, image or the inverse PSF are carried out adapting on continuously varying data. Recently, few methods are developed in simultaneously estimating the PSF and the GTI, where in each estimation step, one of the variables to be estimated (PSF or GTI) is fixed. Estimating an inverse PSF is similar to finding an equalizer in communication systems, where the convolution of the inverse filter and the degraded image produces the estimated GTI. The appeal of investigating multi-channel image restoration follows from the diverse range of applications which produce multiple degraded images of the same scene and the fact the exploitation of cross correlations between channels ensure perfect source recovery.

When the PSF is estimated through direct methods or by adapting on the available data, it is used for non-blind image restoration. Non-blind image restoration techniques have been categorized in number of ways in the literature. Our categorizations attempts to incorporate most of the successful techniques, which include traditional as well as novel methods. Under non-blind deterministic restoration techniques, although the early classifications included recursive approaches such as Kalman filtering [1, 30], we do not address those techniques in our classification, as they are beneficial mostly on spatially adaptive techniques, and significant improvements in applying recursive techniques are not evidenced in the recent literature.

2.4.3 Direct PSF Identification

Identifying PSF in astronomical blind image restoration is considered to be straightforward as an isolated star could be used to reveal the PSF, whereas PSF identification with motion and out of focus blur is much more complex. Early approaches for identification of the motion blur involve methods in which identification is performed separately from the restoration process [36] and were applied for mostly simple parametric blur, where the estimation of missing parameters identifies the blur [30, 37]. Generalized cross validation is yet another method used to identify simple parametric blur kernels through estimation and validation [38]. However, recently more advanced methods have been developed to address the identification
Figure 2.2: Classification of blind image restoration
of two dimensional (2D) blur, such as [22], based on transparency maps and [15] based on the sharp edges of the blur image, which accommodates different type of blurs spatially variant or invariant. The list of PSF estimation techniques also includes enhanced techniques of searching zero crossings of the Fourier domain blur function [39, 40] and estimating the length and angle of the motion blur [16, 17].

In addition, numerous hardware techniques also have emerged in estimating the PSF in consumer photography [18, 23, 41], where [41] and [18] attach a low-resolution video camera to a high-resolution still camera to help in recording the blur kernel. While some techniques, such as [27], which modifies the kernel by minimizing the loss of high spatial frequencies, do not directly contribute to PSF estimation, it helps in getting a better PSF estimation. With these effective ways in estimating the PSF directly, we believe that non-blind deconvolution plays a major role in blind image restoration and image estimation becomes the second step of blind image restoration.

2.5 Adaptation Based Restoration

2.5.1 Inverse PSF Estimation

The fundamentals in inverse PSF estimation for image restorations are mainly borrowed from the field of data communication systems. In data communications, digital signals are generated and transmitted through an analog channel to the receiver. Analog media such as telephone cables and radio channels typically introduce distortion to the transmitted signal. This linear channel distortion, commonly known as the inter-symbol interference (ISI), can severely corrupt the transmitted signal and make it difficult for the receiver to recover the transmitted data. Channel equalization is the principle of removing this ISI [42, 43] and an equalizer is the inverse system for the channel distortion system.

Traditional channel equalizers often adopt non-blind channel estimation, where a training signal is used to identify the channel using mean square error criterion, which is then used to design an inverse system [42]. In contrast, blind channel equalization does not need the use of a training signal, which is costly in terms of bandwidth in data channels, but using other mechanisms such as higher order
moments of the channel, directly estimates the equalizer. By applying the equalizer, estimated either through non blind or blind techniques, to the channel, the original data can be recovered.

Unlike the non-blind estimation, in blind estimation, the minimization criteria, often referred as the objective function can take a variety of forms. Generally, this function is chosen to quantify the amount of distortion present in the equalizer output by measuring how altered some known features or characteristics of the source signal have become through the transmission and reception process. The equalizer parameters are then tuned according to this function to ultimately minimize the distortion in the received data. One of these objective functions, independently developed by Godard [6] and Treichler [7], named as constant modulus algorithm (CMA), has shown better performance when compared with other existing equalization algorithms.

The leap of faith, espoused by these well established equalization techniques, guided researchers on applying them to blind image restoration. Vural and Sethares in their analysis [44, 45] showed how CMA is extended to two-dimensional blind deconvolution applying to blind image restoration. We continue this analysis and perform an exhaustive evaluation of blind image restoration through CMA in Chapter 4 and Chapter 5.

One of the critical factors to be considered in single channel equalization is the support of the equalizer, which tends to be large, for a blur kernel of small support. This drawback has been overcome in communication systems by using single input multiple output (SIMO) systems through fractionally spaced equalization [46]. In addition, while CMA and other single channel equalization methods are based on the higher order statistics of the channel, under SIMO model, direct blind equalization becomes possible using only the second order statistics, which potentially may have faster convergence rates [47]. Borrowing these results and extending to image restoration, it is shown in [13, 14] that perfect image restoration could be achieved through multichannel blind image restoration. Multichannel restoration is explained in detail in Section 2.8.
2.5.2 Simultaneous Estimation of Image and PSF

Various attempts have been made on estimating the GTI and the PSF simultaneously by imposing constraints such as on the image and PSF in an alternating way. These techniques not only decouple the non-linear complex blind deconvolution to two linear simple models, but also overcome the impossibility in tracking the properties of the PSF and the image simultaneously, leading to fast restoration algorithms. In executing, such approaches cycle between the estimation of image and the PSF, where in the image estimation step, the image is estimated assuming that the PSF is fixed to its last estimate and in the PSF estimation step, the PSF is estimated assuming the image to be fixed to its last estimate. It is believed that by allowing the image and PSF estimation techniques to interact, local minima convergence could be avoided.

Alternative image and PSF estimation has been addressed in numerous ways in the literature, such as through regularization [48, 49] and probabilistic techniques [4]. In [50], a general iterative model is presented, where various image and Fourier domain constraints can easily be incorporated.

In [4], under maximum a posteriori (MAP) technique, likelihood is modeled with higher order derivatives and two prior models, global and local for piecewise smoothness were applied. We discuss MAP technique and the applicability of the higher order likelihood model in more detail in Chapter 3. The optimization energy in [4] is varied overtime and in terms of implementing, it is simplified by decoupling the estimation of PSF and image, which allowed to use fast Fourier domain techniques in some points of estimation. It is shown in [19] that edge re-weighting and iterative update of the likelihood weight have contributed [4] to overcome the delta (no blur) solution of the MAP technique. Under regularization approaches, [49] alternatively minimize restoration cost function regulated through total variation. The total variation regularization is discussed in more detail in Section 2.6.2.

2.5.3 PSF Identification through Adaptation

This can be classified as a sub-category of estimating image and PSF in alternating fashion detailed in Section 2.5.2. The techniques belonging to this class adapt the PSF estimation by imposing strong prior assumptions on the blur kernel and/or the
image. Although traditional PSF identification formulations were based on general assumptions such as positivity of the image, recently it is shown by exploiting stronger priors such as natural image statistics, more complicated blur kernels could be identified [3].

In [3], based on the heavy tailed natural image priors [51], a sparsity prior, modeled by a mixture of zero-mean Gaussians was applied. Based on the variational Bayesian approach [52] to approximate the posterior distribution, the parameters of the distribution of the expected PSF and the gradient of the GTI were updated alternatively through ensemble learning [53] until convergence. This method has proven better results on blur estimation using non-parametric kernels when compared to other existing techniques [24].

2.6 Non-blind Deterministic Restoration

2.6.1 Linear Restoration

In our classification, as we consider a class of imaging system in which the spatial degradation is modeled by a linear space invariant impulse response, using the property of superposition, restoration of those images can be performed by linear filtering techniques. Few of these techniques are discussed below under the classification of “linear” restoration techniques.

Pseudo Inverse Filter

Due to the characteristics of the image restoration problem detailed in Section 2.4.1, the naive method of attempting deconvolution with the inverted PSF is not practical. Hence, numerous variations to inverse filtering have been suggested in the literature and out of those, the Pseudo inverse filter is one of the simplest techniques. This is based on the least-squares framework by defining an error metric to be minimized and then finding a solution via an optimization technique. Mathematically, find an estimate of $g$, denoted as $\hat{g}$, which minimizes the norm

$$\|b - [K]\hat{g}\|^2,$$
where $\| \cdot \|$ is the Frobenius norm and $[K]$ has a Toeplitz structure with Toeplitz blocks (BTB). When the circulant embedding is applied to BTB structure, the corresponding matrix will then possess a block-circulant structure with circulant blocks (BCCB) [26, pp.211-224]. The significance of the BCCB structure is that there exist efficient computational approaches such as frequency domain techniques, which could be applied in solving the above minimization problem. As it is verified in [54], that as the dimension of the matrix grows, a block-Toeplitz matrix is asymptotically equivalent to a block-circulant matrix in the Euclidean norm sense, we use the terms block Toeplitz and block circulant interchangeably.

Under general degradation mechanisms, the kernel $k$ exhibits low pass characteristics, leading the pseudo inverse filter to enhance high frequencies in the image. This over-fitting of data results in large deviations in the solution, even for a small error in the data.

**Wiener Filter**

The Wiener filter is a linear filter, which could be considered as an extension of pseudo inverse filter by making use of prior knowledge about the covariance matrices of the original image and noise to overcome the issues related to over-fitting data [55]. The Wiener filter in Fourier domain $W$, is represented by

$$W = \frac{K}{|K|^2 + \frac{S_{nn}}{S_{gg}}}.$$ 

where $K$ is the frequency response of the blur kernel $k$ and $S_{nn}$ and $S_{gg}$ are the power spectrum of the noise and the original image, respectively. While the noise power spectrum $S_{nn}$ is generally assumed to be flat for independent white noise, estimating the image spectrum $S_{gg}$ has always been a challenge and has been one of the main limiting factors in applying Wiener filter in real image restoration problems.

**Constrained Least Squares Filter**

To address the ill-conditioned nature of inverse problems, Tikhonov and Arsenin [33] introduced the concept of regularization, which formed the basis for
a class of improved deconvolution algorithms. Regularization is conceptually simple and can be used as a unifying framework for image restoration as it can be interpreted in Bayesian terms as well. The purpose of regularization is to allow the inclusion of prior knowledge to stabilize the solution in the face of noise and constrain the restoration to physically meaningful and reasonable estimates. Adding the regularization function \( \Phi(\hat{g}) \) and the positive Lagrange multiplier, \( \lambda \) to the pseudo inverse filter defined in Section 2.6.1, results in minimizing the functional

\[
\|b - [K]\hat{g}\|^2 + \lambda \Phi(\hat{g}),
\]

where the term \( \|b - [K]\hat{g}\| \) defines the class of solutions where the norm of the residual image is bounded. The regularization parameter, \( \lambda \), controls the balance between noise artifacts and consistency with the observed data. Out of a wide class of different regularization functionals \( \Phi(\hat{g}) \) available, a general \( \Phi(\hat{g}) \) takes the form of

\[
\Phi(\hat{g}) = \|D\hat{g}\|,
\]

(2.4)

where \( D \) is known as the regularizing operator. A similar approach was introduced by Miller in [56], where he added an additional constraint on the regularization functional. Applying Tikhonov-Miller procedure to image restoration results in

\[
\hat{g} = ([K]^T[K] + \lambda(D^T D))^{-1}[K]^T b.
\]

When the regularization operator \( D \) belongs to the class of quadratic regularization functionals, it becomes a deterministic regularization technique and is known as the constrained least-squares method [57]. When \([K]\) and \( D \) have circulant structure, these equations are diagonalized by the discrete Fourier transform (DFT) and the problem can be easily solved in the frequency domain. The properties of the regularization operator with improved algorithms based on regularization are discussed in more detail in Chapter 3.
2.6.2 Non-linear Restoration

A linear algorithm is able to produce a result directly from the observed data but requires some form of inverse operator to be applied. By contrast, a nonlinear technique does not explicitly implement the inverse, instead it uses an iterative approach to produce successive improvements to the restoration until a termination condition is reached. Iterative deconvolution algorithms often conduct a extremum (minimum or maximum) on an error metric, where the error represent the violation of the imposed constraints with the estimated. Such algorithms are claimed to have better results in noise suppression. Further, in non-linear restoration, it is possible to introduce spatial adaptivity to impose constraints such as non-negativity and recover missing spatial or frequency components at the expense of efficient implementations and local convergence.

Least Squared Iterative Restoration

In image restoration, two frameworks have been developed for catering the use of priori information as an aid for the recovery: the deterministic formulation and the stochastic model. We discuss the former case in this section and stochastic restoration is discussed in Section 2.7.

Under iterative restoration, the iteration depends on

\[ \hat{g}_{r+1} = \hat{g}_r + \alpha \phi_r, \]

where, \( \hat{g} \) stands for the estimated GTI, \( \alpha \) represents a positive step size, \( r \) represent the iteration number and the function \( \phi \) may take different forms depending on the minimization criterion and optimization approach. For example, instead of solving the least squares criterion directly, it can be solved iteratively by using a gradient descent approach. Then \( \phi \) simplifies into

\[ \phi_r = [K]^T (b - [K]\hat{g}_r), \]

where \([K]^T\) denotes the transpose of \([K]\). This kind of iterative methods have a long history and have been named in different ways such as Landweber [58] and Van Cittert [59] iteration. These traditional iterative algorithms have un-
undergo numerous changes for improvement, such as re-blurring in Van Cittert algorithm [30].

The main limiting factors behind direct inverse algorithms such as Van Cittert and Landweber are the small eigenvalues of the blurring operator, which causes noise to dominate the final result. To overcome this issue, regularization was introduced to the iterative algorithms, which incorporates knowledge about the GTI and/or the PSF to the algorithm [30]. With the iterative version of the Constrained Least Squares filter detailed in Section 2.6.1, $\phi$ becomes

$$\phi_r = [K]^T b - ([K]^T[K] + \lambda D^T D)^g r,$$

and is named as regularized iterative Constrained Least Squares restoration [60].

**Total Variation Restoration**

The total variation (TV) regularization, proposed in [61] is another form of regularization and the optimization in this can be given by

$$\arg\min_g \|b - [K]g\|_2^2 + \lambda \|Dg\|_1,$$

where arg denotes the argument producing the minimum (as opposed to the value of the minimum itself). The purpose of regularization is to allow the inclusion of prior knowledge to stabilize the solution in the face of noise and allow the identification of physically meaningful and reasonable estimates. $\| \cdot \|_1$ is the $L1$-norm and $D$ generally takes the form of a gradient operator. Regularization with TV is preferred over the Tikhonov-Miller regularization as total variation regularized answers can contain localized steep gradients and in turn preserves edges in the restoration. On the downside, the optimization of the above is challenging due to the non-differentiability of the total variation cost. A detailed discussion on the recent developments in image restoration through TV regularization can be found in [62].
Maximum Entropy Restoration

The underlying principle in maximum entropy restoration is finding a solution by giving maximum freedom within the limits imposed by the constraints. By considering an image to be a set of positive numbers, the entropy is a measure of uncertainty and can be defined as

$$S = -\sum_{ij} g(i,j) \log g(i,j),$$

where $S$ stands for the entropy of the image [63]. With the standard data fidelity term in Section 2.6.1, the maximum entropy solution for image restoration is defined as

$$\arg\min_g \|b - [K]g\|^2 + \lambda \sum_{ij} g(i,j) \log g(i,j).$$

In addition to the standard data fidelity term, other choices have also been applied such as chi-squared [64]. Maximum entropy restoration has been applied mostly to the restoration of astronomical and radio images and X-ray and gamma-ray data [65].

2.7 Non-blind Stochastic Restoration

These methods model the original image, the blur, and the noise as random fields and try to estimate the GTI as the most probable realization of a certain random process. For example, the original image is modeled as a two-dimensional autoregressive process, and the blur is modeled as a two dimensional linear system with finite impulse response. When compared with the deterministic models, these models rely heavily on the strong statistical hypothesis and thus sensitive to modeling errors.

At this point, it should be emphasized here that although some statistical properties are used in some of the non-blind deterministic restoration methods (discussed in Section 2.6), such as Wiener filter and maximum entropy restoration, we differentiate that with the methods in this section, as they are primarily used
for regularization, not for modeling the estimation process.

### 2.7.1 Maximum Likelihood Restoration

Bayes theorem in probability is the foundation for most of the stochastic restoration algorithm. In applying Bayes theorem to non-blind image restoration, the conditional probability of the GTI $g$, given the blurred image $b$, is given by

$$ p(g|b) = \frac{p(b|g)p(g)}{\sum_{ij} p(b|g)p(g)}, $$

where $i, j$ represent the pixel points of the image.

The seminal papers on image restoration based on the Bayes theorem by Richardson [66] and Lucy [67] created a major breakthrough in image restoration. The Richardson-Lucy restoration algorithm is based on the assumption that the image data is approximately poisson distributed, attributing to the photon noise in the data and has been most widely used for restoring astronomical images. As the original Richardson-Lucy algorithm is prone to errors due to noise and instability, various enhancements have been introduced such as introducing bilateral regularization [68] and using in multichannel restoration [69].

Another variation of Bayes theorem is the probabilistic framework using maximum likelihood (ML) estimation in [70] with the expectation maximization (EM) algorithm. The EM algorithm is an iterative technique for solving ML problems [71], where the optimization is performed in alternate computation of expectation and maximizing the expectation. Tekalp et al. [72] used ML for estimation by modeling the blurring as a 2D autoregressive moving-average identification problem, where the autoregressive coefficients are related to the image model while the moving-average coefficients relate to the blur model.

However, the ML formulation does not allow the incorporation of prior knowledge, which is essential in order to reduce the degrees of freedom of the available observations in the restoration problem. As a result, in order to make these algorithms to work in practice, a number of deterministic constraints such as PSF support and symmetry had to be used [73]. To overcome these limitations, numerous changes have been adopted to the original ML solution such as introducing
sparse regularization [74], which in turn relates to maximum a posteriori solution, which is detailed in the next section.

### 2.7.2 Maximum a Posteriori Restoration

When the prior distributions are applied to a probabilistic framework, the regularized solution corresponds to *maximum a posteriori* (MAP) solution, which maximizes the posterior distribution. Under MAP technique,

\[
\hat{g} = \arg \max_g p(g|b) = \arg \max_g \log p(b|g) + \log p(g),
\]

where, the data dependent term \(\log p(b|g)\) is called the log-likelihood function and the term \(\log p(g)\), which is dependent only on \(g\) is termed as the prior model. The likelihood function captures the dependence of the data on the field, and enforces fidelity to data while the prior model term captures a priori knowledge about the GTI in the absence of data, and incorporates this information into the estimate. These two terms are similar to the Tikhonov-Miller regularization introduced in Section 2.6.1 and a detailed discussion of this similarity including the performance evaluations can be found in Chapter 3.

While direct methods could be applied to prior models that are equivalent to quadratic regularization stabilizing functionals, with the use of sparse prior models [2], which are generally nonlinear, optimization can be expensive, and requires good initializations to avoid converging to local minima of the cost. In dealing with such non-linear processing, although the earlier iterative methods such as Landweber [58] were found to be simple to understand and implement, their slow convergence rate motivated research to use more powerful iterative methods such as conjugate gradient methods [75].

Recently, an exhaustive treatment of MAP technique on blind image restoration was provided in [19, 24], where few important facts on naive MAP estimation are revealed. One of the critical suggestions in [24], was that, as the failure of direct application of MAP technique in image restoration mostly relates to the choice of the estimator, by estimating the blur kernel through MAP and then employing a non-blind restoration algorithm would result in a more accurate GTI estimation.
As the MAP technique is the most classical and frequent approach to image restoration dating from the early days [76] to the most recent research [2–4,24], we follow the MAP framework in our non-blind deconvolution as described in Chapter 3.

2.8 Multichannel Image Restoration

In signal processing, a multichannel signal is referred to a group of signals which relate in cross-channel similarity or correlation. In image restoration literature, multichannel restoration (MCR) has been addressed as multiple frames where, more than one blurred image of the same scene were acquired. These multiple frames are acquired in various ways such as by different sensors or at different times. In addition, the frames can also be generated as multiple channels, in exploiting the three red, green blue (rgb) channels of color images or different frequency bands.

The benefits in MCR over single channel lie in the diversity and redundancy of the information between the channels, which results in relaxing the assumptions on the image and PSF models (which were discussed throughout from Section 2.4.3 to Section 2.7), accommodating the input to be virtually arbitrary. For example, in [77], the restoration of color images was performed with cross channel correlation, while in [78], they used the additional within-channel correlation under separable PSF. Further, MCR results in a more stable estimate than the single channel restoration techniques and could be performed without regularization for moderate noisy images. Use of multiple images involve inherent multiple image generation in some applications such as remote sensing, or acquired with different camera settings, such as position and focus. When compared with the conventional single-input restoration methods, although MCR methods can improve performance considerably, exploiting the related information in different channels, one of the important fact to consider is the assumption of the channels being spatially aligned. In circumstances, where this condition is violated, the channels have to be aligned through registration or other mechanism [79,80].

Similar to single channel restoration, MCR also could be carried out either as two steps, where the blur kernels for the channels are identified first, which
is followed by restoration or as direct restoration, where the GTI is estimated directly. A brief overview of those two techniques is given below.

2.8.1 Channel Identification

Channel identification has characteristics similar to PSF estimation detailed in Section 2.4.3 and Section 2.5.3, where the PSFs related to channels are first identified. The popular techniques for channel identification in one dimensional (1D) include deterministic techniques such as subspace methods [81] and statistical methods such as ML [47]. These techniques were extended to 2D in [82] and it is shown that in a noise-free environment, when there are at least three channels, the blur filters can be recovered accurately. This is due to the fact that the lack of information in frequencies are balanced by multiple channel filters. The robustness of these subspace techniques to noise was addressed in [35] with the introduction of TV regularization.

In addition to these, recently, other techniques also have been developed. In mentioning few of those, in [69], the blur kernel is estimated with the help of the noisy image through regularization and Landweber iterative method [58] and in [83], with the help of motion blurs in different directions.

2.8.2 Direct Restoration

Instead of estimating the blur functions, direct restoration is also proved to be possible with various techniques. Under direct restoration, one of the highly used techniques is the use of mutually reference equalizers (MRE) [13], which is based on using equalizers as discussed in Section 2.5.1. The superiority of MRE over single channel equalizers comes through the small support of the MRE, which is realistic in most cases. The existence of the inversion filters designed through this technique depends on the facts that

1. the blur kernels in the z-domain have no common factor except for a scalar constant and
2. the filter matrix has full column-rank.
Under these assumptions, in the noiseless case, it is proven that exact image de-
convolution can be performed when the number of channels exceeds two. Further,
it should also be emphasized that even in the presence of noise, which imposes
fundamental limitations on the degree of accuracy, a stable estimate could be
determined through MRE.

Direct restoration has also been developed by other means such as, in [84],
with the help of multiple images, the movement of the object is tracked and a high
resolution deblurred image is estimated by iteratively minimizing the difference
between the predicted and observed images.

The essence of multichannel restoration lies on the ability of producing multiple
images of the same scene, which is not achievable in most of the realistic consumer
photography. Exploring this constraint, in Chapter 6, we develop a novel multi-
channel model using only a single image, exploiting the diversity of structures in
the images.

2.9 Summary

In this chapter, we presented a comprehensive literature review on the field of
image restoration. While the fundamental building blocks underlying the mod-
elts developed are mentioned in the first sections, the latter sections cover the
algorithms and models developed in the literature. In categorizing the numerous
algorithms and models, we developed a novel classification, which explore the suc-
cessful mechanisms in the past, incorporating the recently developed techniques
and emphasizing their contributions. As the existing surveys on image restoration
lack the contribution of recently developed techniques, we believe our classification
has provided researchers, both novice and expert, an overall view of the spatial
invariant image restoration techniques developed during the past few decades.
Chapter 3

Regularized Image Restoration

3.1 Overview

As mentioned in Chapter 2, image restoration literature is highly abundant with different restoration techniques. The classification and the algorithms mentioned in the literature represent the foundations for the continuously evolving techniques in this field. Today, we experience a vast contribution to these investigations and attempts to improve upon these fundamental approaches. A great deal of this novel research is motivated by the desire to find a better estimated ground truth image (GTI). Yet, in some of these techniques, we find the ultimate objective of restoration is obscured and some important factors are being overlooked such as reducing computational complexity and convergence speed which are important in real-time image restoration. This stresses the need of taking an objective look at restoration in order to most effectively utilize the existing knowledge in this area.

Although image restoration literature has identified blind deconvolution to be more challenging than non-blind deconvolution, in recent years, a number of effective ways have emerged to directly estimate the Point Spread Function (PSF) as detailed in Chapter 2. With these effective and efficient PSF estimation techniques, the research trend has moved in the direction of handling blind deconvolution as two steps, with PSF estimation as the first step and image estimation as the second step [24]. This motivates us to find simple and efficient algorithms for non-blind image restoration. Thus, in this chapter, we contribute in,
1. Analyzing existing quadratic regularization functionals in image restoration and developing a class of novel quadratic regularization functionals with better performance.

2. Introducing a novel visual metric for comparison of regularization models in image restoration, which outperforms existing metrics in signal fidelity measurement comparisons.

3. Critiquing Sparse and Laplacian prior model performance with quadratic regularization performance in a new perspective and showing that Sparse and Laplacian prior models do not result in a superior restoration over quadratic regularization.

4. Analyzing and critiquing the existing likelihood models for image restoration under the maximum a posteriori (MAP) framework and finding the most effective and efficient likelihood model for non-blind image restoration. Specifically, we find the high complexity “derivative likelihood” models offer no significant advantage to a properly configured “normal likelihood” model.

3.2 Quadratic Regularization in Image Restoration

3.2.1 Regularization

As explained earlier in Chapter 2, image acquisition is best modeled by a continuous model in infinite dimensional space, which is categorized as a Fredholm integral equation of the first kind [85]. In the sense of Hadamard [86], a solution to a well-posed problem satisfies the conditions of existence, uniqueness and stability. As Fredholm integral equations of the first kind do not meet the criteria for a well-posed problem, image restoration belongs into the general class of problems which are classified as ill-posed problems [33]. The ill-posed nature of image restoration problem implies that, small bounded perturbations in the data may lead to unbounded deviations in the solution [87].
When working with linear algebra to find a solution for ill-posed problems such as image restoration, one of the simplest methods is the use of the pseudo inverse [88], for which the solution fulfills the existence and uniqueness conditions of Hadamard’s well-posed problem [86], but fails in meeting the stability condition. Regularization is one of the most widely accepted and used technique, in which the solution fulfills all three conditions of a well-posed problem. The concept underlying regularization is to find an acceptable solution from imperfect data, for which, the problem should be stated more completely including some extra or priori information [33,56].

Regularization approaches to image restoration are classified as two-fold, where the stochastic regularization approach uses the knowledge of covariance matrices of the GTI and noise, while the deterministic regularization applies the constraint that most images are relatively featureless with limited high-frequency activity [1]. While stochastic regularization has been used extensively in the past, with important contributions to the field such as Wiener filter [55], recently, much emphasis has been on the use of derivative filters with deterministic regularization [2,3]. Thus, our contribution in this chapter relates to deterministic regularization and the term regularization, henceforth, refers to the deterministic regularization techniques.

Among many regularization techniques, Tikhonov [33] regularization is one of the oldest and well-known techniques for stabilization. It was proposed in [33], that the solution for

\[ b = Kg + n, \] (3.1)

where \( b \) is the measured data, \( g \) is the original data (ground truth), \( K \) is the distortion operator or the transformation and \( n \) represents additive random noise, can be achieved by constrained minimization of a functional \( \Phi(g) \), which is called the **stabilizing functional**. Under the stabilizing functional approach, the image restoration problem is formulated as determining \( g = \hat{g} \), which minimizes the functional \( \Phi(g) \) under the condition that the element \( g = \hat{g} \) satisfies

\[ \|b - Kg\|^2 = \delta, \] (3.2)
where $\delta$ is a positive constant and $\| \cdot \|$ stands for the Frobenius norm. The Frobenius norm of an arbitrary $m_1 \times m_2$ real matrix $A$ is defined as

$$
\|A\| = \sqrt{\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} a_{ij}^2},
$$

(3.3)

where $a_{ij}$ is the $(i,j)$ entry in matrix $A$. The constrained minimization problem in (3.2) can be solved by the method of Lagrange multipliers, which is to determine $\hat{g}$, an estimate of ground truth $g$, by minimizing the functional

$$
\|b - Kg\|^2 + \lambda \Phi(g),
$$

(3.4)

where $\lambda > 0$ is the Lagrange multiplier and is often called as the \textit{regularization parameter}.

The first term in (3.4), named as \textit{data-fidelity term} fits to the data, while \textit{stabilizing functional} incorporates “believed” properties of the GTI. Generally the \textit{data-fidelity term} is a standard fixed choice. In contrast, the richness and variety of image restoration techniques comes down to different choices of the regularization term, reflecting different implicit models. As the choice of the \textit{stabilizing functional} can take a variety of forms, in this chapter, we selected two widely used model classes for our analysis: the fast quadratic stabilizing functionals are introduced in Section 3.2.2, while relating the \textit{stabilizing functional} to the \textit{priori} knowledge in a probabilistic viewpoint, we discuss the Sparse and Laplacian prior models in Section 3.3.2, which are claimed to have better performance [2].

When the stabilizing functional $\Phi(g)$ in (3.4) belongs to the class of nonnegative quadratic functionals, the minimization problem can be expressed as

$$
\min \|b - Kg\|^2 + \lambda \|Dg\|^2,
$$

(3.5)

where $D$ is a bounded linear operator [56] and is often called as the \textit{regularization operator} or \textit{stabilizing operator}. It is shown in [57], that the minimization problem in (3.5) can be formulated as a constrained least squares image restoration problem.
when the solution $g$ satisfies the necessary and sufficient condition of

$$(\mathcal{K}^T \mathcal{K} + \lambda \mathcal{D}^T \mathcal{D}) g = \mathcal{K}^T b. \quad (3.6)$$

This leads to the closed form solution for (3.5) in the form

$$\hat{g} = (\mathcal{K}^T \mathcal{K} + \lambda (\mathcal{D}^T \mathcal{D}))^{-1} \mathcal{K}^T b. \quad (3.7)$$

We extend, in a trivial way, the above formulation by considering $\mathcal{D}$ as the combination of $R$ component regularization operators, in the form of

$$\mathcal{D} g \triangleq \begin{pmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \\ \vdots \\ \mathcal{D}_R \end{pmatrix} g. \quad (3.8)$$

With the introduction of $R$ Lagrange multipliers, the general form of (3.5) can be expressed as

$$\min \| b - \mathcal{K} g \|^2 + \sum_{r=1}^{R} \lambda_r \| \mathcal{D}_r g \|^2, \quad (3.9)$$

for which, the closed form solution is given by

$$\hat{g} = (\mathcal{K} \mathcal{K}^T + \sum_{r=1}^{R} \lambda_r (\mathcal{D}_r^T \mathcal{D}_r))^{-1} \mathcal{K}^T b. \quad (3.10)$$

As the images are of limited support and when the corresponding hypothesis of uniformity on image edges can be made, the matrices $\mathcal{K}$ and $\mathcal{D}$ in (3.10) have a special structure and are called block circulant matrices [89]. As circulant matrices can be diagonalized by the discrete Fourier transform, the minimization in (3.9) can be solved extremely quickly using the Fourier domain techniques [57].
3.2.2 Regularization Operators as Components in Quadratic Stabilizing Functionals

The generality of the regularization operator allows the development of a class of linear operators and the minimization in (3.9) will be the source of many regularizing solutions for (3.1) depending on the choice of the regularization operator. This choice is usually based on the known details of the image formation process and plays an important role in the regularization.

The simplest regularization operator is when $\mathcal{D}$ is an identity matrix, where $\mathcal{D}g = g$, and the regularized solution for this was referred as minimum norm restoration [26]. In general, $\mathcal{D}$ often takes the form of a sparsifying operator such as a discrete approximation of a derivative operator. Through the experiments in [90], it was shown that even though the statistics of natural images vary from image to image, the histograms for the response of derivative filters are relatively consistent and scale invariant across the images. Taking these factors into consideration, in this section, we discuss a class of regularization operators based on the partial derivative operators (PDO), which could be used in the quadratic stabilizing functional.

First order partial derivative operator

When first order partial derivative operators (FOPDO) are considered as the regularization operators, $\mathcal{D}g$ in (3.8) can be expressed as

$$\mathcal{D}g = \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} g,$$

where $\partial_x$ and $\partial_y$ are any discrete space, spatially invariant linear operators that emulate first order derivative in $x$ and $y$ directions, respectively [2]. This type of regularization uses two component regularization operators.

Second order partial derivative operator

Second order partial derivative operators (SOPDO), could be derived mainly in two forms.
1. Isotropic SOPDO - When the regularization operator takes the form of

\[ Dg = \begin{pmatrix} \partial_{xx} \\ \partial_{yy} \end{pmatrix} g, \]

it is called the isotropic SOPDO. Though the SOPDO defined above cannot be considered as a true isotropic differential operator, such as the continuous Laplacian operator, it gives the simplest possible isotropic operator with even-order derivatives [91]. Similar to FOPDO, \( \partial_{xx} \) and \( \partial_{yy} \) represent any discrete space, spatially invariant linear operators that emulate second order derivatives.

2. Non-isotropic SOPDO - The non-isotropic SOPDO is formed by

\[ Dg = \begin{pmatrix} \partial_{xx} \\ \partial_{xy} \\ \partial_{yy} \end{pmatrix} g. \]

As the edges and lines in images may occur in any direction, when the differential operator is isotropic it would give better results than the non-isotropic differential operator [91].

**Mixed Partial derivative operator**

In general, considering only even-order derivatives, the use of directional derivatives in more than one dimensional can be expressed as

\[ Dg = \begin{pmatrix} \partial_{s_1}^p \\ \partial_{s_2}^p \\ \vdots \\ \partial_{s_m}^p \end{pmatrix} g, \]

where \( p \) is the order of the derivatives, \( m \) is the number of dimensions and \( s_1 \) to \( s_m \) represent the direction of the derivatives.

Using the above general model, we introduce a new regularization operator,
with different combinations of higher order derivative operators. In this discussion, we limit the use of higher order derivative operators up to the second order, and the new PDO is called first and second order derivative operator (FSOPDO). With FSOPDO, $Dg$ in (3.8) will stand for

$$Dg = \begin{pmatrix}
\frac{\partial}{\partial x} \\
\frac{\partial}{\partial y} \\
\frac{\partial}{\partial xx} \\
\frac{\partial}{\partial yy}
\end{pmatrix} g.$$

These quadratic regularization functionals are compared in a new perspective with the widely used prior models which are believed to result in better performance in Section 3.3.5.

### 3.2.3 Selection of Regularization Parameter

As discussed in Section 3.2.1, under the stabilizing functional approach of regularization, the balance in data fidelity and the stabilizing functional is achieved through the choice of the regularization parameter $\lambda$.

For better restoration results, choosing the correct stabilizing functional by itself is not sufficient. In the regularization process it is the regularization parameter, $\lambda$, that controls the tradeoff between the solution accuracy $\|b - Kg\|^2$ and its degree of regularity $\|Dg\|^2$. For example, perfect fidelity to the data is achieved if $\lambda = 0$, whereas perfect fidelity to the priors is achieved if $\lambda = \infty$. In between, if $\lambda$ is chosen to be too small, the resulting image will exhibit amplified noise, while a large $\lambda$ will result in an image smoothed more than necessary. The more noise an image contains, the higher the regularization parameter needs to be to compensate.

Popular methods in determining the regularization parameter $\lambda$ are the L-curve criterion \[92, 93\] and the Generalized Cross Validation criterion \[93–95\]. With the L-curve method, for varying values of $\lambda$, the log residual image error, $\|b - Kg\|^2$ is plotted against the log regularity $\|Dg\|^2$, which results in a “L” shaped plot. While the more vertical portion of this “L” shape corresponds to under-regularized estimates, the more horizontal portion of the L-curve represents over-
smoothed estimates. Hence, the corner point of the plotted curve can be taken as the optimum regularization parameter, $\lambda$.

While using a regularization parameter is the established method in controlling the balance in fidelity, recent research have used alternative ways, such as regularization functional instead of regularization parameter [96].

### 3.2.4 Non-blind Image Restoration through SOPDO

#### Noisy Image Deconvolution

Although most previous image restoration algorithms have considered FOPDO as the regularization model [2], we claim that SOPDO has better performance in terms of the difference between the ground truth and the estimated data on images which are susceptible to noise. Here, we deal only with additive Gaussian noise, as it effectively models the noise in many different imaging scenarios. In this section, we study in detail a few simulation results which are used to do comparison evaluations with other existing image restoration techniques.

We take non-isotropic SOPDO as the regularization operator for image restoration through least squares restoration as given in Section 3.2.2. In order to compare the performance of the non-isotropic SOPDO prior model, we take two regularization models, FOPDO and a sparse stabilizing functional defined in [2]. Relating regularization to probability, the stabilizing functional in image restoration is also referred as prior model and a detailed discussion of bayesian interpretation to regularization including the Sparse prior model is covered in Section 3.3.2. The deconvolution with non-isotropic SOPDO and FOPDO regularization lead to closed form solutions with highly efficient computation, while the Sparse prior cannot be minimized in closed form [2]. In all the simulations discussed in this section we use

$$\lambda_r = \lambda, \quad \forall \ r \in 1 \ldots R,$$

(3.11)

for the quadratic regularization functionals, where the value of $R$ depends on the respective model such as $R = 2$ for FOPDO and $R = 3$ for non-isotropic SOPDO model.
We claim, using non-isotropic SOPDO prior gives better results for images which are susceptible to noise over the FOPDO. When comparing the non-isotropic SOPDO regularization with the Sparse prior, we found that the non-isotropic SOPDO regularization outperforms Sparse prior significantly in speed. These results are shown in Table 3.1.

For the experiment in Table 3.1, we added “Gaussian” noise to the original “Picasso” image [4] with a standard deviation of 0.0001. The original colored image was first converted to grayscale with the pixel values resulting in the range from 0 to 1 and the original image was considered to be periodic. The term MSE stands for mean square error and for a two-dimensional image, MSE is defined as

\[
MSE \triangleq \frac{1}{L_1 L_2} \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} (g(l_1, l_2) - \hat{g}(l_1, l_2))^2
\]

where \(g\) and \(\hat{g}\) represent GTI and the estimated GTI, respectively while \(L_1\) and \(L_2\) represent the size of the image in \(x\) and \(y\) directions respectively. The MSE values in Table 3.1 are in multiples of \(10^{-4}\) while the time is given in seconds. The results show that the non-isotropic SOPDO outperforms FOPDO on MSE and has a significant advantage over Sparse prior on speed performance.

<table>
<thead>
<tr>
<th>Smoothing weight</th>
<th>Non-isotropic SOPDO</th>
<th>FOPDO</th>
<th>Sparse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE (\times 10^{-4})</td>
<td>Time in seconds</td>
<td>MSE (\times 10^{-4})</td>
</tr>
<tr>
<td>0.1</td>
<td>4.978</td>
<td>0.43</td>
<td>5.168</td>
</tr>
<tr>
<td>0.05</td>
<td>4.646</td>
<td>0.42</td>
<td>4.730</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>6.043</td>
<td>0.42</td>
<td>7.561</td>
</tr>
</tbody>
</table>

Table 3.1: Comparison of stabilizing functional model.
### 3.2 Quadratic Regularization in Image Restoration

<table>
<thead>
<tr>
<th>Restoration Technique</th>
<th>Efficiency in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levin Sparse deconvolution (50 iterations)</td>
<td>556</td>
</tr>
<tr>
<td>Levin Sparse deconvolution (10 iterations)</td>
<td>124</td>
</tr>
<tr>
<td>Shan executable</td>
<td>39</td>
</tr>
<tr>
<td>Non-isotropic SOPDO deconvolution</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.2: Efficiency in non-blind image deconvolution

**Efficiency in Deconvolution**

As SOPDO can use frequency domain deconvolution techniques, it can be implemented highly efficiently than most of the recent non-blind deblurring techniques. The comparison was done with the Sparse deconvolution algorithm in [2] named as “Levin Sparse deconvolution” and the non-blind deconvolution of [4] (distributed online) named as “Shan executable”. The results in Table 3.2 support the claim that non-isotropic SOPDO regularization model results in the best speed performance when compared with “Levin Sparse deconvolution” and “Shan executable” methods.

Further, we tested for the robustness of non-isotropic SOPDO regularization by using different sized images with varying sized kernels. The detailed results are shown in Table 3.3. All the images used for this experiment are color images, having separate rgb (red, green, blue) channels. The image and kernel sizes are given in pixels and the efficiency was measured in seconds. The results clearly show the robustness and the efficiency of the non-isotropic SOPDO regularization model with respect to different scales of image and kernel.

**Performance in Deconvolution**

Several computational experiments were carried out in order to compare non-isotropic SOPDO regularization with “Levin Sparse deconvolution” and “Shan executable”. The performance of these deconvolution techniques on a naturally blurred, highly textured image, given in [4], are shown in Fig. 3.1 and Fig. 3.2. The blur kernel used in this experiment was retrieved through the blind deconvolution package of [4] distributed online. Closer visual inspection of the image results show that non-isotropic SOPDO technique best shows the tree branches and leaves while...
<table>
<thead>
<tr>
<th>Restoration Technique</th>
<th>Image size in pixels</th>
<th>Kernel size in pixels</th>
<th>Efficiency in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levin Sparse deconv.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(50 iterations)</td>
<td>484 × 752</td>
<td>19 × 27</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>99 × 99</td>
<td>556</td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>99 × 99</td>
<td>1240</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2429</td>
</tr>
<tr>
<td>Levin Sparse deconv.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10 iterations)</td>
<td>484 × 752</td>
<td>19 × 27</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>99 × 99</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>99 × 99</td>
<td>408</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>848</td>
</tr>
<tr>
<td>Shan executable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>484 × 752</td>
<td>19 × 27</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>99 × 99</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>99 × 99</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Error</td>
</tr>
<tr>
<td>Non-isotropic SOPDO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regularization</td>
<td>484 × 752</td>
<td>19 × 27</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>99 × 99</td>
<td>2.44</td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>99 × 99</td>
<td>5.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13.17</td>
</tr>
</tbody>
</table>

Table 3.3: Efficiency results on scaling
the other techniques have a blurring effect still remaining on the estimated result. This fact is further discussed and evidenced by evaluating the deconvolution in a new perspective in Section 3.3.5.

### 3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

#### 3.3.1 Problem Formulation

While the development of regularized solutions for ill-posed problems is widely discussed in the signal processing literature, recently by looking at the ill-posed image restoration problem from a probabilistic viewpoint, some researchers claim that the Sparse prior model [2,3] (discussed in Section 3.3.2) outperforms quadratic regularization models (discussed in Section 3.2.2). The analytical study in this section addresses the following problems.

1. Are sparse prior models superior to quadratic regularization models?

2. What is the source of better performance of sparse prior models?
Figure 3.2: Image results for a highly textured image
3. Are fast quadratic regularization models good enough for image restoration?

### 3.3.2 Regularization - Bayesian Interpretation

Inverse problems such as image restoration are seen as probabilistic inference problems, where lack of information is compensated by assumptions. Therefore, it is not surprising, when the nature of the regularization detailed above is taken into consideration, to see that there is a close relationship between regularization and Bayesian estimation. Applying Bayes theorem to the image restoration problem in (3.1), for a known blur kernel, the posterior distribution can be written as

\[
p(g|b) \propto p(b|g)p(g), \tag{3.13}
\]

where \(p(b|g)\) represents the likelihood and \(p(g)\) represents the prior for the ground truth image. The estimation of the GTI based on posterior distribution can be classified in several ways. The minimum mean-square error estimate represents the mean of the posterior density, the MAP estimate stands for the mode of the posterior density while the maximum likelihood (ML) estimate may be viewed as a special case of MAP where no prior distribution is used [97].

Under the MAP technique, estimation of the GTI simplifies to

\[
\hat{g} = \arg \max p(g|b). \tag{3.14}
\]

Considering the non-blind image deconvolution process, we convert (3.14) to an energy minimization problem, where the energy is defined as

\[
E(g) \triangleq -\log(p(g|b)). \tag{3.15}
\]

Different likelihood and prior models on the ground truth have been applied for image restoration in literature. An analysis of existing prior models can be found in [98].

Considering the fact that for a given \(g\), the variation in \(b\) is due to the noise \(n\) [97], together with the above definitions, non-blind image restoration problem
can be recast as seeking the unknown GTI, \( g(i, j) \), that minimizes the functional

\[
\|b - K g\|^2 + \sum_{r=1}^{R} \sum_{ij} \lambda_r \rho(D_r g(i, j)),
\]  
(3.16)

where \( D_r \) is the \( r \)-th of \( R \) linear transformations, \( i, j \) are pixel indices, \( \lambda_r > 0 \) are regularization parameters, \( \| \cdot \| \) stands for the Frobenius norm and \( \rho(\cdot) \) is a scalar memoryless nonlinear mapping, generally taking the form

\[
\rho(z) \triangleq |z|^{\alpha}
\]  
(3.17)

for judicious choice of real parameter \( \alpha \) (not necessarily integer).

Many techniques belong to this class and differ only in:

1. the set of linear operators \( D_r, r = 1, 2, \ldots, R \), and
2. the nonlinear mapping \( \rho(z) \) (or choice of \( \alpha \)).

Numerous image restoration techniques have been developed under this framework continuously from the early days [99, 100] to the most recent research [2–4, 24].

**Sparse prior**

In recent literature, it is shown that, when derivative filters are applied to natural images, the filter outputs tend to be sparse [101, 102]. In other words, the histogram of the derivative filtered image peaks at zero and falls off much faster than a Gaussian distribution. These heavy tailed natural image priors are used in a number of applications in image processing literature, such as denoising [103, 104], reflection separation [105, 106] and deconvolution [4, 12] in which, they are implemented in various ways such as student-\( t \) distributions [104] and scale mixtures of Gaussian distributions [3, 107].

In [2], sparsity is incorporated by having \( D_r \) as the derivative filters and \( \alpha = 0.8 \) in (3.17) as the prior term, which results in

\[
\|b - K g\|^2 + \sum_{r=1}^{R} \sum_{ij} \lambda_r (D_r g(i, j))^{0.8}.
\]  
(3.18)
3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

This can be solved in the spatial domain using Conjugate Gradient algorithms [75].

**Laplacian prior**

Although not as close as the Sparse prior to the natural image priors, Laplacian prior with $\alpha = 1$ in (3.17) is expected to result in a less smooth solution than the Gaussian prior. With the Laplacian prior, the optimization becomes

$$\|b - Kg\|^2 + \sum_{r=1}^{R} \lambda_r \|D_r g\|_1.$$  \hspace{1cm} (3.19)

Recently, much attention has been paid in solving $L^1$ norm regularization problems through compressed sensing. In [108], an efficient method for optimizing a solution to a problem similar to (3.19) was discussed when $D_r$ are invertible.

**Gaussian prior**

When $\alpha = 2$, minimization in (3.16) is called the Gaussian prior deconvolution in [2] and is equivalent to the quadratic regularization problem in (3.9). Thus, in this chapter, we use the terms Gaussian prior and quadratic (specifically isotropic SOPDO) regularization interchangeably.

### 3.3.3 Image Restoration Evaluation

**Metric**

For all the restoration performance analysis and comparisons in this paper, we use a recently developed visual metric called SSIM (Structured SIMilarity) index [21], which has not been used for the comparison of prior models in the image restoration literature to date. The approach of SSIM is motivated by the highly structured characteristics of the natural image, where the strong neighborhood dependencies carry important information about the structures of the objects in the visual scene [20].

Assuming $x$ and $y$ are local image patches representing the same patch in the original and estimated images, the local SSIM index measures the similarities of three elements of the image patches: the similarity $l(x, y)$ of the local patch
luminances (brightness values), the similarity \( c(x, y) \) of the local patch contrasts, and the similarity \( s(x, y) \) of the local patch structures. These local similarities are expressed using simple, easily computed statistics, and combined together to form local SSIM, \( S(x, y) \) [20].

\[
S(x, y) = l(x, y).c(x, y).s(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \left( \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \cdot \left( \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right),
\]

where \( \mu_x \) and \( \mu_y \) are the local sample means of \( x \) and \( y \) respectively, \( \sigma_x \) and \( \sigma_y \) are the local sample standard deviations of \( x \) and \( y \) respectively, and \( \sigma_{xy} \) is the sample cross correlation of \( x \) and \( y \) after removing their means. The items \( C_1, C_2, \) and \( C_3 \) are small positive constants that stabilize each term, so that near zero sample means, variances, or correlations do not lead to numerical instability.

Due to the fact that the underlying principle of SSIM is to extract the structural information which complies with the human visual system, SSIM maps are asserted to be a better signal fidelity measurement over MSE [20]. In evaluating images through MSE, all image pixels are treated equally and content dependent variations in image fidelity are not accounted for. The two main indicators in SSIM evaluations, mean SSIM (MSSIM) and SSIM maps have values in the range from 0 to 1, where 1 indicates the best restoration. Although MSSIM and SSIM maps are generally used as visual fidelity metrics, we evaluate image restoration with the histogram of the SSIM map, as it provides an accurate view of the local restoration.

**Image Restoration Models**

Ignoring the presence of noise in image acquisition represented by (3.1), general image restoration could be represented by the model shown in Fig. 3.3, where \( \mathcal{L} \) represents the deblur process. The notation in (3.1) and the representation in Fig. 3.3 may be deceptive and sometimes mislead research. As already discussed in Chapter 2, from physical intuition, we could see that even though \( g \) is continuous by nature, image recording imposes a limited degree of freedom on \( g \) and \( b \), leading to artifacts on the final estimate of image restoration.
3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

As illustrated in Fig. 3.4, we categorize these spatial artifacts in two ways. The “Model artifacts” are those, which are not present on naturally blurred images, but introduced in blur simulations as a result of sharp intensity transitions at the boundary of a finite image. Generating a blur image from a finite GTI causes unnatural blur distortions in the vicinity of the boundary of the image. Suppression of these “Model artifacts” could be accomplished by preprocessing the observed degraded image with techniques such as truncation and reducing the size of the blurred image. On the other hand, “Process artifacts” come along with the deblur process $\mathcal{L}$ due to finite $b$, which affect the performance of most deconvolution algorithms.

In order to show the effect of ‘Process artifacts’, we restore an image, originally, of size $255 \times 255$ pixels, but truncated in order to remove the “Model artifacts” introduced by a $13 \times 13$ pixels blur kernel, making the final image of size $242 \times 242$. The results of restoration with Sparse, Laplacian and Gaussian priors are shown in Fig. 3.5. In this experiment, deconvolution with Sparse and Laplacian priors were carried out using iterative re-weighted least squares (IRLS) method [109], through the code available online [2], while the Gaussian prior is processed with both IRLS and fast Fourier techniques (FFT) separately. In our simulations, we processed IRLS for 150 iterations beyond which there were no further improvements. Analyzing the results of the performance of the Gaussian prior model with
the FFT and IRLS techniques, we see that process artifacts are better handled by the IRLS technique than the FFT and this result is justified by the IRLS processing of Sparse and Laplacian prior models.

Both the “Model” and “Process” artifacts discussed above are not part of natural images, but are imposed artificially by the image modeling and processing techniques. Thus we claim that the evaluation of image restoration should be carried out excluding these artifacts.

### 3.3.4 Performance of Quadratic Regularization Operators

In order to achieve the objective of studying the performance of different operators in the quadratic regularization as detailed in Section 3.2.2, we carried out some simulations, where we avoided the effect of ‘Model artifacts’ by taking a boundary strip off from the blurred image. In our evaluations, we used FOPDO, SOPDO and FSOPDO models to compare the performance. From this point onwards the

---

**Figure 3.5**: Image restoration results with prior models.
3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

Table 3.4: Choice of regularization parameter ($\lambda$) values for different quadratic regularization operators used in the simulations of Table 3.5

<table>
<thead>
<tr>
<th>Regularization operator</th>
<th>$r$</th>
<th>$\lambda_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOPDO</td>
<td>1, 2</td>
<td>$\beta_1$</td>
</tr>
<tr>
<td>SOPDO</td>
<td>3, 4</td>
<td>$\beta_2$</td>
</tr>
<tr>
<td>FSOPDO</td>
<td>1, 2</td>
<td>$\beta_1$</td>
</tr>
<tr>
<td></td>
<td>3, 4</td>
<td>$\beta_2$</td>
</tr>
</tbody>
</table>

Table 3.5: Performance of quadratic regularization operators under varying regularization parameter values

<table>
<thead>
<tr>
<th>Regularization parameter</th>
<th>MSSIM values for</th>
<th>FOPDO</th>
<th>SOPDO</th>
<th>FSOPDO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1 = \beta_2 = 0.001$</td>
<td></td>
<td>0.9412</td>
<td>0.9597</td>
<td>0.9626</td>
</tr>
<tr>
<td>$\beta_1 = 0.001, \beta_2 = 0.003$</td>
<td></td>
<td>0.9596</td>
<td>0.9597</td>
<td>0.9657</td>
</tr>
<tr>
<td>$\beta_1 = 0.003, \beta_2 = 0.001$</td>
<td></td>
<td>0.9412</td>
<td>0.9674</td>
<td>0.968</td>
</tr>
</tbody>
</table>

Term SOPDO refers to isotropic-SOPDO unless stated otherwise.

The simulations, for which the results are demonstrated in Fig. 3.6, are executed in the same environment as the simulation for Fig. 3.5, but with quadratic regularization models. We evaluated the performance of the regularization models under varying regularization parameter ($\lambda$) values as discussed in Section 3.2.2. While the choice of parameters representing $\lambda$ for FOPDO, SOPDO and FSOPDO are given in Table 3.4, the actual values for the respective parameters are given in Table 3.5. While the overall SSIM values for few of the simulation results under varying $\lambda$ values are shown in Table 3.5, the histogram distribution representing the first line of Table 3.5 is shown in Fig. 3.7. Overall, by analyzing these results, we claim that, in the presence of ‘Process artifacts’, a better performance could be achieved with FSOPDO over FOPDO and SOPDO models. In the next section we compare the performance of these quadratic regularization models by removing the “Process” and “Model” artifacts.
Figure 3.6: Image restoration results with quadratic regularization models.

Figure 3.7: Image restoration results for simulations in Table 3.5
### 3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

Figure 3.8: Image restoration model for a naturally blurred image, where $\mathcal{K}$ is the blur process, $\mathcal{L}$ is the deblur process, $g$, $b$, $z$, $\hat{g}$ stand for ground truth image, blur image, deblurred image with artifacts, and the final estimated ground truth image respectively. The process $P_M(z)$ decouples “Model” and “Process” artifacts from the deblurred image.

#### 3.3.5 Regularization model comparison

As shown earlier in Section 3.2.2, we modeled the regularization of image restoration based on the quadratic regularization terms (sometimes called as the least squares regularization) and in Section 3.3.2, we discussed the existing probabilistic models under a MAP framework. These models form a method of regularization in image restoration. This section is devoted for the comparison of these models. The comparison in this section will guide us for making recommendations for the appropriate regularization technique and is discussed at the end of this section.

**SSIM Performance Comparison**

As the objective of our simulations is to evaluate the contribution of the regularization models towards image restoration, we use the restoration model shown in Fig. 3.8, where we decouple artifact effects from restoration by projecting the estimated image with

$$P_M(z)(i, j) = \begin{cases} 
  z(i, j), & \text{if } i, j \in M \\
  0, & \text{otherwise}
\end{cases}$$

where $M$ is a region without ‘Model’ and ‘Process’ artifacts. To be consistent with the SSIM map region in Fig. 3.5, we take a large image of support $1024 \times 1024$ and project the final image to a $242 \times 242$ region within the inner region of the estimated image, which is least affected by the artifacts. The restoration was carried out with FFT processing of the Gaussian prior and IRLS processing of Sparse and Laplacian priors. The comparison of the performance of the priors
is shown in Fig. 3.9. In it we note that Gaussian prior with FFT processing has performed as well as or better to the Sparse and Laplacian prior models.

As the literature claims that iterative algorithms such as conjugate gradient algorithms suppress noise and perform better in noisy blur image restoration, we simulated a noisy blurred restoration under the same conditions given for Fig. 3.9, but with different regularization parameter values, as more weight should now be given to the prior over data. The noise added was Gaussian with zero mean and 0.01 variance. The optimal results we obtained for varying \( \lambda \) are shown in Fig. 3.10. With these results, we claim that Gaussian prior handles noisy images as better as the Sparse and Laplacian prior models.

Thus, these results pave a new path of thinking and we claim that quadratic regularization with SOPDO model, when appropriately configured and used in a realistic context, free from unnatural artifacts, is comparable to Sparse prior model in terms of image restoration performance under the SSIM criterion.
3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

Figure 3.10: Image restoration results for a noisy image under the same environment in Fig. 3.9.
Table 3.6: Efficiency of regularization operators. The times taken for restoration of grey-scale and colored images are given in seconds for each of the regularization operators.

<table>
<thead>
<tr>
<th>Image size in pixels</th>
<th>Time taken for Quadratic regularization with FFT processing given in seconds</th>
<th>Time taken for Sparse prior processing given in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50 iterations</td>
<td>100 iterations</td>
</tr>
<tr>
<td>grey-scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>255 x 255</td>
<td>0.08</td>
<td>30</td>
</tr>
<tr>
<td>1024 x 1024</td>
<td>1.07</td>
<td>303</td>
</tr>
<tr>
<td>colored</td>
<td></td>
<td></td>
</tr>
<tr>
<td>484 x 752</td>
<td>1.7</td>
<td>371</td>
</tr>
<tr>
<td>1107 x 1694</td>
<td>9.0</td>
<td>2180</td>
</tr>
</tbody>
</table>

Efficiency Comparison

As the optimization problem in least-squares regularization is convex and as the fast Fourier techniques could be applied for the computation, for an image of size \( L \times L \) pixels, the restoration through least-squares regulation has a complexity of \( O(L \log L) \) operations. On the other hand, when a Sparse prior is used, the optimization problem is no longer convex and cannot be minimized in closed form. Using the conjugate gradient method [75], or IRLS method, the optimization can be solved in \( O(Li_{\text{max}}) \) where \( i_{\text{max}} \) represent the maximum number of iterations.

A few simulation results on efficiency are shown in Table 3.6, where all the values are in seconds and represent the time taken for the restoration using each of the respective model. While the quadratic regularization deconvolution was carried out using Fourier domain techniques, the Sparse deconvolution was carried out using the IRLS method. Under the IRLS algorithm, it is experienced that in order to achieve an acceptable result, the number of iterations should be at least 50 and better results could be achieved when the number of iterations are above 100. From the results shown, it is evident that when the size of the image increases, the relative efficiency of the restoration through Sparse prior model becomes extremely low.
3.3 Are Sparse Prior Models Superior to Quadratic Regularization Models?

Regularization recommendations

In addition to lower efficiency and not-superior performance, Sparse prior models lack in proper theoretical guidelines for selecting the best regularization parameter. In contrast, the quadratic regularization models have well-established methods as detailed in Section 3.2.3. Difficulties in selecting the optimal converging point in non-convex minimization techniques such as IRLS also is an issue.

According to the theoretical and experimental details provided above, we propose that if we could decouple image restoration and Process artifact handling, then the use of quadratic regularization models will result in more efficient and effective image restoration in comparison to Sparse and Laplacian prior models. The decoupling of image restoration and Process artifact handling could be achieved through techniques such as tiling [110], which enables the uses of the efficient least squares regularization.

Thus, coming back to our problem formulation in Section 3.3.1, we claim that

1. Sparse prior models are not superior to quadratic regularization models in terms of performance in image restoration.

2. In terms of efficiency, Sparse prior models are significantly inferior to quadratic regularization models.

3. The performance through Sparse prior model increases over quadratic regularization models when boundary effects are not addressed and processing artifacts are not compensated for.

4. Quadratic regularization models provide the best image restoration for large images in terms of efficiency and effectiveness while they provide a good enough solution for other images when the boundary artifacts are taken care of.

Analyzing the above items further, if the improvements of the Sparse prior model are in artifact handling, not in image restoration, we can pose the following questions.
“Do more complicated prior models such as Sparse, which are asserted be better matched to natural images, actually help image restoration in terms of restoring natural image features?”

“If those complicated prior models hold no significant advantage, is it worth the effort spend on them compared to simple and efficient prior models which restore closer or better than those prior models?”

3.4 Likelihood Model Analysis

Different likelihood models in the prior model in (3.13) have been studied in various ways. The fact that most of these models are not justified with proper theoretical foundations encouraged us to analyze and understand the variations and the validity and accuracy of the (implicit) underlying assumptions, which could explain the different performances.

This investigation guides our development of a new scheme for the multiple image likelihood model described in Section 3.4.1. The likelihood model analysis is carried out using this new model and the theoretical analysis is corroborated by the computational experiments detailed in Section 3.4.3.

3.4.1 Likelihood Models in Image Restoration

Likelihood Model for a Single Image

In image restoration literature, the likelihood for a single image is defined by modeling the image noise \( n \) as a set of independently and identically distributed (i.i.d.) random variables following a Gaussian distribution for all pixels, which is given by

\[
p(b|g) = \prod_{l_1=1}^{L_1} \prod_{l_2=1}^{L_2} N(n(l_1, l_2) | 0, \sigma),
\]

where \( N(\cdot|\mu, \sigma) \) denotes a Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \), while \( L_1 \) and \( L_2 \) represent the image support.
Likelihood Model for Multiple Images

Based on the above likelihood model for a single image, we develop a new model for the likelihood of multiple images as detailed below.

Given a set of \( R \) degraded images of a common GTI \( g \), the posterior distribution for the GTI can be derived by extending (3.13), resulting in

\[
p(g|b_1, b_2, \ldots, b_R) \propto p(b_1, b_2, \ldots, b_R|g)p(g),
\]

where, generalizing (3.1),

\[
b_r = \mathcal{K}_r g + n_r, \quad r = 1, 2, 3, \ldots, R
\]

and \( \mathcal{K}_r \) are operators representing possibly different but known blurs, and \( n_r \) are noise images. Under the assumption that \( n_p \) is independent of \( n_q \) for all \( p \neq q \), the likelihood in (3.22) is

\[
p(b_1, b_2, \ldots, b_R|g) = \prod_{r=1}^{R} N(n_r).
\]

Thus, for a group of \( R \) images satisfying the noise independency condition in (3.24), the likelihood can be modeled as

\[
p(b_1, b_2, \ldots, b_R|g) = \prod_{r=1}^{R} \prod_{l_1=1}^{L_1} \prod_{l_2=1}^{L_2} N(n_r(l_1, l_2)|0, \sigma_r),
\]

where \( \sigma_r \) represent the standard deviation of the Gaussian distribution for \( n_r \). This new model for the likelihood of multiple images will be used for the analysis of likelihood models in the next section.

Likelihood Models for Analysis

Out of the various likelihood models introduced in the literature of image restoration, we consider two recent approaches in [2] and [4] for our analysis.
In [2], the single image likelihood conforms to (3.21) and is explicitly given by
\[
p(b|g) \propto e^{-\frac{1}{2\sigma^2} \|Kg-b\|^2},
\]
where \( \| \cdot \| \) stands for the Frobenius norm.

In [4], the likelihood is defined with different orders of partial derivatives, denoted by operator \( \partial_r \), of a single degraded image. For ease of understanding, we represent their model in the form
\[
p(b|g) = \prod_{\partial_r \in \Theta} \prod_{l_1=1}^{L_1} \prod_{l_2=1}^{L_2} N(n(l_1, l_2)|0, \sigma) N(\partial_r n(l_1, l_2)|0, \sigma_{\omega(\partial_r)}), \quad r = 1, 2, 3, \ldots, R
\]
where \( \Theta \) is a set of partial derivative operators having,
\[
\Theta \triangleq \{ \partial_1, \partial_2, \partial_3, \ldots, \partial_R \}.
\]
For example, in [4], the set \( \Theta \) has the elements \( \{ \partial_x, \partial_y, \partial_{xx}, \partial_{xy}, \partial_{yy} \} \), in which, \( \partial_x \) is the first order derivative in \( x \) direction and \( \partial_y \) is the first order derivative in \( y \) direction and similar interpretations hold for higher order derivatives.

Further, [4] shows that the partial derivatives of \( n \) also follow Normal distributions with standard deviation values based on the order of the partial derivative operator. The standard deviations of the partial derivatives are specified in the form
\[
\sigma_{\omega(\partial_r)} = (\sqrt{2})^{\omega(\partial_r)} \sigma,
\]
where \( \omega(\partial_r) \) represents the order of the partial derivative operator \( \partial_r \). Few example elements of the set \( \Theta \) in (3.28) with the respective standard deviation values are given in Table 3.7.

As there was no analysis presented behind using the higher order partial derivatives of noise in [4] leading to (3.27), we provide an interpretation of formula, based on our new general likelihood model for a group of degraded images of a common ground truth \( g \) in (3.25).
Table 3.7: An example of set $\Theta$ in (3.28) with $R = 5$

<table>
<thead>
<tr>
<th></th>
<th>$\partial_r$</th>
<th>$\partial_r n$</th>
<th>$\omega(\partial_r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\partial_1$</td>
<td>$\partial_x n$</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$\partial_2$</td>
<td>$\partial_y n$</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>$\partial_3$</td>
<td>$\partial_{xx} n$</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>$\partial_4$</td>
<td>$\partial_{xy} n$</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>$\partial_5$</td>
<td>$\partial_{yy} n$</td>
<td>2</td>
</tr>
</tbody>
</table>

Guided by the likelihood expression (3.25), we can define a virtual group of images for the likelihood model in (3.27) as

$$b_r = \partial_r b, \quad r = 1, 2, 3, \ldots, R$$  \hfill (3.30)

and in order to align with model (3.23), define

$$K_r g \triangleq \partial_r (g * k), \quad r = 1, 2, 3, \ldots, R$$  \hfill (3.31)

$$n_r \triangleq \partial_r n, \quad r = 1, 2, 3, \ldots, R$$  \hfill (3.32)

where $*$ stands for the convolution operator and $k$ is the blur kernel.

From this, we infer that the likelihood (3.27) implicitly assumes $\partial_p n$ is independent of $\partial_q n$ for all $p \neq q$. Since all virtual images are derived from a single degraded image, we can infer this is a strong assumption made to simplify the likelihood expression. In principle, it should be possible to formulate a model without recourse to the derivative images which add limited new information. We corroborate this claim in Section 3.4.3 with experiments.

### 3.4.2 Frequency Domain Deconvolution

In this section we approach image deconvolution with FOPDO regularization and with different likelihood models discussed above. For our analysis, we consider the likelihood models of (3.21) and (3.27) using terminology “normal likelihood” and “derivative likelihood” with the notation using subscripts “$n$” and “$d$”, respectively. With our experiments, we limit the set $\Theta$ in (3.28), going up to second
order partial derivative operators in (3.27) and we take elements of \( \Theta \) from the following values

\[
\Theta = \{ \partial_x, \partial_y, \partial_{xx}, \partial_{xy}, \partial_{yy} \}.
\]  

(3.33)

Normal Likelihood Deconvolution

Under FOPDO regularization as detailed in Section 3.2.2, the stabilizing functional \( \Phi(g) \) takes the form

\[
\Phi(g) \triangleq \| \partial_x g \|_2^2 + \| \partial_y g \|_2^2.
\]  

(3.34)

Applying this stabilizing functional to the MAP framework detailed in Section 3.3.2, the energy functional under “normal likelihood”, can be derived as

\[
E_n(g) = \| g * k - b \|_2^2 + \lambda \Phi(g).
\]  

(3.35)

According to the convolution theorem, the convolution operation in the spatial domain becomes an element-wise product in the frequency domain making \( \mathcal{F}(g * k) = G \star K \) where \( \mathcal{F}(\cdot) \) stands for discrete Fourier transform, \( G \) for \( \mathcal{F}(g) \), \( K \) for \( \mathcal{F}(k) \) and “\( \star \)” denotes element-wise product. Based on the above property, transforming (3.35) into frequency domain and applying Plancherels theorem [111], we derive the energy in the frequency domain for (3.35) as follows.

\[
\mathcal{F}(E_n(g)) = \| G \star K - B \|_2^2 + \lambda \mathcal{F}(\Phi(g)),
\]  

(3.36)

where

\[
\mathcal{F}(\Phi(g)) = \| \mathcal{F}(\partial_x) \star G \|_2^2 + \| \mathcal{F}(\partial_y) \star G \|_2^2,
\]

\( B \) stands for \( \mathcal{F}(b) \) and given \( \partial_x \) takes the form of a (convolution) matrix, then \( \mathcal{F}(\partial_x) \) denotes its Fourier transform.

Minimizing the energy in (3.36) and solving for estimated \( G \) denoted as \( \hat{G} \) results in

\[
\hat{G}_n = \frac{B \star \overline{K}}{K \star \overline{K} + \lambda \Psi},
\]  

(3.37)
where,
\[ \Psi = \mathcal{F}(\partial_x) \ast \overline{\mathcal{F}(\partial_x)} + \mathcal{F}(\partial_y) \ast \overline{\mathcal{F}(\partial_y)}, \]

\( \hat{G}_n \) is the Fourier transform of the estimated GTI under "normal likelihood", \( (\cdot) \) stands for the complex conjugate and the division is performed element-wise. The estimated ground truth image \( \hat{g}_n \) can be derived by taking the inverse Fourier transform of \( \hat{G}_n \).

With the above derivations, it is evident that Fourier domain expression used to estimate the GTI is:

1. simple and leads to a closed form solution and
2. amenable to Fast Fourier Techniques leading to highly efficient solution.

**Derivative Likelihood Deconvolution**

We now give an analogous derivation for the “derivative likelihood”.

The energy functional in this case is derived similar to (3.35),

\[ E_d(g) = \sum_{\partial_r \in \Theta} \frac{1}{2\omega(\partial_r)} \| \partial_r (g \ast k) - \partial_r b \|^2_2 + \lambda \Phi(g). \tag{3.38} \]

Transforming (3.38) into the frequency domain results in

\[ \mathcal{F}(E_d(g)) = \sum_{\partial_r \in \Theta} \frac{1}{2\omega(\partial_r)} (\| \mathcal{F}(\partial_r) \ast G \ast K - \mathcal{F}(\partial_r) \ast B \|^2_2) + \lambda \mathcal{F}(\Phi(g)), \tag{3.39} \]

where, \( \partial_r \) is a matrix convolution operator representing a partial order derivative operator and \( \mathcal{F}(\partial_r) \) denotes its Fourier transformation.

By minimizing the energy (3.39), we compute the estimated \( G \),

\[ \hat{G}_d = \frac{B \ast \overline{K} \ast \Omega}{K \ast \overline{K} \ast \Omega + \lambda \Psi}, \tag{3.40} \]
Table 3.8: Comparison of likelihood models

<table>
<thead>
<tr>
<th>$\lambda \times 10^{-5}$</th>
<th>Normal likelihood $\times 10^{-4}$</th>
<th>Derivative likelihood $\times 10^{-4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2.8448</td>
<td>4.4371</td>
</tr>
<tr>
<td>5</td>
<td>2.4148</td>
<td>4.0862</td>
</tr>
<tr>
<td>0.25</td>
<td><strong>2.2292</strong></td>
<td>4.1852</td>
</tr>
</tbody>
</table>

where

$$\Omega \triangleq \sum_{\partial_r \in \Theta} \frac{1}{2\omega(\partial_r)} \mathcal{F}(\partial_r) \ast \overline{\mathcal{F}(\partial_r)}.$$ 

By taking the inverse Fourier transforms of (3.40), we could get the estimated GTI, $\hat{g}_d$, under “derivative likelihood” model similar to “normal likelihood” model.

### 3.4.3 Likelihood Model Analysis

In order to come up with the most effective and efficient restoration algorithm, we investigate the contribution of each of the likelihood models for estimating the GTI: (3.37) corresponding to “normal likelihood” and (3.40) corresponding to “derivative likelihood” respectively.

We used the same “Picasso” image which was used in [4] for experiments using the likelihood model in (3.27). The ground truth images are estimated using the Fourier domain techniques, specifically applying (3.37) and (3.40) for the “Normal” and “Derivative” likelihood models respectively. The experiment results are given in Table 3.8. In order to eliminate the “Model” and “Process” artifacts as discussed in Section 3.3.3, in all our simulations, the blurring was carried out under the assumption that the images are periodic.

The MSE values in the table are given as multiples of $10^{-4}$, while the value of $\lambda$ is given in multiples of $10^{-5}$. The values in **bold** in Table 3.8 refer to the optimal MSE values the respective likelihood model could reach for varying $\lambda$. As the results show clearly, the “normal likelihood” model has a better estimate for
the GTI than the “derivative likelihood” model, we claim that applying “normal likelihood” in the image restoration algorithm results in a better restoration.

Our investigation was further extended to analyze whether higher order derivatives of noise contribute to the spatial randomness of noise as claimed in [4]. The noise maps given in Fig. 3.11 are computed for different values of $\lambda$ in (3.37) and (3.40).

As per the results Fig. 3.11(c) and Fig. 3.11(d), when the effect of the prior becomes smaller (i.e., the weight on the data fitting term or the likelihood becomes larger), the noise estimate is more spatially random, but with the increase in the weight of the prior, the noise estimate becomes structured (signal dependant), see Fig. 3.11(e) and Fig. 3.11(f). We experienced these results regardless of the likelihood model we used. Based on the above results, we claim that using higher order partial derivatives in the likelihood model for non-blind deconvolution does not result in a better noise map estimation while the same noise map estimation can be achieved through the normal likelihood model with the appropriate Lagrange multiplier.

Hence, through the likelihood model analysis based on benchmark image, we conclude that higher order derivatives in the likelihood model are not required for better performance whereas applying single image likelihood model with appropriate regularization results in a more effective non-blind image restoration.

3.5 Contributions

In this chapter, we have contributed to regularization based image restoration techniques in the following:

1. We have developed a general class of quadratic regularization models based on partial derivative operators (PDO), Section 3.2.2. Out of those models, we have shown that the Second Order Partial Derivative Operator (SOPDO) model performs better than First Order Partial Derivative Operator (FOPDO) model for images susceptible to noise, while the novel First and Second Order Partial Derivative Operator (FSOPDO) model performs better than both FOPDO and SOPDO models.
Figure 3.11: Noise maps for Likelihood models
2. We have used the Structured Similarity index (SSIM) map, Mean SSIM (MSSIM) value and histograms of SSIM maps as novel visual metrics for comparison and evaluation of regularization models in image restoration, Section 3.3.3.

3. We have critically evaluated Sparse and Laplacian prior models against Quadratic regularization models using the novel visual metrics in Section 3.3.5. By eliminating the effects of processing and modelling artifacts, not present when capturing actual blurred natural images, we have shown that Sparse and Laplacian derivative prior models, which are claimed to be consistent with natural images, do not significantly contribute in restoring natural image features and have significantly slower relative restoration performance.

4. Finally, we have analyzed and evaluated the existing likelihood models under MAP/ML framework with a novel model to represent the likelihood based on multiple images, Section 3.4.1. By using this novel model, we demonstrate that complex higher order derivative likelihood models are not required for better performance in image restoration.

3.6 Conclusion

With the list of preceding contributions in image restoration through regularization and under Maximum A Posteriori (MAP) framework, we conclude that amidst the claims in the recent literature, highly complex restoration models do not add significant value over the more simple restoration models.
Chapter 4

Image Deconvolution through CMA

4.1 Overview

Blind deconvolution techniques are of great importance in modern high-efficiency communication systems. As a source signal propagates through a communication system, commonly referred as a channel, it gets distorted by the channel effects. Recovering the source signal in communication systems can be mainly categorized in two ways, system identification and equalization. System identification is the process of figuring out the characteristics of the distortion channel, which could be performed by observing the result of a known signal through the channel. Alternatively, equalization is the attempt of directly recovering the source signal, where the equalizer needs to closely emulate the inverse characteristics of the channel.

Channel equalization in data communication systems is generally performed by using a linear filter to equalize a dispersive channel impulse response. Conventionally, an equalizer is employed with the aid of a training sequence known to both the transmitting and receiving ends. This training session, however, can be rather costly or even unrealistic. To improve the overall throughput of a transmission system, the use of a training period is avoided by performing blind equalization [10].

To succeed with equalization, the equalizer often needs to have adjustable parameters, as the distortion characteristics are often varied with different channel
conditions. Further, the invention of automatic and adaptive channel equalization is considered to be a significant advance in data communications technology [112]. A general adaptive blind equalization system is shown in Fig. 4.1, where the equalizer parameter adjustments are based on the output signal through an objective function. The objective function evaluates the adjustments to the equalizer coefficients by measuring the alteration of the known characteristics of the source signal through the transmission process.

Blind adaptation is considered to be a challenging problem, as it must rely on only the system output and a limited amount of side information regarding the input. In addressing the poor performance in early blind adaptive algorithms, Sato in 1975 used the sign of the outputs instead of the actual outputs to guide adaptation [5]. Based on this, in 1980, Godard [6] made an important contribution to the field of adaptive equalization, where the equalization was designed on minimization of channel dispersion. This algorithm was also suggested by Treichler [7] and later referred as the Constant Modulus Algorithm (CMA). Despite its name, one of the most important features of CMA is that it can equalize constant modulus as well as non-constant modulus signals [113]. It is shown in [25], that CMA equalizers show better performance in terms of mean square error (MSE) and therefore in reducing inter symbol interference (ISI). Exploiting these features, Vural and Sethares employed CMA for blind image restoration in the image deconvolution context [44, 45, 114].
In this chapter, we follow and extend the work of Vural and Sethares [44, 45, 114], which looks at using CMA in the context of image restoration. In particular, we:

1. characterize the qualitative differences between deconvolution in the image processing context and deconvolution in the communications context,

2. develop a broad class of blind image deconvolution algorithms, in particular algorithms for blurring kernels with a separable property, which show significantly faster convergence than conventional algorithms and

3. study the properties of dispersion, an important quantity in blind image deconvolution algorithm cost functions, and derive scaling properties and optimal values.

Further, we corroborate the strengths of our novel separable CMA deconvolution algorithm to classical CMA deconvolution [45] through various simulations.

4.2 Image Deconvolution and channel equalization

Though CMA is widely used and recognized for its effectiveness and efficiency within the domain of digital communication systems, it cannot be directly applied in the context of image processing, as there are a number of qualitative and quantitative differences in the image deconvolution problem from the one-dimensional (1D) channel equalization problem. Few of those aspects which demand attention are:

1. **Multi-dimensional Data:** Clearly images are two-dimensional (2D) but also more recent trends in medical imaging are directed at three-dimensional imaging. The methods explored in this chapter are defined for the two-dimensional case and should generalize to the three-dimensional case.

2. **Limited/Poor Data:** One does not normally regard an image as occupying limited memory, but in comparison with the roots of blind deconvolution in
communication systems, the situation is quite different. A typical communication system is data rich, with typically millions of data packets, or a continuous stream, available to drive adaptation and aid identification. For an image the data is limited by the image size which is relatively small. The issue is statistical estimation is affected by the amount of data and so multiple processing passes over the one image are expected in image processing.

3. **Kernel Structure:** In communications, the distorting convolution generally does not exhibit much structure, such as symmetries in time. Of course, there are higher level structural aspects such as the communication channel tending to be minimum phase or approximately exponentially decaying, or sparse, etc. In contrast, the blurring convolution kernel in image processing can exhibit strong explicit symmetry, such as being isotropic (no preferred direction), or being separable in the two variables (horizontal and vertical) despite being anisotropic, or may be radially symmetric.

4. **Unstructured Source Modeling:** The source to be estimated in communication context is generally well modeled in a statistical sense, for example, the input may be binary or multi-level Pulse Amplitude Modulation (PAM). In contrast for images, there is no well defined sense that an image or set of images should have a fixed or known histogram *a priori*.

These aspects motivated us for developing blind image deconvolution algorithms with improvements in speed and accuracy of the processing relative to conventional blind communications-style deconvolution processing.

### 4.3 CMA on Image Deconvolution

#### 4.3.1 Adaptive System

The system used for adaptive blind deblurring, is shown in Fig. 4.2, where $x$ is the source to be estimated, $y$ is the known distorted source and $z$ is the output of the deconvolution process which estimates the source. This system of a blur-blind deblur system in image processing context can be considered as equivalent to a
channel-blind equalizer system in communication context. In the communications context, the source $x$ is the transmitted signal and the output $z$ becomes the equalized signal, while in the image processing context, source $x$ refers to the original image and the output $z$ is the deblurred image. In both contexts, CMA algorithm can be applied to recover an estimate for the unknown source $x$ from the known distorted source $y$. For simplicity, we have not considered the effect of noise in our blur-deblur model. In order to come up with the best estimate for the deblur, the weights of the “Adaptive Blind Deblur” are adjusted according to the CMA cost minimization, assuming the deblur to be a linear filter.

### 4.3.2 Blind Cost Functions

The CMA cost function [7], which is the archetype of a class of a set of cost functions [6], is given by the scalar positive functional

$$\mathcal{J}(z) \triangleq \frac{1}{4} E\{(z^2 - \gamma)^2\},$$

(4.1)

where $z$ is the output and $\gamma \in \mathbb{R}$ is the dispersion constant which depends on the source image histogram. More generally, (4.1) is seen to be a special case of a
broader case defined by Godard [6], parameterized by positive integer \( p \in \mathbb{Z} \), viz:

\[
J_p(z) \triangleq \frac{1}{2p} E\{(|z|^p - \gamma_p)^2\},
\] (4.2)

where \( p = 2 \) is the CMA cost, as given in (4.1), and \( p = 1 \) is the Sato algorithm [5, 115]. Godard mainly analyzed the cases \( p = 1 \) and \( p = 2 \), as higher orders gave no better performance. In (4.2), the notation indicates that the dispersion, \( \gamma_p \), depends on \( p \), indeed it is usually derived as [6]

\[
\gamma_p \triangleq \frac{E\{|x|^{2p}\}}{E\{|x|^p\}},
\] (4.3)

where \( x \) is the source image (pixels) to be estimated, see Fig. 4.2. This expression, (4.3), implies there is some probabilistic model for the source image pixels that characterizes all inputs. This makes sense in the communications context but is not well justified in the image case. That is, (4.3) is really a condition on the image histogram and it would be a poor assumption that all images have the same histogram. In Section 4.4.1, we explore the importance or otherwise of choosing the dispersion constant correctly\(^2\).

In the following, we drop the notation \( \gamma_p \) and simply use \( \gamma \) when the appropriate \( p \) can be inferred, whence

\[
J_p(z) \triangleq \frac{1}{2p} (|z|^p - \gamma)^2.
\] (4.4)

### 4.3.3 Blind Error Functions

From the Godard cost, (4.4), we can define the error functional

\[
\phi_p(z) \triangleq \frac{\partial J_p(z)}{\partial z} = \frac{1}{2p} \left\{ \frac{1}{2p} (|z|^p - \gamma)^2 \right\} = |z|^{p-1} (|z|^p - \gamma) \sgn(z),
\] (4.5)

\(^{\text{1}}\)In the expressions, \(| \cdot |\) can be interpreted as modulus as the formulation extends to the complex case. In the real case, \(| \cdot |\) is the absolute value and when raised to an even power is superfluous.

\(^{\text{2}}\)It’s not important.
where $\text{sgn}(\cdot)$ is the signum function.\(^3\) This expression can be further simplified as

$$
\phi_p(z) = z |z|^p - 2 (|z|^p - \gamma).
$$

(4.6)

### 4.3.4 Algorithms

Blind algorithms based on the use of a cost function can be formulated as a gradient descent technique. To define the algorithm from the cost function, or equivalently the error function, we compute the gradient

$$
\frac{\partial J_p(z)}{\partial w(i)} = \frac{\partial J_p(z)}{\partial z} \frac{\partial z}{\partial w(i)}, \quad \forall i \in \mathcal{I}
$$

(4.7a)

$$
= \phi_p(z) \frac{\partial z}{\partial w(i)}, \quad \forall i \in \mathcal{I}
$$

(4.7b)

where \(\{w(i)\}_{i \in \mathcal{I}}\) are parameters in our blind adaptive deconvolution filter, \(\mathcal{W}\), and \(i\) varies over some abstract index set \(\mathcal{I}\) (our adjustable parameters)\(^4\). From the gradient we have the update equation, \(\forall i \in \mathcal{I}\),

$$
w(i)_{k+1} = w(i)_k - \alpha \phi_p(z_k) \frac{\partial z_k}{\partial w(i)} \bigg|_{w(i) = w(i)_k},
$$

(4.8)

which can be written using (4.7b) as

$$
w(i)_{k+1} = w(i)_k - \alpha \phi_p(z_k) \frac{\partial z_k}{\partial w(i)} \bigg|_{w(i) = w(i)_k}
$$

(4.9)

where \(\alpha \in \mathbb{R}\) is a sufficiently small positive step size, \(k\) is an index of iteration (usually time in the communications context but here space in the image context), and \(z_k\) is the deconvolved signal at index \(k\). We can rewrite (4.9) in vector form

$$
\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha \phi_p(z_k) \frac{\partial z_k}{\partial \mathbf{w}} \bigg|_{\mathbf{w} = \mathbf{w}_k}
$$

(4.10)

---

\(^3\)Taking values \(\pm 1\) and zero for the zero argument.

\(^4\)This index set abstractly represents either a 1D or 2D setup, and the indices are not meant to be the same as those in a tapped delay line.
\[ \beta^{-1/p} \quad \beta^2 \quad \beta > 0 \]

Figure 4.3: Scaling properties of Blind Cost Function, \( J_p(\cdot) \), given in (4.4), with respect to scaling the dispersion by \( \beta > 0 \).

\[ \beta^{-1/p} \quad \beta^{2-1/p} \quad \beta > 0 \]

Figure 4.4: Scaling properties of Blind Error Function, \( \phi_p(\cdot) \), given in (4.6), with respect to scaling the dispersion by \( \beta > 0 \).

where \( \partial z_k / \partial w \) can be interpreted as the regressor vector whose dimension is the cardinality of the index set, \(|I|\) or equivalently the number of parameters that can be adjusted in the blind adaptive deconvolution filter. This regressor depends only on the parametrization structure of the \( \mathcal{W} \) adaptive filter, see Fig. 4.2.

### 4.4 Dispersion Analysis

We claim the determination of \( \gamma \), is not critical because it does not affect, on average, the deconvolution aspect and only affects the scale or dynamic range (equivalent to contrast in the image processing context) — it is straightforward to adjust the final deconvolved image to have an acceptable contrast. In fact we can say something stronger about the importance or otherwise of the choice of the dispersion \( \gamma \) based on the following development and analyzed in later sections.
4.4.1 Scaling Properties of the Dispersion

Define

\[ J_p^{(\beta)}(z) \equiv \frac{1}{2p}(|z|^p - \gamma \beta)^2, \]  

(4.11)

which introduces a scaling factor of \( \beta > 0 \) into the dispersion. Therefore, comparing with expression (4.4), \( J_p(z) = J_p^{(1)}(z) \), where \( \beta = 1 \). Further, into (4.4), we introduce \( \beta^{-1/p} \),

\[ J_p(\beta^{-1/p}z) = \frac{1}{2p} (\beta^{-1} |z|^p - \gamma)^2 \]  

(4.12a)
\[ = \frac{1}{2p\beta^2} (|z|^p - \gamma \beta)^2, \]  

(4.12b)

that is,

\[ J_p^{(\beta)}(z) = \beta^2 J_p(\beta^{-1/p}z). \]  

(4.13)

So, as depicted in Fig. 4.3, scaling the dispersion by a positive real multiplier, that is, going from \( \gamma \) to \( \gamma \beta \), is equivalent to scaling the input of the cost function by \( \beta^{-1/p} \) and scaling the output by \( \beta^2 \).

Similarly, define

\[ \phi_p^{(\beta)}(z) \equiv z |z|^{p-2}(|z|^p - \gamma \beta), \]  

(4.14)

from which we can infer, from (4.6), with some calculation,

\[ \phi_p^{(\beta)}(z) = \beta^{2-1/p} \phi_p(\beta^{-1/p}z), \]  

(4.15)

that is, as depicted in Fig. 4.4, scaling the dispersion going from \( \gamma \) to \( \gamma \beta \), is equivalent to scaling the input of the error function by \( \beta^{-1/p} \) and scaling the output by \( \beta^{2-1/p} \).

4.4.2 Effect of Dispersion on Adaptation

With reference to Fig. 4.5, we investigate the effect of different values of the dispersion on adaptation. This further develops the results we had on scaling the cost function, (4.13), and error function, (4.15).

Our “Reference System”, in the upper portion of Fig. 4.5, is the system de-
Figure 4.5: Diagram showing the effect of dispersion on blind adaptation. The “Reference System” is the upper portion and the “Scaled System” is the lower portion. The two $\mathcal{W}$ adaptive filters are identical systems and may differ only in the specific values taken on by the adapted weights.

scribed by (4.9). We make an important observation regarding the regressor term in (4.9), which we write here as

$$\frac{\partial z_k}{\partial w(i)_k} \triangleq \frac{\partial z_k}{\partial w(i)} \bigg|_{w(i)=w(i)_k}, \quad i \in \mathcal{I}. \tag{4.16}$$

As a function (4.16) depends only on the $\mathcal{W}$ filter implementation (structure and parametrization) and is independent of the adaptation algorithm, and, therefore, independent of $\beta > 0$. For example, if $\mathcal{W}$ is a tapped delay line, see Section 4.5.1, and the weights are the delay coefficients, then this regressor takes the form $\partial z_k / \partial w(i)_k = y_{k-i}$ (the input to the filter at $k-i$). Using (4.16), we rewrite (4.9) as

$$w(i)_{k+1} = w(i)_k - \alpha \phi_p(z_k) \frac{\partial z_k}{\partial w(i)_k} \tag{4.17}$$

for all for $i \in \mathcal{I}$. Our “Scaled System”, in the lower portion of Fig. 4.5, is then

$$w(i)_{k+1}^{(\beta)} = w(i)_k^{(\beta)} - \alpha^{(\beta)} \phi_p^{(\beta)}(z_k^{(\beta)}) \frac{\partial z_k^{(\beta)}}{\partial w(i)_k^{(\beta)}}$$
\[ w(i)_k^{(\beta)} - \alpha^{(\beta)} \beta^{2-1/p} \times \]
\[ \phi_p(\beta^{-1/p}z_k^{(\beta)}) \frac{\partial z_k^{(\beta)}}{\partial w(i)_k^{(\beta)}}, \] (4.18)

where we have used (4.15). We can now state a basic result that relates (4.18) to (4.17):

**Theorem 1** (Dispersion Scaling). For the two systems in Fig. 4.5, a Reference System with dispersion given by \( \gamma \) and a Scaled System with dispersion given by \( \gamma \beta \), having identical \( W \) adaptive filters with identical initial states at iteration \( k = 0 \), and whose respective weights satisfy

\[ w(i)_0^{(\beta)} = \beta^{1/p}w(i)_0, \quad \forall i \in I \] (4.19)

and respective step sizes satisfy the relation

\[ \alpha^{(\beta)} = \beta^{2/p-2}\alpha \] (4.20)

then for all \( k > 0 \)

\[ w(i)_k^{(\beta)} = \beta^{1/p}w(i)_k, \quad \forall i \in I \] (4.21)

and, therefore,

\[ z_k^{(\beta)} = \beta^{1/p}z_k. \] (4.22)

**Proof.** By the indicated substitutions, (4.18) becomes equivalent to (4.17), carefully noting the definition of (4.16).

**Comments**

1. The choice of dispersion \( \gamma \), or \( \gamma \beta \), does not make an essential change to the adaptation.

2. There is a natural scaling of the outputs, captured by (4.22), which will be a tendency during adaptation rather than equality if the initializations do not satisfy (4.19).
3. Any equilibria of the Scaled System are at a scaled factor times the equilibria of the Reference System.

4. The speed of convergence is not affected by the value of $\beta > 0$ whenever the step sizes satisfy (4.20).

5. There is no affect of dispersion on adaptation once appropriate scalings are accounted for. This finding is relevant for the image processing context since there is no sensible way to assign a consistent dispersion value to different images. Fortunately, with CMA, there is no need to.

4.5 Separable Linear Deconvolution

In this section, we develop a novel separable implementation of the linear filter $W$. First we recall the standard direct parametrization, Section 4.5.1, to compare with the new separable parametrization, Section 4.5.2.

4.5.1 Direct Parametrization

Normally, the parameters $w \equiv \{w(i)\}_{i \in I}$ are used as independent weights, corresponding to each delay, in a tapped delay line structure to explicitly effect the deconvolution, as a linear filter $W$, see Fig. 4.2. This is pervasive in communication algorithms (1D case) and is easily generalized to the 2D image deconvolution, as is done in the CMA image deconvolution [114]. Under adaptation, each weight is independently adjustable.

The direct parametrization is the 2D convolution

$$z(r, s) = \sum_{i=-N}^{N} \sum_{j=-N}^{N} w(i, j) y(r - i, s - j), \quad (4.23)$$
comprising of \((2N + 1)^2\) parameters \(\{w(i, j)\}\).\(^5\) Whence, in the notation of (4.16),

\[
\frac{\partial z(r, s)}{\partial w(i, j)} = y(r - i, s - j)
\]  

(4.24)

and, then (4.17) takes the form (direct parametrization)

\[
w(i, j)_{k+1} = w(i, j)_{k} - \alpha \phi_p(z(r, s)) y(r - i, s - j),
\]  

(4.25)

at the \(k^{th}\) iteration of the algorithm update which is evaluated at pixel \((r, s)\). Equation (5.6) is evaluated for all \((2N + 1)^2\) combinations of adaptation weights indexed by \(-N \leq i \leq N\) and \(-N \leq j \leq N\).

### 4.5.2 Separable Parametrization

We define a separable parametrization as the 2D convolution

\[
z(r, s) = \sum_{i=-N}^{N} \sum_{j=-N}^{N} w(i)w(j) y(r - i, s - j)
\]  

(4.26)

or equivalently, for \(W: y \mapsto z\),

\[
z(r, s) = \sum_{i=-N}^{N} w(i) \sum_{j=-N}^{N} w(j) y(r - i, s - j)
\]  

(4.27)

comprising of \((2N + 1)\) parameters \(\{w(i)\}\). It is a rank one kernel representation and is a suitable parametrization when the blurring kernel itself is separable. Whence, in the notation of (4.16),

\[
\frac{\partial z(r, s)}{\partial w(i)_k} = \sum_{j=-N}^{N} w(j)_k (y(r - i, s - j) + y(r - j, s - i))
\]  

(4.28)

\(^5\)Note that having doubly indexed weights can be made trivially consistent with the abstract index set, \(I\) which has a single index.
and, then (4.17) and (4.23) takes the form (separable parametrization)

\[
    w(i)_{k+1} = w(i)_k - \alpha \phi_p(z(r, s)) \frac{\partial z(r, s)}{\partial w(i)_k}, \quad \forall i
\]  

\[
    z(r, s) = \sum_{i=-N}^{N} w(i)_k \sum_{j=-N}^{N} w(j)_k y(r - i, s - j),
\]

at the \( k \)th iteration of the algorithm update which is evaluated at pixel \((r, s)\).

### 4.5.3 Dispersion under Separability

For the direct parametrization (4.23), it is known that the dispersion is given by (4.3), [6]. Under the separable parametrization, (4.27), which leads to a more complicated regressor expression it is less clear that this dispersion relation applies.

**Theorem 2** (Dispersion under Separability). For the separable blind adaptation defined through equations (4.29) and (4.30), then the ideal dispersion, implicit in \( \phi_p(z(r, s)) \) given in (4.6), is given by

\[
    \gamma \triangleq \frac{E\{|x|^{2p}\}}{E\{|x|^p\}},
\]

where \( x \) is the source image.

**Proof.** The objective is to choose \( \gamma \) such that (4.29) has an equilibrium at the ideal convolution setting. Let \( w(i) = w(i)^{eq} \) be an equilibrium, not necessarily the desired ideal equilibrium. With these equilibrium weights, the regressor takes the form

\[
    \frac{\partial z(r, s)}{\partial w(i)} \bigg|_{w(i)=w(i)^{eq}} = \sum_{j=-N}^{N} w(j)^{eq} \left( y(r - i, s - j) + y(r - j, s - i) \right).
\]

Equilibria occur when (4.29) has stationary updates, that is,

\[
    E \left\{ \phi_p(z(r, s)) \frac{\partial z(r, s)}{\partial w(i)} \right\} \bigg|_{w(i)=w(i)^{eq}} = 0, \quad \forall i.
\]
Substituting for the regressor, (4.32), leads to

\[
E\left\{ \phi_p(z(r, s)) \sum_{j=-N}^{N} w(j)^{eq}(y(r-i, s-j) + y(r-j, s-i)) \right\} = 0, \forall i. \tag{4.34}
\]

which implies the linear combination is also zero:

\[
\sum_{i=-N}^{N} w(i)^{eq} E\left\{ \phi_p(z(r, s)) \sum_{j=-N}^{N} w(j)^{eq}(y(r-i, s-j) + y(r-j, s-i)) \right\} =
E\left\{ 2z(r, s)\phi_p(z(r, s)) \right\} = 0
\]

where we have used (4.30). Substituting in for (4.6), we obtain

\[
E\left\{ z(r, s) \phi_p(z(r, s)) \right\} = E\{|z(r, s)|^2 |z(r, s)|^{p-2}(|z(r, s)|^p - \gamma)\}
\]

\[
= E\{|z(r, s)|^p (|z(r, s)|^p - \gamma)\}. \tag{4.35}
\]

Equation (4.35) holds for any equilibrium. At the desired weights, \( z(r, s) = x(r, s) \), and this can be made an equilibrium by choosing \( \gamma \) as in (4.31).

### 4.5.4 Simulations

The simulations were carried out on a 256 \( \times \) 256 image. For processing, the image was scaled and shifted to make it zero mean. The input for our simulation algorithm was the blurred image created by convolving the original image with the two dimensional separable blur kernel with components

\[
k = \begin{bmatrix} 1 & 0.5 & 0.25 & 0.125 & 0.0625 & 0.03125 \end{bmatrix}',
\tag{4.36}
\]

that is, the kernel blur matrix \( k \) is the outer product \( kk' \). The system was tested with different dispersion values, with initializations done as per (4.19). The result in Fig. 4.6 shows the difference (\( d \)) between the actual (\( w_{\text{actual}} \)) and estimated (\( w_{\text{estimated}} \)) deblur kernels when going through the adaptive blind deblur system.

\[
d = \| w_{\text{actual}} - w_{\text{estimated}} \|_F, \tag{4.37}
\]
Figure 4.6: Difference between the actual and estimated deblur kernels with the dispersion change
where $F$ stands for Frobenius Norm. From the results, it is evident that the speed of convergence is not affected by the changes of the dispersion value.

Next, the performance of the two algorithms, one based on the novel separable parametrization and the other on the direct parametrization were compared. The results in Fig. 4.7 shows the difference $d$ on CMA deconvolution with direct and separable parametrization.

Theoretically, the convolution of the blur kernel and the deblur kernel should result in an impulse response, as the purpose of deblur is to inverse the effect of the blur kernel. Considering this, we carried out few simulations to test the error $e$, defined as

$$e = \|\delta - k \otimes w_{estimated}\|_F,$$

(4.38)

where, $\delta$ is the impulse response of the ideal blur-free system and $\otimes$ stands for the convolution. The estimation of $e$ on CMA deconvolution with direct and separable parametrization is shown in Fig. 4.8. All these simulations show clearly that the
separable parametrization algorithm has a faster rate of convergence than the
direct parametrization algorithm in corroborating the aforementioned theoretical
analysis.

The CMA deconvolution performance with respect to direct and separable
parameterizations on the “Lena” image are shown in Fig. 4.9. The superiority of
the separable parametrization estimate requiring less iterations is clearly evidenced
in the results, while the direct parametrization estimate becomes more comparable
to the separable result with the increasing number of iterations.

4.6 Contributions

By analyzing blind image restoration with CMA in this chapter, we contribute in

1. analyzing the importance of an accurate choice of the dispersion constant,
As the statistics of the source image is not available at the time of deconvolution, this analysis is essential in order to come up with an accurate estimate of the ground truth image. Using Theorem 1, we prove that there is no affect of dispersion on adaptation once appropriate scalings are accounted for.

2. developing a novel, efficient, two dimensional (2D), blind deconvolution algorithm for restoring images corrupted by an unknown 2D blurring kernel satisfying a separable property. By exploiting the separable property of kernels there was a substantial speedup relative to an unstructured 2D blurring kernel. That is, for a $2N+1 \times 2N+1$ kernel, the complexity is improved by a factor of $O(N)$, where, the reduction in parameters greatly improves speed of convergence, robustness and accuracy of the deconvolution.

4.7 Conclusion

In this chapter, we developed a broad class of blind image restoration algorithms using CMA, which show significantly faster convergence than conventional algorithms.
Figure 4.9: Lena original image, blurred image and a comparison between the classical CMA with direct parametrization and our novel CMA with separable parametrization.
Chapter 5

Analysis of CMA Blind Image Deconvolution in Practice

5.1 Overview

While the Godard class of algorithms have achieved considerable success in channel equalization, their convergence to undesirable equilibria has been emphasized by many researchers [115–118]. The ill-convergence of Constant Modulus Algorithm (CMA) in channel equalization mainly relates to presence of noise in the channel, poor initialization and the violation of the assumptions made on the source signal, namely zero-mean, independent, uniformly distributed and constant-modulus (circularly symmetric when complex) [11, 119]. Specially, in [120], it is shown that a temporally correlated source inhibits the perfect channel equalization under CMA. Out of the assumptions made on the source, as it is shown in [113] that constant modulus (CM) criterion works almost as well with non-CM sources, we consider the assumptions of source being zero-mean, independent and uniformly distributed as the critical assumptions which are required to assist the good convergence of CMA deconvolution.

These assumptions made on CMA in the communication context stress the importance of investigating them in the image processing context, as there are significant differences between image deconvolution and channel equalization [121]. Hence, the main thrust of this chapter deals with the examination of the as-
assumptions made on the source in CMA image restoration and the impact of their violations. In addition to that analysis, we develop a novel model to address the hindrance caused by spatial correlation in images on blind image restoration through CMA. Summarizing, in this chapter, we contribute in:

1. Analyzing the effects of assumptions made on CMA deconvolution in image restoration, mainly the assumptions of
   (a) source of uniform distribution,
   (b) white source, and
   (c) source being zero-mean.

2. Modeling the relationship between pixel correlations and higher order moments, thereby developing a segmentation based CMA deconvolution to overcome the spatial correlation effects.

3. Discussing the use and effectiveness of higher order moments on natural image deconvolution.

5.2 Preamble

Consider a blind image deconvolution system shown in Fig. 5.1, where $g$ is the unknown ground truth image (GTI), $k$ is the unknown blur kernel, $b$ is the blurred image, $w$ is the adaptive deblur kernel and $z$ is the estimated GTI.
For a blurred image $b$ of finite support $[L_1, L_2]$, and a deblur kernel $w$ of support $[M_1, M_2]$, the deblur process can be mathematically represented as

$$z(l_1, l_2) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} w(m_1, m_2) b(l_1 - m_1, l_2 - m_2),$$  \hspace{1cm} (5.1)$$

where $(l_1, l_2) \in [0, L_1 - 1] \times [0, L_2 - 1]$.

As detailed in Chapter 3, applying CMA deconvolution to the system in Fig. 5.1, the parameters of the adaptive deconvolution filter $w$, could be updated iteratively with

$$w_{n+1} = w_n - \alpha \phi(z_n) \frac{\partial z_n}{\partial w} \bigg|_{w=w_n},$$  \hspace{1cm} (5.2)$$

where

$$\phi(z_n) = z_n \left( z_n^2 - \gamma \right),$$  \hspace{1cm} (5.3)$$

the dispersion constant, $\gamma$ is defined as

$$\gamma \triangleq \frac{E\{|g|^4\}}{E\{|g|^2\}},$$  \hspace{1cm} (5.4)$$

$\alpha \in \mathbb{R}$ is a sufficiently small positive step size and $n$ represents the index of iteration. With the deconvolution defined as in (5.1), under direct parametrization on classical CMA adaptation, discussed in Chapter 3, the regressor vector simplifies into

$$\frac{\partial z(l_1, l_2)}{\partial w(m_1, m_2)} = b(l_1 - m_1, l_2 - m_2),$$  \hspace{1cm} (5.5)$$

resulting in (5.2) taking the form

$$w_{n+1}(m_1, m_2) = w_n(m_1, m_2) -$$
$$\alpha \phi(z_n(l_1, l_2)) b(l_1 - m_1, l_2 - m_2),$$  \hspace{1cm} (5.6)$$

at the $n^{th}$ iteration of the algorithm update which is evaluated at pixel $(l_1, l_2)$. 
5.3 Statistical Distribution of Ground Truth Image

One of the assumptions made under CMA and related algorithms for perfect source recovery (PSR) is the non-Gaussian distribution of the source signal \([113,115]\). This assumption is further studied in \([119]\), showing the importance of source statistics in channel equalization through CMA. In this section, we analyze the effects of the distribution of the GTI in image restoration through CMA.

5.3.1 Kurtosis

Kurtosis is a concept which is closely related with the cost function of CMA. Kurtosis, \(C_x\) of a random variable \(x\), is defined as the normalized fourth central moment assuming zero mean,

\[
C_x = \frac{E\{|x|^4\}}{E\{|x|^2\}^2}.
\]  
(5.7)

Then the relationship between the dispersion constant for ground truth image \(g\) and kurtosis becomes

\[
\gamma = \sigma_g^2 C_g,
\]  
(5.8)

where

\[
\sigma_g^2 = E\{|g|^2\}.
\]

Kurtosis can be regarded as a simple operational measure of sparseness, where the normal (Gaussian) distribution has a kurtosis value of 3\(^1\) and the uniform distribution has a kurtosis value of 1.8. Distributions which have kurtosis lower than the value of 3 are called platykurtic (sub-Gaussian) distributions, while those with kurtosis values higher than 3 are called leptokurtic (super-Gaussian) distributions. Based on this classification, we carried out a series of simulations as to test the effects of source kurtosis on CMA deconvolution. Most of the images used for these simulations are extracted from the data used for the evaluations in \([123]\) and

\(^1\)Some books and papers define kurtosis as \(C_x = \frac{E\{|x|^4\}}{E\{|x|^2\}^2} - 3\), making the Gaussian distribution having a kurtosis of value 0. However, we will follow the definition of (5.7) as found in [122].
are shown in Fig. 5.2.

5.3.2 Histogram Equalization

Out of many differences between the deconvolution in the image processing context and equalization in the data communication context [121], one of the main differences is the structure and distribution of the source. While in data communication context, it is a standard assumption to take the source as of uniform distribution, in image processing context, the natural GTI distribution could vary over a range of distributions from platykurtic to leptokurtic.

In order to meet the assumption of uniform source distribution in image restoration through CMA, previous works [45, 114] have applied histogram equalization [124] on the GTI as a pre-processing step to CMA deconvolution. The adaptive blind deblur system with histogram equalization as a pre-processing step is shown in Fig. 5.3.

We investigate the effects of source distribution on the convergence of CMA image deconvolution through the two systems shown in Fig. 5.1 and Fig. 5.3. For impartial comparisons, we maintained the same environment for both systems by using the same step size $\alpha$ and other parameters. The deconvolution results are shown in Fig. 5.3, where Fig. 5.3(a) to Fig. 5.3(i) correspond to the GTIs of Platykurtic distribution (Fig. 5.2(a), Fig. 5.2(d) and Fig. 5.2(e)), Meso-kurtic distribution (Fig. 5.2(b), Fig. 5.2(f) and Fig. 5.2(g)) and Leptokurtic distribution (Fig. 5.2(c), Fig. 5.2(h) and Fig. 5.2(i)) in order. Each of the deconvolution result plots the difference ($d$) between the actual ($w_{\text{actual}}$) and estimated ($w_{\text{estimated}}$) deblur kernels against the number of pixel estimations when adapting through the blind deblur system. The evaluation of $d$ corresponds to

$$d = \|w_{\text{actual}} - w_{\text{estimated}}\|_F,$$

where $F$ stands for Frobenius norm.
Figure 5.2: Natural images used for testing distributional aspects.
5.3.3 Source Distribution Analysis

The results shown in Fig. 5.3 are analyzed in three categories based on the distribution of the GTI.

1. Platykurtic source: The CMA deconvolution results shown in Fig. 5.3(a) to Fig. 5.3(c) are for GTIs with platykurtic distributions. As evidenced by the results, when the source distribution is closer to a uniform distribution (Fig. 5.3(a) and Fig. 5.3(b)), the deconvolution ends up with similar results for systems shown in Fig. 5.1 and Fig. 5.3. In communication context, it is shown that, for PSR in CMA, the source distribution has to be sub-Gaussian [11, 113]. In image restoration through CMA, we experienced results supporting this claim. The reasons behind our deconvolution results not achieving PSR will be discussed in detail in Section 5.4.

2. Meso-kurtic source: A standard result in blind equalization schemes including CMA is that a Gaussian source causes ill-behavior in the equalization [125]. The term meso-kurtic is introduced in [119] referring to the class of distributions (including Gaussian sources) which has a kurtosis around value 3 that cause such ill-behavior. In our simulations with meso-kurtic sources (Fig. 5.3(d) to Fig. 5.3(f)), we experienced undesirable qualities in the CMA convergence if the source is not histogram equalized. Thus, the GTIs with meso-kurtic distributions lead to slow CMA convergence and poor performance in image restoration.

3. Leptokurtic source: It is shown in [120] that leptokurtic sources result in
disastrous performance, maximizing inter-symbol interference (ISI) in channel equalization. Our simulations on blind image restoration through CMA indicate such behavior as shown in Fig. 5.3(g) to Fig. 5.3(i). It can be clearly seen that the concerns of poor CMA performance for Meso-kurtic sources are heightened more when the GTI becomes more sparse, having a higher kurtosis value. Hence, natural GTIs with sparse distributions will end up having prolonged convergence times and undesired behavior under CMA image deconvolution.

As a similar trend of results is observed on CMA deconvolution with a range of natural images, we infer that the shape of the statistical distribution of the GTI hinders the CMA image deconvolution ability.

Upon investigating the impact of the source distribution in real world blind image restoration problems, one can see that it becomes a major concern, as in practice, the deconvolution system neither can access GTI nor its statistical properties. Comparing the two systems shown in Fig. 5.1 and Fig. 5.3, it reflects that CMA deconvolution without histogram equalization more accurately models the image restoration system in practice. As histogram equalization is a non-linear transformation, the process of blur and histogram equalization in Fig. 5.3 could not be interchanged as to make the system applicable in realistic situations.

With the above analysis, we have demonstrated that the statistical distribution of the GTI affects the results of CMA image restoration in practice, except for those GTIs which have their distribution closer to uniform distribution (Platykurtic).

5.4 Spatial Correlation of Images

The Godard class of blind equalization algorithms including CMA are based on the critical assumption of temporal independence [113, 119]. It is shown in [11] that while any amount of source correlation violates the global convergence requirements of CMA, large amounts of correlation may even cause additional false minima to appear on the constant modulus cost surface in the parameter space.

Although the effect of source correlation is analyzed in detail in communication context [119, 126], no mathematical analysis has been performed on the effect of
5.4 Spatial Correlation of Images

(a) $C_g = 2.0570$

(b) $C_g = 2.1082$

(c) $C_g = 2.4737$

(d) $C_g = 3.2449$

(e) $C_g = 3.0596$
Figure 5.3: CMA deconvolution results for the images in Fig. 5.2.
source correlation in image deconvolution through CMA. Hence, we do a detailed analysis on the effect of image pixel correlation on CMA deconvolution.

### 5.4.1 Source Correlation Effect in Image Deconvolution through CMA

The temporal independence in channel equalization becomes spatial independence in image restoration and it is shown that PSR through CMA image restoration could be achieved when the spatial correlation of the image is eliminated through randomization [127]. The representation in (5.1) can be represented in matrix notation as [26, pp. 216-218] for an image of support \([L_1, L_2]\), and a deblur kernel of support \([M_1, M_2]\),

\[
\begin{bmatrix}
  z_0 \\
  z_1 \\
  \vdots \\
  z_{M_1+L_1-1}
\end{bmatrix}
\triangleq
\begin{bmatrix}
  [W_0] & [W_0] & \cdots & [W_0] \\
  [W_1] & [W_1] & \cdots & [W_1] \\
  \vdots & \vdots & \ddots & \vdots \\
  [W_{M_1-1}] & [W_{M_1-1}] & \cdots & [W_{M_1-1}]
\end{bmatrix}
\begin{bmatrix}
  b_0 \\
  b_1 \\
  \vdots \\
  b_{L_1-1}
\end{bmatrix},
\]

(5.9)

where the rows of the lexicographic orderings, stacked as columns are subscripted:

\[
\begin{align*}
  z(J) &= \begin{bmatrix}
  z(j, 0) \\
  z(j, 1) \\
  \vdots \\
  z(j, M_2 + L_2 - 1)
\end{bmatrix} \\
  b(i) &= \begin{bmatrix}
  b(i, 0) \\
  b(i, 1) \\
  \vdots \\
  b(i, L_2 - 1)
\end{bmatrix}
\end{align*}
\]
The matrices \([W_j]\) are formed from the \(j^{th}\) row of the deblur matrix \(w\) according to

\[
[W_j] = \begin{bmatrix}
w(j, 0) & w(j, 0) \\
w(j, 1) & w(j, 0) \\
\vdots & \ddots & \ddots \\
w(j, M_2 - 1) & \vdots & \ddots & w(j, 0) \\
w(j, M_2 - 1) & w(j, 1) \\
\vdots & \ddots & \ddots & \ddots \\
w(j, M_2 - 1) & & & & w(j, M_2 - 1)
\end{bmatrix}
\]

With the above notation, (5.9) can be written in the form of

\[
z = [W]b,
\]

where \([W]\) is called the block-Toeplitz-block (BTB) matrix of \(w\). Similarly, the blur process can also be represented as

\[
b = [K]g,
\]

resulting in

\[
z = [W][K]g.
\]

Under gradient descent algorithms, the CMA error surface stationary points are given by \(\frac{\partial J}{\partial w} = 0\). The gradient of the CMA cost function

\[
J(z) \triangleq \frac{1}{4}(z^2 - \gamma)^2
\]

computes into

\[
\frac{\partial J(z)}{\partial w} = \frac{\partial J(z)}{\partial z} \frac{\partial z}{\partial w},
\]

\[
= \phi(z) \frac{\partial z}{\partial [W]} \frac{\partial [W]}{\partial w},
\]

\[
= E\{\phi(z) [K]g [H]\},
\]
where $[H]$ is a BTB matrix with

$$
[H] \triangleq \frac{1}{M_1} \begin{bmatrix}
  [H_0] \\
  [H_1] & [H_0] \\
  \vdots & \vdots & \ddots \\
  [H_{M_1-1}] & \cdots & [H_0] \\
  [H_{M_1-1}] & [H_1] & \cdots \\
  \vdots & \vdots & \ddots \\
  [H_{M_1-1}] & \cdots & [H_0] \\
\end{bmatrix},
$$

(5.15)

where each $[H_j]$ represents a Toeplitz matrix with the structure of $[W_j]$ having identity values for the diagonal starting with row $j$. For example,

$$
[H_0] = \begin{bmatrix}
  1 \\
  0 & 1 \\
  \vdots & \vdots & \ddots \\
  0 & \cdots & 1 \\
  \vdots & \vdots & \vdots \\
  0 & \cdots & 0
\end{bmatrix}.
$$

Further simplifying (5.14), the gradient of the CMA cost function can be written in terms of the blur ($[K]$) and deblur ($[W]$) kernel parameterizations as

$$
\frac{\partial J}{\partial w} = E\{([W][K]g)^3 - \gamma [W][K]g[K]g[H]\}.
$$

(5.16)

As $[K]$ and $[H]$ are full column rank matrices, (5.16) implies that the stationary points of CMA error surface are given by,

$$
E\{([W][K]g)^3 - \gamma [W][K]g)g = 0.
$$

(5.17)

Let $[D]$ be the combined blur-deblur kernel response, represented by

$$
[D] = [W][K].
$$
Then, (5.17) can be simplified into

\[ E\{(|D|g)^3 - \gamma [D]g\} = 0. \tag{5.18} \]

Taking the expectation operator with respect to ground truth image, (5.18) can be expanded into a system of equations to solve for the stationary points in CMA cost function. For example, let \( d_0, \ldots, d_r \) be the elements of the first row of the matrix \([D]\), then the first multivariate polynomial equation results in

\[
E\{(d_0 g(0, 0) + d_1 g(0, 1) + \ldots + d_r g(L_1 - 1, L_2 - 1))^3 - \gamma (d_0 g(0, 0) + d_1 g(0, 1) + \ldots + d_r g(L_1 - 1, L_2 - 1)) [g(0, 0) g(0, 1) \ldots g(L_1 - 1, L_2 - 1)]^T\} = 0,
\]

where \( T \) denotes the transpose of a matrix.

In most of the channel equalization systems, the source is considered to be symmetric where \( E\{g\} = E\{g^3\} = 0 \) assuming zero-mean, but in image processing context, the source cannot be based on such assumptions as images generally do not follow a well defined structure. Hence, it clearly shows that the stationary points of the CMA cost function are affected by the intensity distribution of the ground truth image, specially from first order moment to fourth order moment.

In order to see the effect of pixel correlation on CMA convergence, we deconvolved a natural blurred image and a randomized blurred image based on the same ground truth image and the randomization is done in such a way that it preserves the histogram of the ground truth image. As shown in Fig. 5.4, it is clear that, while an uncorrelated source results in PSR, a correlated source such as a natural image will not result in PSR under classical CMA adaptation.

### 5.4.2 Natural Image Statistics

It is intuitively clear that neighborhood pixel values tend to be very similar in natural images resulting in inherent spatial correlation [123]. The recent advances in research on natural image statistics encouraged us to explore on alternative ways to overcome the spatial correlation effect on CMA deconvolution.

Recently, natural image statistics has been an area of much research and atten-
Figure 5.4: Classical CMA deconvolution of a natural image and a randomized natural image
tion. This appeal of study on natural image statistics follows from the vast range of applications they are involved in, starting from our own visual system [128] to plethora of image processing systems such as compression algorithms [51, 129]. The principal goal of visual coding and image compression is to reduce the redundancies in the visual presentation. It is shown in [51] and [130], when the statistics of a set of data are stationary, which is the case in natural images, all the redundancies are reflected in terms of correlations between the pixels and are captured by the amplitude spectra of the data. Furthermore, several studies [129–131] have shown that, generally, the spectral power of natural images falls with frequency according to power law, $1/f^p$, having $p$ closer to 2 for natural images. We borrow the results of these studies to mitigate the effect of spatial correlations in image restoration through CMA.

5.4.3 Whitening and CMA

Whitening is the process of transforming an original image to a set of variables, where the transformed variables are uncorrelated. Out of number of whitening algorithms proposed, we concentrate on one of the three special linear transforms given in [132], namely, principal component analysis (PCA) filter, over the zero-phase component analysis (ZCA) and independent component analysis (ICA) filters. The motivation behind our choice relates to the sparsity of the transformation, where PCA results in the least sparse distribution [132], which will ensure better CMA convergence over ZCA and ICA, as detailed in Section 5.3. The PCA filter, we use is based on the Fourier filters, ordered according to the amplitude spectrum of the image in justifying the aforementioned natural image statistics.

The CMA image restoration system with whitening as a pre-process is shown in Fig. 5.5. In order to isolate the effects of pixel correlation on CMA convergence, we apply histogram equalization to the whitened source, making the source approximately uniform distributed. For comparisons, we use the systems shown in Fig. 5.3 and Fig. 5.5, where the blurred images of natural image and the randomized image based on the natural image are deconvolved through the system in Fig. 5.3, while the blurred image of whitened image based on the same natural image is deconvolved with the system in Fig. 5.5. The difference $d$ against the pixel
estimations is plotted in Fig. 5.6. As portrayed by the plot, while the randomized natural image deconvolution achieves PSR, through whitening, we could achieve satisfactory PSR in CMA deconvolution.

5.4.4 Whitening in Practice

Inspired by the results in Fig. 5.6, we extended our analysis to test the potential of whitening as a pre-process in realistic CMA image restoration. As the first step in this analysis, the implications of histogram equalization on whitening was tested by deconvolving the blurred images of natural image, randomized natural image and whitened natural image through the system shown in Fig. 5.1. An unsatisfactory behavior of the whitened natural image can be observed through the results shown in Fig. 5.7.

The reasons behind this sensitivity of CMA image deconvolution to histogram equalization on a whitened source can be simply explained by the shape of the whitened and non-whitened source distribution. The log histogram of the original and whitened images are shown in Fig. 5.8, where the departure of the whitened source from the uniform distribution to a sparse distribution is clearly visible. These results further stress the significance of shape of the source distribution on CMA convergence as already discussed in Section 5.3.

Hence, we claim that even though natural image statistics can be successfully used for spatial de-correlation in images, the greater degree of sparsity of the whitened image distribution restricts CMA performance in realistic image deconvolution.
Figure 5.6: CMA deconvolution results on spatial correlation tests.
Figure 5.7: CMA deconvolution results on whitening and histogram equalization.
Figure 5.8: Log histograms of natural ‘Lena’ image ($C_x = 2.1759$) and whitened ‘Lena’ image ($C_x = 11.4199$).
5.5 Deconvolution of Natural Images using Segmentation Based CMA

The findings in Section 5.4 encouraged us to develop a novel model to overcome the limitations of spatial correlation in CMA deconvolution which is detailed in this section.

5.5.1 Pixel Correlations and Higher Order Moments

Although Kurtosis has been used extensively in sparse coding of natural images [51], no attempt has been made it to be used in relation with image restoration. By looking at kurtosis in a different perspective, as a measurement of pixel correlatedness, we develop a novel model to address the issue of spatial correlation in CMA deconvolution. For this, we use the model shown in Fig. 5.9, where the ground truth image $g$ is modeled as a collection of image segments $x_1, \ldots, x_r$ which are smoothened by the filters $f_1, \ldots, f_r$. Hence, the unknown ground truth image in the image deconvolution problem can be considered as a collection of regions with varying levels of smoothness.

The initial source of image segments $x_1, \ldots, x_r$, which are considered to be having zero or less correlated pixels, will have low kurtosis values. Once these images are filtered through the bank of smoothing filters and a blurring kernel,
the blurred image $b$ will be having a range of regions with varying kurtosis values, where the least correlated segments will have the lowest kurtosis values [133]. As $w$ is determined through CMA, in order to lower the effect of correlation on CMA convergence, we perform the CMA adaptation only on the minimum kurtosis regions of the blurred image. Once the deconvolution kernel $w$ is determined through minimum kurtosis regions, the whole blurred image $b$ is convolved with $w$ to get the estimated ground truth image $z$.

### 5.5.2 CMA Adaptation Based on Kurtosis Minimization

In classical CMA image deconvolution under direct parametrization, the regressor vector represents the pixels in the blurred image. With our segmentation based CMA deconvolution model, the regressor vector stands for the pixels of the segmented blurred image, in which, the segmentation is based on the minimum kurtosis.

Let $\mathcal{C}(r_1, r_2)$ be the kurtosis for each pixel in the blurred image $b$, where $\mathcal{C}(r_1, r_2)$ is estimated for a $\ell \times \ell$ neighborhood window, centering the pixel $b(r_1, r_2)$, where $(r_1, r_2)$ are selected in such a way, avoiding spurious kurtosis values near the borders of the image. In order to segment the blurred image based on the minimum kurtosis, $\mathcal{C}(r_1, r_2)$ is calculated for all the pixels in the blurred image, then the minimum kurtosis for the image is defined as

$$
\mathcal{C}_{\text{min}} = \min(\mathcal{C}(r_1, r_2)) \quad \forall \{r_1, r_2\} \in \mathcal{I},
$$

where $\mathcal{I}$ is the set of image pixels.

Under CMA adaptation on minimum kurtosis segments, the update equation (5.6) changes to

$$
\begin{equation}
\begin{aligned}
w(m_1, m_2)_{n+1} &= w(m_1, m_2)_n - \alpha \phi(z(l_1, l_2)_n) s(l_1 - m_1, l_2 - m_2),
\end{aligned}
\end{equation}
$$

where

$$
\begin{equation}
\begin{aligned}
s(r_1, r_2) &\triangleq \{b(r_1, r_2) | \mathcal{C}(r_1, r_2) < (\mathcal{C}_{\text{min}} + \mathcal{T})\},
\end{aligned}
\end{equation}
$$

and the threshold, denoted by $\mathcal{T} \in \mathbb{R}$ is chosen based on the values of $\mathcal{C}$ for
optimum results.

### 5.5.3 Simulations on Correlation Model

Based on the model shown in Fig. 5.9, first, we simulated the CMA adaptation on a randomized source. For simplicity, we use \( r = 1 \), making half of the image to be filtered through the filter bank. We take \( l \) to be 64 and \((r_1, r_2)\) start from \((32, 32)\) to avoid boundary problems. The blur kernel filtering is then applied to the image globally and classical CMA adaptation and CMA adaptation on minimum kurtosis were carried out in deconvolution. The results shown in Fig. 5.10 prove that, with CMA adaptation based on the minimum kurtosis, we could estimate the deblur kernel accurately, which result in PSR.

In reality, a natural image is far from a randomized image. In [51], it is shown that while randomized white noise results in a uniform probability density through
its “state space”, a natural image does not fill the “state space” uniformly. Motivated by the above research in natural images, we extended our experiments to take an image with less redundancy as the initial source for the model shown in Fig. 5.9, where the redundancy reduction is performed through whitening. The CMA adaptation on minimum kurtosis is carried out on a whitened and a randomized image based on the same ground truth image for the model in Fig. 5.9. The results are shown in Fig. 5.11, with the classical CMA adaptation results for comparison.

As evidenced by the results, CMA adaptation on minimum kurtosis sections results in better deblur kernel estimation than the classical CMA adaptation. In addition, the results also verify that images with less redundancies could achieve closer to PSR in minimum kurtosis CMA deconvolution. This is a very important result as one finds many types of images with less redundancies in natural images and in other fields, for which, the minimum kurtosis deconvolution could be applied effectively.
5.5.4 Natural Image Deconvolution on Minimum Kurtosis Segmentation

Taking the minimum kurtosis image segmentation model detailed in Section 5.5.2 as the basis, various simulations were performed for a large number of natural images. The image database we use has a variety of images having kurtosis values in the range of 1.1921 to 11.2778. One of the deconvolution results from our simulations are shown in Fig. 5.13\(^2\). As seen in Fig. 5.13(f) and experienced by other simulations, we verify that minimum kurtosis CMA adaptation converges faster than classical CMA adaptation. But in some cases, we experienced lack of robustness of minimum kurtosis CMA deconvolution on deconvolving natural images.

Upon investigating the reasons for such behavior, we found several issues underlying the deconvolution of natural images based on fourth order moments, which are detailed below.

1. Limited data - As already detailed in Chapter 3, the limited data in natural images affects our modified CMA deconvolution algorithm. As the segmented CMA deconvolution is based on the fourth order moment, there need to be enough data to get a reliable measurement of the fourth order moment.

The experiments we carried out show that when the \( \ell \times \ell \) neighborhood window used for kurtosis calculation is larger (i.e., \( I \) is larger), the results are better. The results of minimum kurtosis CMA deconvolution for a 515 \( \times \) 515 image with 32 \( \times \) 32 and 64 \( \times \) 64 neighborhood windows are shown in Fig. 5.12. When deconvolving with the 64 \( \times \) 64 window, we used the scale invariant property of image statistics [130, 134] and image was scaled twice upwards as to come up with a reasonable number of estimations. These results verify that when the neighborhood window is larger, deconvolution results are better and the limitation caused by finite images can be overcome by upward scaling of the known image.

2. Meso-kurtic input - As already discussed in Section 5.3, Meso-kurtic input

\(^2\)The image for the simulation is downloaded from http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds.
Figure 5.12: Difference between the actual and estimated deblur kernels for $515 \times 515$ image deconvolved via classical CMA adaptation and minimum kurtosis adaptation based on the model in Fig. 5.9
cause ill-behavior in CMA image deconvolution. Under our pixel correlation model shown in Fig. 5.9, it can be seen that the less correlated source images after filtered through the bank of smoothing filters, will have monotonically increasing kurtosis values. Based on the central limit theorem, the smoothing will asymptotically make the blurred image to reach a Meso-kurtic distribution. When this happens, we experienced that we could not achieve the expected results in our simulations.

3. Outlier sensitivity - It is shown in [51, 123] that the kurtosis measurement might be outlier-prone for some digital images. For example, this is experienced when the digital image consists of few bright spots and those outliers are not properly calibrated through gamma correction.

4. Blur kernel effect - In communication systems, linear equalizers are proved to be very inefficient as well as ineffective when they are applied to channels with zeros on the unit circle or to Single-Input-Multiple-Output (SIMO) channels which have common zeros in their $z$ transform of the impulse response. [125, 135, 136]. This applies to CMA equalization as well, resulting that CMA gives perfect blind equalization results (in the absence of noise), provided that the $z$-transform of the channel impulse response has no zeros on the unit circle [137]. We experienced similar situation in our experiments, when the blurring function has zeros on the unit circle.

In conclusion, even though the modified minimum kurtosis based algorithm is not robust in catering the perfect source recovery of all the different highly correlated natural images, it is better than classical CMA deconvolution for less correlated images in terms of efficiency and effectiveness.

5.6 Zero-mean Ground Truth Image

As already mentioned in Section 5.1, it is shown in [11], that a gradient descent of the CM criterion, such as CMA will achieve perfect equalization when the source meets the conditions of zero-mean, white, and sub-Gaussian. In Section 5.3 and Section 5.4, we have discussed the deviations of CMA from the ideal conditions,
Figure 5.13: Comparison between the classical CMA adaptation and our novel segmentation based CMA adaptation.
when the source is non-platykurtic and non-white in CMA image deconvolution. Thus, in this section, we discuss the tolerability of zero-mean condition of the source on CMA image deconvolution.

We build a mathematical relationship between the blurred image created from a zero-mean GTI denoted as $b_{gz}$ and the zero-mean blur image denoted as $b_{bz}$ as follows:

$$
b_{gz} = k \otimes (g - \mu_g)$$
$$= k \otimes g - \mu_g \otimes k$$
$$= b - \frac{1}{L_1 \times L_2} \sum_{l_1=0}^{L_1-1} \sum_{l_2=0}^{L_2-1} g(l_1, l_2) \otimes k$$
$$= b - \frac{1}{L_1 \times L_2} \sum_{l_1=0}^{L_1-1} \sum_{l_2=0}^{L_2-1} b(l_1, l_2)$$
$$= b - \mu_b = b_{bz},$$

where $\otimes$ represent the convolution operator, $L_1$ and $L_2$ are the maximum support of the GTI $g$ and the blurred image $b$, while $\mu_g$ and $\mu_b$ stand for the mean of GTI $g$ and blur image $b$ respectively. As (5.6) depends only on the pixel values of the blurred image, with the above analysis, it is clear that the same result of a zero-mean source could be achieved through making the blur image zero-mean.

This mathematical analysis is corroborated by the result shown in Fig. 5.14. While the convergence plot of the non zero-mean source justifies the necessity of a zero-mean source for global convergence of the CMA, this limitation is overcome by having the blur image zero-mean, which has the same CMA convergence results as a zero-mean GTI.

## 5.7 Contributions

Though the global convergence analysis of CMA has been exhaustive in the data communication context, it is still in its infancy in image processing context. This prompted us to examine the underlying assumptions on the source on CMA image restoration and we come up with the following contributions.
Figure 5.14: CMA deconvolution results on zero-mean tests.
1. In Section 5.3, we analyze the effect of source distribution on image restoration through CMA. We show that the statistical distribution of the ground truth image (GTI) affects the results of CMA image restoration in practice, except for those images, which have their distribution closer to uniform distribution.

2. The analysis of effect of spatial correlation on source in Section 5.4 show that spatial correlation is the most stringent factor for the undesirable convergence in blind CMA image restoration. As discussed in Section 5.4.4, even though there are applicable methods such Principle Component analysis (PCA) to de-correlate natural images, due to the relationship between de-correlation and sparse distribution, CMA cannot directly use such de-correlation techniques in practise.

3. Thus, to overcome the negative implications of spatial correlation in natural images in blind CMA deconvolution, in Section 5.5, we develop a modified CMA algorithm based on the fourth order moment of the blurred image to overcome the correlation effect. Although the modified algorithm is not robust in catering the perfect source recovery of all the different highly correlated natural images, it is better than classical CMA deconvolution for less correlated images in terms of efficiency and effectiveness.

4. Out of the assumptions of the source that we applied on CMA image restoration, we prove in Section 5.6 that only the assumption that source being zero-mean is valid and does not carry any weight on the ill-convergence of blind image restoration through CMA.

## 5.8 Conclusions

In this chapter we examine validity of key underlying assumptions on the source, namely, the source being zero-mean, independent and uniformly distributed for perfect source recovery in CMA blind image deconvolution. By theoretical expositions and experiments, we claim that while spatial correlation is the most stringent factor for the undesirable convergence in blind CMA image restoration, statistical
distribution of the source also restricts the performance of blind image restoration through CMA.
Chapter 6

Multichannel Image Deconvolution with a Single Image

6.1 Overview

The signal processing community identifies a multichannel signal as a signal which can be separated out to a group of signals which relate in cross-channel similarity or correlation. Deconvolution is one of the fields that exploit these relationships in order to develop higher performing algorithms. As there are significant differences in the processing of multichannel signals from processing of single channel systems, the extension of single channel deconvolution algorithms to the multichannel deconvolution context is non trivial and would not deliver the optimal performance. Hence, the need for explicit development of multichannel deconvolution algorithms has been lately addressed in the context of data communications, producing highly efficient one dimensional deconvolution algorithms [46, 47, 138]. Extending this research into two-dimensional signals, effective image restoration has been performed when there are multiple images for the same scene [13, 14].

Based on the acquisition of images, multichannel restoration (MCR) is classified as single-input multiple-output (SIMO) restoration model and multiple-input multiple-output (MIMO) restoration model, where the former relates to channels
with frequency bands, resolution levels or color channels, while the latter represent different frames under varying environment conditions. We confine ourselves to the SIMO model exclusively and any reference to the term MCR hereafter denotes the SIMO model.

Though there are specific applications such as astronomical [78], microscopic, broadband imaging and several other types of imagery which requires/generates multiple copies of the same scene, in consumer photography, it is impractical to find/create multiple copies of the same scene, unless, using a special optical device or there is a special requirement to do so. Hence, as multichannel imagery is not inherent in most of the general imagery, the use of the efficient MCR techniques is restricted in practical image restoration. This stimulated us to research into the applicability of MCR techniques for a single image.

In this chapter, we contribute in:

1. developing a novel model for the restoration of a single blurred image through multichannel restoration techniques.

2. demonstrating the use of our novel model in simultaneous processing of deconvolution and demosaicing in the processing of color images.

3. using finite impulse response (FIR) deconvolution filters as opposed to infinite impulse response (IIR) deconvolution filters, which are generally required in single image deconvolution.

### 6.2 Equalization

As already explained in Chapter 4, equalization is the technique used in communication systems to restore the original signal, which is distorted during the channel transmission. In other words, when the channel is modeled as a linear system, the equalizer is the system, which inverts the response of the channel. In Chapter 4, we have shown that blind equalization is generally carried out by adapting the equalizer parameters according to an objective function. When the desired equalizer output is not available, this objective function is built based on some known characteristic or statistical property of the source. Most of the successful single
channel equalization techniques such as constant modulus algorithms (CMA) are based on the higher order moments of the source [6–8], as second order statistics is not sufficient to recover the signal successfully [47]. In contrast to these algorithms, it was shown by Slock [46], that under the multichannel model, perfect equalization could be achieved by using only second order statistics, where the multichannel model is built by over sampling and equalized through fractionally spaced equalizers.

Thus, this section is devoted to provide a theoretical exposition of main aspects in multichannel system processing. To make the theories and concepts behind equalization simpler, in the following sub sections we have assumed noiseless channel-equalizer systems.

### 6.2.1 Fractionally Spaced Linear Equalization

Fractionally spaced equalizer (FSE) is commonly used in communications context, where the tap spacing of the equalizer is a fraction of the baud spacing (in time) or the transmitted symbol period. The continuous-time model of a simple noiseless version of the communications model is shown in Fig. 6.1, where the index $n$ represent the baud spacing and index $k$ represent the fractionally spaced signals [11].

The source signal $s_n$ is transmitted through a channel and the received signal $r(t)$ can be represented as [139]

$$ r(t) = \sum_{n=-\infty}^{\infty} s_n h(t - nT), $$

where $h$ represent the channel impulse response and $T$ is the baud space interval. In order to get the same rate of output from the equalizer as the input rate, FSE,

![Figure 6.1: Single channel system](image-url)
$e_k$ is tap spaced by $T/L$ where $L$ is the over sampling factor. Finally, the FSE output $x_k$ is decimated to create the $T$-spaced output sequence $y_n$.

Consider for example, a system with $L = 2$ where we get a $T/2$ spaced $r$.

$$r(k \frac{T}{2}) = \sum_{n=-\infty}^{\infty} s_n h(k \frac{T}{2} - nT).$$

For $k$ even (i.e., $k = 2m$, $i = 0, 1, 2, ...$) and odd (i.e., $k = 2m - 1$), we can define $T$ spaced subsequences of $r$.

\[
\begin{align*}
\quad r_{m}^{even} &= r(mT) \\
&= \sum_{n=-\infty}^{\infty} s_n h((m - n)T) \\
\quad r_{m}^{odd} &= r(mT + \frac{T}{2}) \\
&= \sum_{n=-\infty}^{\infty} s_n h((m - n)T - \frac{T}{2})
\end{align*}
\] (6.1)

For a length $2N$ finite impulse response (FIR) FSE with tap spacing $T/2$,

$$x_k = \sum_{i=0}^{2N-1} e_i r((k - i) \frac{T}{2}).$$

(6.2)

Once decimated, the equalized output becomes [11],

\[
\begin{align*}
\quad y_n^{odd} = x_{2n+1} \\
&= \sum_{i=0}^{2N-1} e_i r(nT + \frac{T}{2} - \frac{T}{2}) \\
&= \sum_{i=0}^{2N-1} e_i r(nT - i \frac{T}{2} + \frac{T}{2}) \\
&= \sum_{i=0}^{N-1} (e_{2i} r(nT - iT + \frac{T}{2}) + e_{2i+1} r(nT - iT))
\end{align*}
\]
Figure 6.2: Two channel system

\[
\begin{align*}
    s_n &\rightarrow [h_{n}^{\text{even}} r_{n}^{\text{even}} e_{n}^{\text{odd}}] + y_n \\
    [h_{n}^{\text{odd}} r_{n}^{\text{odd}} e_{n}^{\text{even}}] &\rightarrow
\end{align*}
\]

where \( h_{n}^{\text{even}} = h(nT)\) and \( h_{n}^{\text{odd}} = h(nT + \frac{T}{2})\). Combining all together, and considering that in decimation, we keep only the odd samples, makes \( y_n = y_n^{\text{odd}} \), and

\[
\begin{align*}
    y_n &= \sum_{i=0}^{N-1} e_{i}^{\text{even}} (\sum_{l} s_{l} h_{n-i-l}^{\text{odd}}) + \sum_{i=0}^{N-1} e_{i}^{\text{odd}} (\sum_{l} s_{l} h_{n-i-l}^{\text{even}}) \\
    &= s_n \otimes (e^{\text{even}} \otimes h^{\text{odd}} + e^{\text{odd}} \otimes h^{\text{even}}),
\end{align*}
\]

which results in multichannel model as shown in Fig. 6.2. Generalizing the upsampling factor to \( L \) results in Fig. 6.3.

### 6.2.2 Multirate System

As shown in the above section, FSE involves changing the sampling rate of the system. If we are to change the sample rate of the system by converting the
Figure 6.3: Multichannel system with upsampling factor $L$
6.2 Equalization

digital signal to a continuous signal and then back to digital with the changed sampling rate, it may introduce many errors in the processing through non-ideal filters and analog to digital converters and vice versa. Thus, in this section, we build the above multichannel system in discrete domain through interpolation and decimation.

Consider the interpolation of a signal $p$ by a factor of $L$. The output signal $y$ can be represented as

$$y(n) = p\left(\frac{n}{L}\right), \quad \forall n \in w$$

$$= 0, \quad otherwise$$

where $w$ is the set of $n$ which are multiples of $L$. The $z$ transform of the interpolator output $y(n)$ is given by [140]

$$Y(z) = P(z^L).$$

In general, interpolation is followed by an interpolation filter to eliminate the imaging effect created during interpolation [141]. In the case of decimation, the decimator is preceded by a bandlimiting filter, whose purpose is to avoid aliasing. As implementing these changes in straightforward filtering requires a large amount of computing, multirate signal processing techniques are used in improving the efficiency of the FSE systems.

In general, multirate signal processing techniques refer to the use of upsampling, downsampling, expanders and compressors in a variety of ways to increase the efficiency of the system. As a basic result of multirate signal processing, the operation of linear filtering and up-sampling or down-sampling can be interchanged if the linear filter is modified as shown in Fig. 6.4 and Fig. 6.5 [142, pp 179-180].

The multirate operations of decimation and interpolation shown in Fig. 6.4 and Fig. 6.5 will now be used to decompose the system function $H(z)$ into its polyphase representation. The polyphase decomposition of a sequence is obtained through the representation of the sequence as a superposition of $L$ subsequences. The advantage of polyphase decomposition of multirate signal processing systems lies in the efficient implementation structures for linear filters [142, pp 180-184].
Let \( H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n} \) be the transfer function of a digital filter. We can write \( H(z) \) in the \( L \)-component polyphase form [141] as

\[
H(z) = \sum_{k=0}^{L-1} z^{-k}C_k(z^L),
\]

where \( C_k(z) \) are called polyphase components and the coefficients of the filter \( C \) are given by

\[
c_k(n) = h(nL + k) \quad 0 \leq k \leq L - 1.
\]

Applying this polyphase decomposition into the interpolation system in Fig. 6.4, we can come up with an efficient implementation of the interpolation system as shown in Fig. 6.6.

Consider the single channel system with an upsampling factor \( L = 2 \), for which, subsequences of \( r \) are given in (6.1). As these subsequences are the result of polyphase decomposition on the \( T/2 \)-spaced channel impulse response [141], the equalizer system can also be modeled as a multirate system as given in Fig. 6.7.
6.2 Equalization

\[ C_0(z) \uparrow L + \cdots + C_{(L-1)}(z) \uparrow L \]

Figure 6.6: Polyphase decomposition of an interpolation system

\[ p_n \uparrow 2 a_n h_k r_k e_k x_k \downarrow 2 y_n \]

Figure 6.7: Multirate system for \( L = 2 \)
Applying similar concepts as in Fig. 6.6, this multirate system can also be drawn as a polyphase decomposition of the filters as shown in Fig. 6.8. The filters $h$ and $e$ are considered as $T/2$ spaced, making $h_{\text{even}}$, $h_{\text{odd}}$, $e_{\text{even}}$ and $e_{\text{odd}}$ to be $T$ spaced vectors.

Further extending Fig. 6.7, the single channel system in Fig. 6.1 can be represented by an equivalent discrete-time multirate system model given in Fig. 6.9. The coefficients of each of the multichannel filters are a decimated version of the filters used in the multirate system. For notational simplicity, from this point onwards, the indexes $n$ for baud-spaced and $k$ for fractionally spaced are been dropped.

### 6.2.3 Zero-Forcing FSE

In communication context, the desired system response from transmitter to receiver in most applications is a pure delay, $z^{-\delta}$ for some integer $\delta > 0$. If the combined channel-equalizer response is denoted by $c$, then $c_{\delta}$ represents a pure delay. Such a pure delay is not achievable in the baud-space equalization system [139]. In contrast to the baud-space equalization, FSE can achieve a pure delay and for the system shown in Fig. 6.2, the combined channel-equalizer response in $z$ domain
can be represented as
\[ H^{\text{even}}(z^{-1})E^{\text{odd}}(z^{-1}) + H^{\text{even}}(z^{-1})E^{\text{odd}}(z^{-1}) = z^{-\delta}, \] (6.3)

where \( H \) and \( E \) stand for the z-domain representation of \( h \) and \( e \) respectively. This is called zero-forcing equalization, since the combined channel-equalizer response is forced to zero for all delays except \( \delta \). The generalization of this representation leads to the Bezout relationship [143] given by
\[ H_0(z^{-1})E_0(z^{-1}) + H_1(z^{-1})E_1(z^{-1}) + \ldots + H_{L-1}(z^{-1})E_{L-1}(z^{-1}) = z^{-\delta}. \] (6.4)

It is shown in [11, 139], that (6.4) can be written in Toeplitz form having a compound matrix \( H \) for the blur filters \( h_0 \cdots h_{L-1} \) and a compound vector \( e \) for the equalizers \( e_0 \cdots e_{L-1} \). For length \( M \) blur kernels,
\[ c_\delta = He, \] (6.5)
\[ H = [[H]_0 \ [H]_1 \ \cdots \ [H]_{L-1}], \]
\[ e = \begin{bmatrix} e_0 \\ e_1 \\ \vdots \\ e_{L-1} \end{bmatrix}, \]
\[ [H]_j = \begin{bmatrix} h_j(0) & h_j(1) & \cdots & h_j(M-1) \\ h_j(1) & h_j(0) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h_j(M-1) & \cdots & h_j(1) & h_j(0) \end{bmatrix}, \]

where \( c_\delta \) is the impulse response having a nonzero co-efficient in the \( \delta \) position.
Perfect Source Recovery Conditions

Perfect source recovery (PSR) or perfect source equalization is, when the output of a system \( y \) is related to the input of the system \( s \) with \( y_n = s_{n-\delta} \) for some fixed delay \( \delta \). In FSE, this is achieved with several conditions as given below [11, 139].

1. Noise - There is no additive noise

2. Equalizer length - It is proved in [135, 144, 145] that for perfect equalization, \( H \) should be full row rank resulting in the equalizer length requirement as follows.

\[
N \geq \frac{M - 1}{L - 1},
\]

where \( N \) is the length of the equalizer, \( L \) is the sampling factor and \( M \) is the length of the channel.

3. Channel Disparity - In satisfying the full row rank condition on (6.5), it is shown in [11, 135, 146], that there should not be a root (factor) common to all the sub-channel polynomials.

6.3 Multichannel Image Restoration

Multichannel image restoration has been an active field of research in the last two decades. It has been applied in the fields of astronomical [78, 147], microscopic [148], satellite and several other types of imagery including consumer photography. The benefit gained through multichannel processing to single channel processing is mainly due to exploiting the redundancy and diversity of information available through various acquisitions. In considering the stochastic image restoration algorithms discussed in Chapter 2, the major drawbacks of single channel restoration such as dependency on prior knowledge of the image-blur model [12], strong assumptions on image-noise model [4], less efficiency and convergence issues due to iterative processing [149] and others could be made minimal or completely avoided by the use of multichannel processing. Not only in stochastic image restoration, but inverse filtering also gains advantage on multi channel filtering, as
in single channel deconvolution, if the blur filter is FIR filter, the deconvolution filter becomes infinite impulse response filter, and this could be completely overcome by the use of multi channels. As discussed in Chapter 2, MCR can be addressed either as direct restoration or with the two step method of channel identification and non-blind deconvolution. The multiple channels in MCR can be categorized in various ways, in frequency, in color, in resolution and in time.

6.3.1 Notations and Preliminaries

In absence of noise, MCR can be represented as in Fig. 6.10, where \( g \) stands for the ground truth image (GTI), \( L \) for the number of channels, \( k_0 \cdots k_{L-1} \) for the blur kernels, \( b_0 \cdots b_{L-1} \) for the blurred images, \( e_0 \cdots e_{L-1} \) for the equalizers and \( \hat{g} \), for the estimated GTI. For a GTI \( g \) of size \([L_1, L_2]\), the blur process in a single channel with a blur kernel \( k \) of size \([N_1, N_2]\) can be represented as

\[
b(l_1, l_2) = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} k(n_1, n_2)g(l_1 - n_1, l_2 - n_2),
\] (6.7)
where \((l_1, l_2) \in [0, L_1 - 1] \times [0, L_2 - 1]\). In MCR, for a given set of blurred images \(b_0 \cdots b_{L-1}\), we try to estimate the set of equalizers \(e_0 \cdots e_{L-1}\), which in turn will be used to estimate the GTI.

We extend (6.4) as a multiplication of block-Toeplitz-block (BTB) matrix of \(k\) and lexicographic ordering of \(e\) [26] to form the Bezout identity in two dimensions. Consider a single channel in the multichannel framework where the blur kernel \(k^i\) of \(i\)th channel is of size \(N_1 \times N_2\). We are interested in finding an equalizer \(e^i\) of order \(M_1 \times M_2\). The BTB matrix for \(k^i\) supporting \(e^i\) denoted as \(H_i\), can be defined as [26, pp. 216-218]

\[
[H]_i \triangleq \begin{bmatrix}
[K_0^i] & [K_1^i] & \cdots & [K_{N_1-1}^i] \\
[K_N^i] & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
[K_{N_1-1}^i] & [K_{N_1}^i] & \cdots & [K_0^i]
\end{bmatrix},
\]

(6.8)

where the matrices \([K_j^i]\) are each of size \((N_2 + M_2 - 1) \times M_2\) and each is formed from the \(j^{th}\) row of the PSF matrix \(k^i\) according to

\[
[K_j^i] = \begin{bmatrix}
k^i(j, 0) & k^i(j, 0) & \cdots & k^i(j, 0) \\
k^i(j, 1) & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
k^i(j, N_2 - 1) & \cdots & \cdots & k^i(j, 1)
\end{bmatrix}.
\]

Similar to (6.4) we write the combined blur-deblur response as

\[
c_\delta = He, \quad (6.9)
\]

\[
H = [[H]_0 \quad [H]_1 \cdots [H]_{L-1}],
\]
\[ e = [e^0 \ e^1 \ \ldots \ e^{L-1}]^T, \]
\[ c = [c^0 \ c^1 \ \ldots \ c^{L-1}]^T, \]

where \( T \) stands for the matrix transpose and \( c^i \) and \( e^i \) are formulated by the rows of the lexicographic orderings, stacked as columns, which are denoted as

\[ c^i_j = \begin{bmatrix}
  c^i(j,0) \\
  c^i(j,1) \\
  \vdots \\
  c^i(j,N_2 + M_2 - 1)
\end{bmatrix};
\]
\[ e^i_j = \begin{bmatrix}
  e^i(j,0) \\
  e^i(j,1) \\
  \vdots \\
  e^i(j,M_2 - 1)
\end{bmatrix}, \]

where \( i \) represents the channel and \( j \) represents the rows of the channel matrix. This system of equations can be used for estimating the equalizers for each channel in MCR.

### 6.3.2 Perfect Source Recovery Conditions

It is shown in [13], that the existence and uniqueness of solutions to (6.9) depends on the following factors:

1. Noise - There is no additive noise

2. Number of channels - Under the strong co-primeness requirement, \( L \geq 3 \), implying that at least three different channels are needed for perfect restoration.

3. Equalizer size - When \( H \) is full row rank, which is a sufficient condition for PSR, the equalizer size \([M_1, M_2]\) is conditioned with

\[
[M_1, M_2] \geq [N_1 - 1, \left\lfloor \frac{2N_2}{L-2} \right\rfloor - 1] \quad \text{or} \quad [M_1, M_2] \geq \left\lfloor \frac{2N_1}{L-2} \right\rfloor - 1, N_2 - 1, \]

where \( \lfloor x \rfloor \) denotes the largest integer less than or equal to \( x \).
6.4 Novel Image Restoration Model

As discussed in Chapter 2, most of the MCR techniques developed in the past were based on the multiple copies of the same scene, acquired through a specific application such as astronomy or microscopic or using a special optical device. As these options are not always available in applications such as consumer photography, the application of MCR techniques have been limited in the image restoration literature. In addressing this limitation, we introduce a novel image restoration model, which could be utilized to achieve the benefits of multichannel processing but through a single image. We believe that our novel multichannel image restoration model has the potential of overcoming most of the drawbacks in single image restoration, while bringing in the advantages of multichannel processing.

Our model of multichannel image restoration is formed by several assumptions, which are given below:

1. Finite support - We assume that the image to be processed is of finite support, where the area outside the image is unknown. The edges demarcate the image and the unknown area.

2. Blur kernel - The blur kernels are taken as FIR filters.

3. Noise - The initial model was developed assuming a noiseless distortion.

In the absence of noise, we model the conventional blur-deblur process in a different way, which is shown in Fig. 6.11. The conventional blur model given in (6.7) is shown in the upper part of Fig. 6.11, while the proposed deblur model is shown in the lower part. As discussed in Chapter 2 and Chapter 3, the classical single image restoration techniques could be applied as $\mathcal{D}$ in the conventional deconvolution process to obtain the estimated GTI $\hat{g}$. With our novel image restoration model, we consider the GTI, $g$, as a result of a transformation process of a pre-ground truth image (pre-GTI), $p$.

Although there are numerous transformation processes that could be applied to get the GTI, $g$, in our restoration model we consider the transformation to be an upsampling process. Let $p$ be a pre-ground truth image which has essentially the same information as $g$ but in a more compact form. The transformation process $\mathcal{T}$
transforms \( p \) to the GTI \( g \). Specifically, \( \mathcal{T} \) can be considered as the ground truth model, where the actual ground truth image is formed by interpolating \( p \), followed by an interpolating filter as shown in Fig. 6.12, where \( L \) is the up-sampling factor and \( f \) is the interpolation filter.

Taking the new blur-deblur model together with the ground truth image model shown in Fig. 6.11 and Fig. 6.12 into account, our objective is to design an effective and efficient deblurring model \( \mathcal{D}_m \), which allows us to estimate the pre-ground truth image \( \hat{p} \), from which, we can estimate the GTI \( \hat{g} \).
6.4.1 Image Restoration using Bezout Equalizers

When the ground truth model in Fig. 6.12 is introduced to the conventional blurring process defined in (6.7), we can define the blur image as

\[ b(l_1, l_2) = \sum_{r_1=0}^{R_1-1} \sum_{r_2=0}^{R_2-1} w(r_1, r_2)p'(l_1 - r_1, l_2 - r_2), \quad (6.11) \]

where \( p' \) is the up-sampled image of \( p \) and \( w \) is a blur filter of order \((R_1, R_2)\) defined as the convolution

\[ w(r_1, r_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} k(n_1, n_2)f(r_1 - n_1, r_2 - n_2). \quad (6.12) \]

We extend the concepts presented in Section 6.2 and Section 6.3, to our new image restoration model shown in Fig. 6.13. The pre-image \( p \) is upsampled by a factor of \( L \) making \( p' \). The upsampled image is then filtered by an interpolation kernel \( f \) resulting in the GTI \( g \). The blur image \( b \) is created by convolving \( g \) with the blur filter \( k \).

In the deconvolution process, the blurred image is separated into \( L \) sub images in order to create the multichannel system for image restoration. In order to find the set of equalizers with the concept shown in (6.9), we downsample \( w \) to \( L \) sub filters. Let \( w^i \) be the sub filter for the \( i \)th channel. The corresponding equalizer \( e^i \) could be found by solving a similar set of equations as in (6.9) with

\[
H_i \overset{\Delta}{=} \begin{bmatrix}
[W^i_0] & [W^i_1] & \cdots \\
[W^i_1] & [W^i_0] & \cdots \\
\vdots & [W^i_1] & \ddots \\
[W^i_{R_1-1}] & \cdots & [W^i_0] \\
\end{bmatrix},
\]

where the matrices \([W^i_j]\) is formed from the \( j^{th} \) row of the filter matrix \( w^i \) according
Figure 6.13: Multichannel image restoration with a single image
to

$$[W^i_j] = \begin{bmatrix}
    w^i(j, 0) & w^i(j, 0) & \ldots & w^i(j, 0) \\
    w^i(j, 1) & w^i(j, 1) & \ldots & w^i(j, 1) \\
    \vdots & \vdots & \ddots & \vdots \\
    w^i(j, R_2 - 1) & \ldots & w^i(j, 0) & w^i(j, R_2 - 1)
\end{bmatrix}.$$

The set of downsampled blur images are deconvolved with the corresponding equalizer, resulting in the pre-ground truth image $\hat{p}$.

$$\hat{p}(l_1, l_2) = \sum_{i=1}^{L} \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} e^i(m_1, m_2)b^i(l_1 - m_1, l_2 - m_2), \quad (6.13)$$

where $(M_1, M_2)$ is the order of the equalizer $e^i$ and $b^i$ is the sub blur image on the $i$th channel. Finally, the GTI $g$ can be estimated by applying the ground truth model as shown in Fig. 6.12, to the estimated pre-GTI $\hat{p}$.

### 6.4.2 Example Model

Consider a pre-GTI, $p$, of size $S_1 \times S_2$. By applying the ground truth image formation model as shown in Fig. 6.12, we up-sample $p$ by a factor of 2, row-wise and column-wise making the resulting image $p'$ to be 4 times the number of pixels in $p$. Let $w$ be the convolved result of the interpolation filter $f$ and the blur filter $k$. If $w$ is of size $R_1 \times R_2$, by applying the theory detailed in section Section 6.3, we could derive the set of equalizers. In this case, as $L = 4$, the size of four equalizers $[M_1, M_2]$ will be constrained by,

$$[M_1, M_2] \geq [R_1 - 1, R_2 - 1],$$

as shown in (6.10). These equalizers will be applied to the sub blur images, which are formed by down sampling the blur image $b$ by a factor of 4 and the pre-GTI will be estimated by applying (6.13).
6.5 Ground Truth Image Model

Our single image multi channel restoration model depends on the GTI formation model as shown in Fig. 6.12. In this section, we detail out the underlying practical and theoretical aspects in this GTI model.

6.5.1 Sensor resolution

Complementary metal oxide semiconductor (CMOS) image sensors are now widely used in digital optical devices and have recently attracted attention in the field of high-sensitivity imaging [150]. Scaling of CMOS technology throughout the recent past has improved spatial resolution and image quality. There have been several analysis on the tradeoffs in image quality with the CMOS technology scaling [151].

The image formation process in an optical system can be written as [151]

\[
I_{image}(x, y, \lambda) = PSF(x, y, \lambda) \ast I_{ideal}(x, y, \lambda),
\]

where \( \ast \) stands for convolution operation and \( I(x, y, \lambda) \) is the spectral irradiance for a scene radiance at a spatial location \((x, y)\) for a \( \lambda \) incident light wavelength. The real image \( I_{image} \) is a distorted version of the \( I_{ideal} \). The point spread function (PSF) includes all the optical aberrations and diffractions in the optical system. Though aberrations contribute to the distortion of the ideal image, lens designers reduce it to a certain extent leaving diffraction to be the main limiting factor on resolution and contrast.

The on-axis blur of a circular diffraction-limited optical system is represented by the form of an Airy pattern [152, pp.76-78]. The diffraction limit is given by the radius \( r \) of the Airy pattern, which is defined as

\[
r = 1.22 \times \lambda \times \frac{f_l}{D},
\]

where \( f_l \) is the focal length of the lens and \( D \) is the diameter of the lens’ aperture. The term \( f_l/D \) is referred as \( f \)-number (\( f/\# \)) in photography.

An image is a sampled version of the real world in the sense that, the spectral radiance or the energy of the original scene is measured at discrete positions known
as pixels. According to the Nyquist-Shannon sampling theorem [63], the sampling operation generates aliasing if the sampling frequency of the sensor is less than twice the maximum frequency of the continuous signal. This translates to at least two pixels per Airy disk diameter for a monochrome sensor in optics. Nyquist limit, when taken together with the incident light wavelength and $f/#$ of the lens, we could determine maximum resolution and the pixel pitch of the sensor.

The microlens and spatial arrangement of the pixels are both critical to the pixel shrink technology [153]. Assuming that the microlens is designed to focus light ensuring high quantum efficiency and less crosstalk, our ground truth image model allows for a smaller pixel pitch. Under the example model discussed in section 6.4.2, the spatial arrangement of a pixel of the upsampled pre-GTI can be represented as shown in Fig. 6.14, where $(i, j)$ are ordered pairs of integers denoting the pixel locations. With this spatial arrangement, the pixel size can be shrunk by a factor of 2 times the normal pixel size. Hence our GTI model proposed in this chapter is practically applicable in the cases where the over sampled image resolution is limited by the optical diffraction limit.

### 6.5.2 Demosaicing

Demosaicing is an important part of the image processing chain in the digital camera, in order to produce a high quality digital image. To obtain full color images as well as to reduce size and cost, most digital cameras have a single sensor with a color filter array (CFA) placed in front of it. Out of several patterns for the CFA, Bayer array is one of the most commonly used in digital cameras [154].

In Bayer CFA, a luminance dominated layering is achieved by intermixing lumi-
nance (green) and first and second chrominance (blue and red), where luminance elements occur at every other chrominance element. As this concept agrees with the arrangement of the human visual system, Bayer CFA is considered as one of the most optimal spatial arrangements [155]. Demosaicing is the process of estimating the full color image, from the output of the CFA. An extensive survey has been performed on demosaicing in [156], where a detailed comparison of the demosaicing algorithms is presented.

We find that the ground truth model we apply in section 6.4.1 is analogous to the CFA and demosaicing applied in digital cameras. Let $p'$ be the raw image processed through Bayer CFA. The separated red, blue and green channels represent the spatial arrangement of $p'$ when $L = 4$, as shown in Fig. 6.14 and Fig. 6.15, where $R$ represent the corresponding pixel in the red channel. In this case red and blue channels can be considered as the row-wise and column-wise upsampled image of the pre-image $p$, whereas green channel is upsampled only row-wise. Hence, by the equalization method detailed in section 6.4.1, we can derive the set of four equalizers for red and blue channels and can consider green channel as a combination of two channels (i.e., following the same pattern as red and blue channels).
6.6 Image Restoration with the Proposed Model

6.6.1 Color Image Restoration

We perform color image restoration together with demosaicing as explained in Section 6.5.2. For simulation purposes, we create the output from the Bayer CFA with the GTI multiplied by the corresponding modulation functions as given in [155]. Hence,

$$p'(x, y) = \sum_c g_c(x, y) m_c(x, y), \quad \forall c \in \{R, G_1, G_2, B\} \quad (6.16)$$

where

$$m_R(x, y) = (1 + \cos(\pi x))(1 + \cos(\pi y))/4$$
$$m_{G_1}(x, y) = (1 + \cos(\pi x))(1 - \cos(\pi y))/4$$
$$m_{G_2}(x, y) = (1 - \cos(\pi x))(1 + \cos(\pi y))/4$$
$$m_B(x, y) = (1 - \cos(\pi x))(1 - \cos(\pi y))/4$$

and $g_R$ and $g_B$ represent the red and blue channels of $g$ while $g_{G_1}$ and $g_{G_2}$ represent the green channel of $g$ as shown in Fig. 6.16.

In this simulation we use one of the simplest interpolation kernels as $f$, known
as bilinear interpolation filter \cite{155},

\[
\begin{align*}
    f_{R,B} &= \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} / 4 & f_{G_1,G_2} &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} / 4
\end{align*}
\]

With the above interpolation filters (i.e., \(f_{R,B}\) for red and blue channels while \(f_{G_1,G_2}\) for the green channel), we perform non-blind image restoration on a colored image with the proposed model detailed in Section 6.4.

The first simulation was carried out with a blur kernel representing a linear motion of the camera with an angle of 85° in a counterclockwise direction. The restored images quality is evaluated by the objective measure of Structural Similarity (SSIM) maps and mean SSIM (MSSIM) index value proposed in \cite{21}. The results shown in Fig. 6.17 clearly evidence the PSR of the three color channels of the estimated pre-GTI, \(\hat{p}\). The Fig. 6.17(d) shows the log distribution of MSSIM index values in the SSIM maps, where the maximum MSSIM value of 1 has been achieved by all three channels in the estimated pre-GTI, \(\hat{p}\). The final estimated image, \(\hat{g}\), shown in Fig. 6.17(c) has MSSIM values of 0.9805, 0.9944 and 0.9769, respectively for red, green and blue color channels. The log distribution of the MSSIM values of the SSIM maps of the three color channels in Fig. 6.17(e) show that due to the interpolation process, PSR has not been achieved by the three channels of the estimated GTI, \(\hat{g}\).

In our next simulation, we used the same GTI with a sparse kernel representing the camera motion in real world as shown in Fig. 6.19(c). We experienced similar results in this simulation as well. The estimate of the pre-GTI was achieved with PSR while the GTI has MSSIM values as given in the previous simulation results. Hence, it shows that the process of demosaicing lessens the accuracy of the estimate of the GTI. As we have used bilinear interpolation kernel which is a simple interpolation kernel, we conclude that by using a more robust demosaicing process, we could achieve PSR for the GTI.
Figure 6.17: Results of image restoration with the proposed model.
6.6 Image Restoration with the Proposed Model

Figure 6.18: Results of image restoration with a sparse kernel.

Figure 6.19: Sparse blur kernels used for the results in Table 6.1.

6.6.2 Comparison

Even with the simplest demosaicing process, we see that our proposed model is superior than most of the existing non-blind restoration techniques based on the examples studied. We have corroborated the efficiency and effectiveness of our proposed restoration technique with the following simulation results. For the comparison, we use the techniques of iterative deconvolution with sparse prior [2], Laplacian prior [157] and quadratic prior [157] which can be executed efficiently in the Fourier domain. In order to compare the support for different blur kernels, we used two different sized sparse kernels resembling the camera motion, which are shown in Fig. 6.19.

The deconvolution results are summarized in Table 6.1, where the effectiveness of the deconvolution techniques are measured in terms of MSSIM values for the three color channels and the blur kernel sizes are given in pixels. As detailed in
the results, our proposed deconvolution technique results in consistent accuracy even for large blur kernels, while the other techniques result in poor performance in support of large blur kernels.

### 6.7 Contributions

The contributions in this chapter involves in:

1. Developing a novel novel multichannel image restoration model in Section 6.4. This restoration model eliminates the major constraint of using multiple images of the same scene in applying multichannel algorithms for a single image restoration problem.

2. Demonstrating the use of a small support inverse filter in our novel restoration model. As in conventional deconvolution algorithms which use inverse filters, the support of the inverse filter becomes unmanageable, this becomes a major advantage in image restoration algorithm.
6.8 Conclusions

In this chapter we have developed a novel image restoration model, which exploits the diversity of structures in images and the advantages of multichannel restoration algorithms. We have also shown the performance of image restoration using our novel model in conjunction with demosaicing. The simulation results justify the superiority of our model even with the simplest demosaicing filter. This model can be extended to a variety of applications which are detailed in Chapter 7.
Chapter 7

Conclusions and Future Research Directions

In this chapter we state the general conclusions drawn from this thesis. The summary of specific contributions can be found at the end of each chapter and are not repeated here. We also outline some future research directions arising from this work.

7.1 Conclusions

In conclusion, we elaborate our research work on image restoration under two main sections, development of novel deconvolution algorithms and analysis of the validity and the applicability of existing restoration techniques and claims made under them.

7.1.1 Novel algorithms

With the objective of achieving near perfect source recovery, we have introduced a novel multichannel image restoration model, which restores using only a single image. We claim this to be a major contribution, as it exploits all the benefits in multichannel restoration, while eliminating the constraints of having to use multiple images and registration of those multiple copies before applying the restoration algorithm. In addition, this thesis details out the development of efficient decon-
volution algorithms with low computation complexity and convergence speed as given below.

- Quadratic regularization functionals - We extend the existing first order and second order derivative operator functionals to class of quadratic regularization functionals with better performance.

- Constant Modulus deconvolution algorithms - A broad class of blind image deconvolution algorithms based on Constant Modulus Algorithm (CMA) are developed, which take the advantage of separable property of blur kernels.

- Segmentation based CMA deconvolution algorithm - To overcome the spatial correlation effects in CMA deconvolution and to achieve optimal results through CMA deconvolution, we develop a novel CMA deconvolution algorithm based on the higher order moments of the image.

7.1.2 Analysis of existing algorithms

Image restoration is a research area of continuous contributions. If not properly analyzed, these contributions might mislead future research if they are based only on experimental validations. In this thesis, we consider few such claims made under the maximum a posteriori (MAP) framework and demonstrate the validity of them through theoretical expositions. Specifically, we prove that high complexity derivative likelihood models offer no advantage to a properly configured normal likelihood model and properly configured linear deterministic deconvolution algorithms could deliver more efficient and effective results than the nonlinear stochastic algorithms.

In terms of the research on the applicability of the existing restoration techniques, we show that

- the dispersion constant, an important variable in CMA deconvolution algorithm does not affect on adaptation, once appropriate scalings are accounted for,
• the assumptions of source of uniform distribution and white source affects the performance of the CMA deconvolution algorithm and

• due to the relationship between de-correlation and sparse distribution, the whitening techniques such as Principal Component Analysis (PCA) could not be used to eliminate the spatial correlation effects on CMA deconvolution.

7.2 Future Research Directions

We believe there is much to be gained by using our novel multichannel image restoration model. With the combination of other disciplines such as demosaicing and whitening, there is a potential to build powerful deconvolution algorithms. Our novel multichannel restoration model is redrawn in Fig. 7.1 for clarity.

The main challenge one encounters in applying this model in practical applications is the design of filter $f$. If a generic form of $f$ could be found/designed, then perfect blind image restoration can be achieved.

As shown in the diagram, $f$ is the interpolation filter, which could take many forms. When $f$ is a demosaicing filter, the three rgb (red, green and blue) channels are deconvolved separately. As demonstrated in Chapter 6, one can investigate in employing better demosaicing filters to achieve better performance on the end result.

Another form of $f$ can be a de-whitening filter. This is justified when the pre-ground truth image $p$ is a less correlated image, where the spatial correlation effects are introduced through $f$. When $f$ is a de-whitening filter, the deconvolution algorithm can be further extended to use the benefits of CMA.

In discussing the blind image deconvolution through CMA, as shown in Chapter 5, perfect image restoration could be achieved when the whitening filter does not result in a sparse distribution. Research into whitening techniques which could deliver a less sparse distribution together with de-whitening filters would make perfect image restoration a reality.
Figure 7.1: Novel multichannel image restoration model
Bibliography


[98] M. Mignotte, “A segmentation-based regularization term for image decon-

[99] S. Geman and D. Geman, “Stochastic relaxation, gibbs distributions and 
the bayesian restoration of images,” IEEE Transactions on Pattern Analysis 

[100] D. M. Greig, B. T. Porteous, and A. H. Scheult, “Exact maximum a poste-

[101] B. A. Olshausen et al., “Emergence of simple-cell receptive field properties by 
learning a sparse code for natural images,” Nature, vol. 381, no. 6583, 

and synthesis,” in Conference Record of the Thirty-First Asilomar Confer-

[103] E. P. Simoncelli, “Bayesian denoising of visual images in the wavelet do-

priors,” in Proceedings of the 2005 IEEE Computer Society Conference on 
Computer Vision and Pattern Recognition (CVPR’05) - Volume 2, Wash-

[105] Y. Weiss, “Deriving intrinsic images from image sequences,” in 9th Interna-
Computer Society.

image using a sparsity prior,” IEEE Transactions on Pattern Analysis and 


