MULTIFRAME SUPERRESOLUTION TECHNIQUES FOR DISTRIBUTED IMAGING SYSTEMS

by

Premchandra Maravanthe Shankar

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2 0 0 8

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Premchandra Maravanthe Shankar
entitled Multiframe Superresolution Techniques For Distributed Imaging Systems
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 19 May 2008
Mark A. Neifeld

Date: 19 May 2008
Michael W. Marcellin

Date: 19 May 2008
Raymond Kostuk

Date: 19 May 2008
Nathan Goodman

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

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

Date: 19 May 2008
Dissertation Director: Mark A. Neifeld

STATEMENT BY AUTHOR

SIGNED: ____________________________

Premchandra Maravanthe Shankar

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Prof. Mark A. Neifeld for shaping me to become a good researcher. I am indebted to him for his constant encouragement, numerous discussions, and continuous support. I would like to thank professors Michael W. Marcellin, Raymond Kostuk, and Nathan Goodman for teaching me many topics in electrical engineering and optics. Their patience and time to be on my dissertation committee are very much appreciated.

Special thanks to Tami Whelan and ECE office staff for all the help with administrative matters. Thanks to my labmates of OCPL for creating a lively environment to work. I enjoyed interacting with each and everyone of them.

I am thankful to all my friends for making my life pleasant in Tucson. I am fortunate for my family's encouragement to pursue doctoral program. I wouldn't have made it so far without the constant support of my parents, brothers, and sister.

To the memory of my brother, Naveenchandra

1973-2004

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF TABLES

ABSTRACT

Multiframe image superresolution has been an active research area for many years. In this approach image processing techniques are used to combine multiple low-resolution (LR) images capturing different views of an object. The fidelity of the high-resolution image thus obtained is better than that of an object estimate from one LR image. These multiple images are generally under-sampled, degraded by optical and pixel blurs, and corrupted by measurement noise. We exploit diversities in the imaging channels, namely, the number of cameras, magnification, position, and rotation, to undo degradations. Using an iterative back-projection (IBP) algorithm we quantify the improvements in image fidelity gained by using multiple frames compared to single frame, and discuss effects of system parameters on the reconstruction fidelity. As an example, for a system in which the pixel size is matched to the diffraction-limited optical blur size at a moderate detector noise level, we can reduce the reconstruction root-mean-square-error by 570% by using 16 cameras and a large amount of diversity in the deployment geometry.

We develop a new technique for superresolving binary imagery by incorporating finite-alphabet prior knowledge. We employ a message-passing based algorithm called two-dimensional distributed data detection (2D4) to estimate the object pixel likelihoods. We present a novel complexity-reduction technique that makes the algorithm suitable even for channels with support size as large as 5×5 object pixels. We compare the performance and computational complexity of 2D4 with that of IBP. In an imaging system with an optical blur spot matched to pixel size,

2×2 undersampled images, and four LR images, the reconstruction error measured in terms of the number of pixel mismatches for 2D4 is 300 times smaller than that for IBP at a signal-to-noise ratio of 38dB.

We also present a transform-domain multiframe image superresolution algorithm to efficiently incorporate sparsity as a form of prior knowledge. The prior knowledge that the object is sparse in some domain is incorporated in two ways: first we use the popular ℓ_1 norm as the regularization operator. Secondly we model wavelet coefficients of natural objects using generalized Gaussian densities. The model parameters are learned from a set of training objects and the regularization operator is derived from these parameters. We compare the results from our algorithms with an expectation-maximization (EM) algorithm for ℓ_1 norm minimization and also with the linear minimum mean squared error (LMMSE) estimator. Using only eight 4×4 pixel downsampled LR images the reconstruction errors of object estimates obtained from our algorithm are 5.5% smaller than the EM method and 14.3% smaller than the LMMSE method.

CHAPTER 1

Introduction

This dissertation addresses the topic of multiframe image superresolution (MIS). MIS is a technique to obtain an estimate of a high-resolution (HR) object by combining multiple low-resolution (LR) image measurements [1]. The multiple cameras capturing these LR images have a common field-of-view and hence, a region of the scene or object is common to all images. We can improve the image fidelity in the common region by fusing the multiple images. A sequence of LR images may be obtained by using a customized array of cameras pointing at the scene ("multi-aperture"), or a sensor network of cameras ("distributed"), or as images captured at different times from a camera on a satellite orbiting the earth, or as consecutive frames from a video camera ("multiframe"). In this dissertation we use terms "multiframe", "multi-aperture", and "distributed" imaging interchangeably. The knowledge of camera positions, imaging blurs (e.g., optical and pixel) and statistics of pixel noise are used to formulate mathematically the image formation. Image restoration refers to techniques where knowledge of image formation is used to estimate the original object by deblurring and denoising a measured image. Restoration methods using multiple images can yield an object estimate having much higher fidelity than a traditional single frame image restoration, and hence they are referred to as MIS techniques. We discuss the reasons for a higher image fidelity from multiple frames in chapter 2.

A typical digital camera measures a representation, often isomorphic, of an object. Optical blur of the camera lens, sampling by the pixel array, and measurement noises are a few example degradations that limit the image resolution. Often the degradation operators in an imaging system are linear and hence a LR image measurement can be expressed as a system of linear equations. Even the multiframe imaging process can be represented by a system of linear equations as we will see in chapter 4. Image restoration techniques, either linear or nonlinear, offer solutions to this inverse imaging problem. In addition to restoration, the image registration is a critical issue in multiframe methods. Because of different locations of cameras a particular object pixel in the scene is mapped to different image pixels in different cameras. The method for aligning these images with respect to a reference image co-ordinates is called image registration. This operation requires the knowledge of camera parameters that describe its position and orientation with respect to a reference camera. When the camera parameters are unknown or inaccurate, they need to be estimated from the measured LR images. The image registration can be either performed as a separate step preceding image restoration or can be combined with image restoration.

MIS techniques may be employed in many defense and commercial applications including astronomy, medical imaging, satellite imaging, surveillance, etc. Because of diverse application fields there is extensive literature on the topics of image registration, and multiframe image superresolution. The main problem remains the same in all these superresolution applications; however often the importance is shifted to a particular component. For example in medical imaging data may obtained from cameras having different modalities and hence photometric image registration and hyperspectral image restoration may be necessary.

1.1 Previous Work

The original research on image MIS was carried out by Tsai and Huang [2]. They were the first to coin the phrase "multiframe" superresolution. They considered restoring a HR image from a sequence of LR, undersampled satellite images of an aerial scene. Each image is assumed to be shifted with respect to a reference image, and images were assumed to be free from degradations due to blur and noise. They used Fourier domain methods to estimate the spectrum of the HR image. Ur and Gross [3] included blur in the imaging model and assumed that both the imaging process and the relative shifts among LR images are known precisely. The LR images are interpolated onto a finer grid and the resulting blurry image is restored using pseudo-inverse techniques. Irani and Peleg [4] considered LR images that are both shifted and rotated. They proposed an iterative back projection (IBP) method, based on principles found in computer aided tomography, to restore a HR image. The algorithm iteratively back projects the difference between the actual LR images and predicted LR images from the model to update the current HR object estimate. A recursive regularized least-squares method in the discrete Fourier domain is proposed by Kim et al. [5, 6] and to combat the effects of noise and imaging blur, the regularization function is iteratively updated. Tom et al. [7, 8] presented two approaches for superresolution from multiple frames with unknown sub-pixel shifts. In the first approach an expectation-maximization algorithm is used to perform registration and restoration (without interpolation) and a regularized interpolation is performed as a separate step to enhance the image resolution. The second approach performs registration, restoration and interpolation simultaneously. Elad and Feuer proposed an unified methodology toward the multiframe superresolution problem using three estimation/restoration tools [9]. They analyzed the superresolution problem from the points of view of maximum-likelihood (ML), maximum a-posteriori, and projection onto convex sets (POCS) approaches. A hybrid method combining ML and POCS is presented. Komatsu et al. ([1],Chapter 7) showed that utilizing multiple cameras having different pixel shape and size increases the superresolution image reconstruction fidelity. It complicates the image registration however, and an iterative alternate minimization strategy is used to solve the registration.

To reduce the computational complexity of the superresolution problem an efficient preconditioner has been derived by exploiting the structure of the imaging operator [10]. The conjugate gradient method is employed to solve the Tikhonov regularized inverse imaging problem. Farsiu et al. [11] proposed a steepest descent based superresolution technique that minimizes the ℓ_1 norm instead of the L_2 norm of the difference image (i.e., between actual and predicted). A total variation type regularization has also been used to make the superresolution technique robust. Vandewalle et al. [12] reconstruct a HR image from multiple images that are sampled below the Nyquist rate by solving the linear inverse imaging problem in the frequency domain. A subspace method in the frequency domain is used to compute the shifts for image registration. The performance of multi-frame superresolution is also analyzed statistically by using the Cramer-Rao inequality to bound the L_2 error norms [13, 14]. A gradient-based image registration method is also analyzed in terms of estimator bias and variance.

Superresolution techniques have also been applied to sequences of image data from a video source. Elad et al. [15] treated video superresolution as a sequence of image superresolution problems, and and approximate version of the Kalman filter is used to efficiently reconstruct images. Shekharforoush et al. [16] proposed a method to superresolve images from video sequences based on generalized sampling theory. The camera blur functions are utilized to form optimal Riesz bases on which the image frames are projected to overcome sampling ill-posedness. Schultz et al. [17] utilize both spatial and temporal information present in a short video sequence to create a HR still image. Bayesian restoration method with a discontinuity-preserving prior is used to extract a single high-resolution frame from a short low-resolution sequence. A problem of restoring a high-resolution frame from a low-resolution compressed video sequence is addressed by Segall et al. ([1], chapter 9). They include the knowledge of underlying compression scheme and the compression artifacts are reduced along with increasing the resolution. Chen et al. [18] used a Huber-Markov random field model based Bayesian method to extract a high-resolution video stills from MPEG video sequence. Mateos et al. [19] proposed a join estimation of motion parameters and superresolved image from compressed low-resolution video. The algorithm exploits the existing correlation between the high and low resolution frames.

An imaging system is generally ill-conditioned because of degradations due to optical blur, sampling, and noise [20]. A given measured image can be generated by several candidate object estimates using the image formation model. It is well-known that any available prior knowledge about the object space may be incorporated to improve the image fidelity. Example image priors include non-negativity constraint on the object pixel intensity, finite spatial support, etc. A smoothness constraint for the object can be incorporated in a Tikhonov regularized image restoration methods [21]. Objects having significant discontinuities or edges can be formulated efficiently using edge-preserving priors such as total variation or Laplacian priors [22, 23, 24, 25, 26]. The first (mean) and second order (autocovariance) object statistics can be incorporated in linear minimum mean squared error estimation, and the image restoration problem is solved often efficiently in frequency domain approaches, e.g, Wiener filter [20]. Sophisticated object models can be developed in a statistical framework using pixel probability distributions. Image restoration incorporating statistical object priors are performed using Bayesian methods such as maximum a-posteriori estimation. Markov random fields, Huber-Markov fields, and/or Gauss-Markov fields are often used to capture the local correlation structure of object pixels [27, 17].

Often the object statistics can be better modeled in a transform domain different than the space-domain. For example a part of the object frequency spectrum may be known a-priori, and it can be utilized in an image restoration algorithm based on POCS [28]. Multiresolution analysis tools such as wavelets allow decomposition of an object in multiple scales and orientations. Superresolution algorithms can be modified to incorporate such transform-domain object prior knowledge and to yield a high-resolution object estimate in that domain. Edge-preserving priors are incorporated in wavelet subbands in a regularized frameworks to superresolve an object from multiple low-resolution images [29]. Nguyen et al. proposed a wavelet based interpolation-restoration algorithm for image superresolution [30]. This technique takes advantage of regularity and structure of non-uniformly sampled data points from multiple LR frames. Pseudo-inverse based methods in the wavelet-domain are also reported for image superresolution [31]. Capel et al. proposed an image superresolution technique that constrains the object estimate to lie in a sub-space determined by principal component analysis learnt from object examples [32].

Other object priors that can be utilized to improve image restoration performance include finite pixel values, object sparseness, objects belonging to a known set. If the object consists of elements (e.g., numbers, letters or faces) from a known set the superresolution procedure may first involve recognizing the element from the set of low-resolution images. Once all elements in the object are recognized, the image can be superresolved by using the known high-resolution copy of elements [33].

Superresolution methods have also been applied in astronomy where multiple snapshots of a distant object are obtained by dithering the telescope [34]. The variable pixel linear reconstruction method is described to combine dithered images in the presence of cosmic rays. In an imaging sensor network the sensor nodes transmit low-resolution and compressed images of the object, and superresolution techniques are used as a "decoding" step to enhance image resolution at the receiving station [35]. Superresolution techniques based on Walsh-functions are used to enhance the resolution of satellite imagery by utilizing translated images [36]. Superresolution methods are also used in several medical imaging modalities such as MRI, fMRI etc. [37]. They demonstrate a superior SNR improvement in images reconstructed using superresolution. Several optical architectures have been proposed to increase the image capture diversity in a single imaging device. Tanida et al. presented Thin Observation Module by Bound Optics (TOMBO) that simulates a compound-eye imaging system [38]. It acquires multiple low-resolution images of a target object that are shifted with respect to one another. The image is reconstructed using sampling and back projection methods. An adaptable steerable field-of-view version of TOMBO, called PANOPTES is introduced by Christensen et al. [39]. It utilizes arrays of microelectromechanical mirrors to adaptively redirect the fields of view of multiple LR cameras.

1.2 Overview of the dissertation

Chapter 2 of this dissertation motivates the multiframe superresolution techniques by justifying the need for improvements in image fidelity. We formulate mathematically the process of image formation, discuss the sources of image degradations, and identify the system parameters that affect image fidelity.

Chapter 3 describes a multiframe image restoration method called iterative back projection (IBP). In this chapter we quantify the image fidelity in terms of error norms and object frequency contents. We will analyze how the amount of optical blur, size of image pixels, and the number of cameras control the achievable image fidelity. It will be shown that not only the number but also the diversity in geometric deployment of cameras influence the fidelity. Image registration is not the central issue of this dissertation, however this chapter provides results showing degradations in the restored image fidelity if the camera parameters are not known accurately. We present a camera array that is built on a flat surface and provide example images for demonstrating the advantages of distributed superresolution. The results observed in this chapter can be used as design strategies for implementing camera arrays with flat form-factor. The chapter also demonstrates that the effects of pixel blur and measurement noise can be alleviated better than that of optical blur in multiframe superresolution systems.

The image restoration technique described in chapter 3 utilizes a system model and measured LR images. It is well known in signal/image estimation theory that any available object prior knowledge should also be incorporated to improve the restoration fidelity. In chapter 4 we explain a communication-theoretic algorithm called two-dimensional distributed data detection (2D4) to incorporate finite-alphabet constraint of object pixels to superresolve binary imagery. The algorithm estimates the likelihoods of each object pixel based on the measured pixels and the neighboring object pixels in a message-passing fashion. The exponentially growing computational complexity of the algorithm limits its applications to small images and small channel support. We propose a novel object-independent and data-driven complexity reduction technique that makes the algorithm suitable for imaging channels with large support. We compare the 2D4 algorithm with linear minimum mean square error estimate (LMMSE) that incorporates the object autocorrelation prior knowledge. The comparisons of algorithm performance and computational complexity are provided. The 2D4 demonstrates significant improvement in image fidelity over the LMMSE in terms of number of pixel errors. Observing that both algorithms incorporate object priors, results in this chapter not only signify advantages of using object prior knowledge but also the form in which it is used. In case of binary imagery finite-alphabet is a superior form of object prior knowledge than its power spectrum density. The complexity reduction technique proposed in this chapter is applicable in any general maximum likelihood estimation of binary sequences or 2D patterns.

Chapter 5 discusses a transform domain approach for image superresolution. Often objects can be better modeled and analyzed in a transform domain instead of the space-domain. Transform domain techniques are used for spectral analysis, data compression, dimensionality reduction, etc because an object can be represented compactly in that domain. The compact representation may be quantified in terms of object sparsity. The number of non-zero pixels (space-domain) or coefficients (transform-domain) in an object represents the sparsity. Recently there has been tremendous interest in regularized signal/image restoration using the object sparsity constraint. Alternate metrics have been devised to quantify the object sparsity as the strict metric (the number of non-zero elements) leads to a combinatorial optimization problem. We formulate the image superresolution problem as maximum a-posteriori (MAP) estimation problem and we incorporate the object prior knowledge that many natural objects are sparse in the wavelet-domain. An object estimate is obtained by solving a series of linearized regularized least squares problems. We also model the wavelet coefficients of natural images at multiple scales and orientations using generalized Gaussian functions. The prior model parameters are learned from a set of training images. We compare our algorithm with a ℓ_1 norm minimization technique and LMMSE estimation, and show that it performs better in terms two important image fidelity metrics. The algorithm described in this chapter is based on a doubly-iterative strategy, and may be applicable to solve general problems with non-quadratic objective cost functions. We also describe a laboratory experimental setup to demonstrate the distributed image superresolution. Each camera is calibrated and its response is estimated. We apply the new technique on the low-resolution images gathered from these cameras, and present example restored images.

CHAPTER 2

Multiframe Superresolution Imaging Model

High-Resolution (HR) images are desired in many imaging applications such as surveillance, medical imaging and remote sensing. Such a HR image can directly be obtained from high precision optics together with dense photodetector (e.g., CCD) arrays. This approach is too expensive to be suitable in many commercial applications. Also, by reducing the pixel size, the amount of light available for each pixel decreases. This reduces the signal to noise ratio (SNR) of each pixel measurement [40]. An alternate solution is to use multiframe image superresolution as we mentioned in the previous chapter [1]-[41]. It obtains a HR image by combining multiple low-resolution (LR) images. A schematic diagram of the setup of a multiframe superresolution system is shown in Fig. 2.1. Inputs to the system are the K number of LR frames and the output is a single HR estimate of the object. Multiframe image restoration technique can either be linear or nonlinear, and may use knowledge of the imaging process, camera geometry, and noise statistics. The images shown in Fig. 2.2 are examples of HR estimates obtained from (a) single frame, and (b) 32 frames. As seen in this example, the image obtained by combining 32 LR images has better visual quality than that from one LR image.

2.1 Motivation

To understand the motivations for using multiframe superresolution techniques and reasons for improvement in the image fidelity, we discuss a simplified analysis. Con-

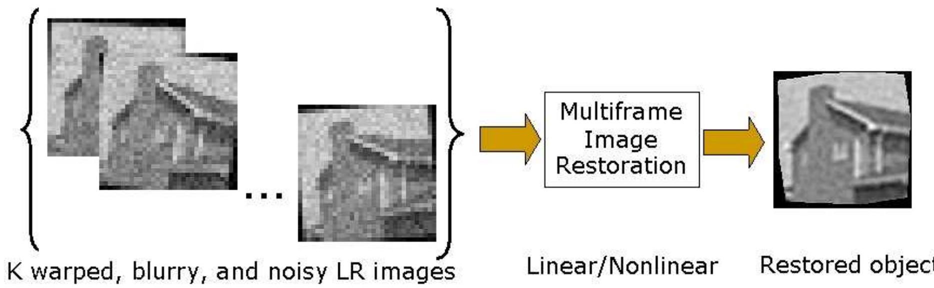


Figure 2.1: Schematic diagram of a multiframe image superresolution system.

sider a rectangular array of pixels from a LR camera shown as dark squares in Fig. 2.3. To obtain an object estimate on a finer HR pixel grid these regular pixels can be interpolated using techniques such as pixel replication, bilinear/bicubic interpolation, etc [42]. If we have pixel arrays from two more LR cameras they may be mapped as shown in diamonds and circles in Fig. 2.3 depending on cameras' positions and orientations. The object can now be estimated using these larger number of irregularly spaced pixel measurements. One might expect the estimate to be of better fidelity than the one obtained using a single frame. Hence the "diversity" in these cameras can be seen as providing additional sub-pixel information about the object. We analyze the effect of camera diversity on the image fidelity in chapter 3.

Another advantage of multiframe systems can be understood by considering two imaging systems with the detector structures depicted in Fig. 2.4. The finer grid represents the object pixels whereas the thicker grid represents the image pixels. The camera pixels in these two detectors arrays have different widths D_1 and D_2 ($D_1 > D_2$). Alternatively, these systems might also correspond to cameras with two different magnification settings: the imager in (b) having higher magnification than (a). Consider the one-dimensional pixel transfer functions associated with these two imagers, $H_k\left(w\right) = sinc\left(D_k w\right),\ k = 1, 2$ where w is the spatial frequency. Ex-

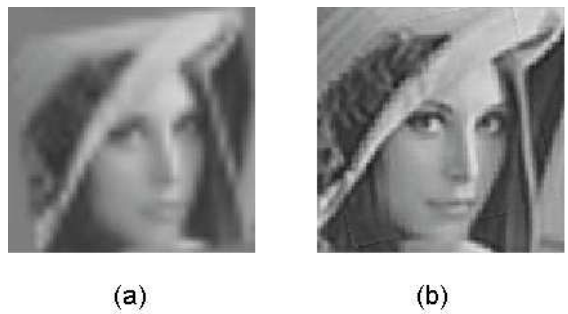


Figure 2.2: Example images reconstructed using (a) single frame and (b) 32 frames. The number of pixels are increased by 4× in both image coordinates.

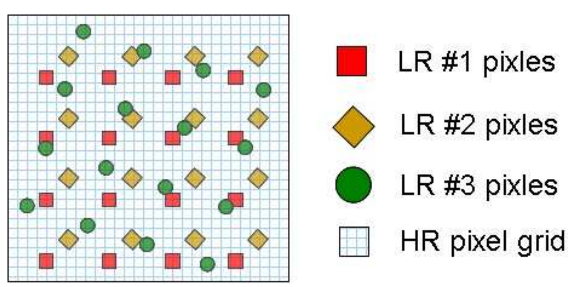
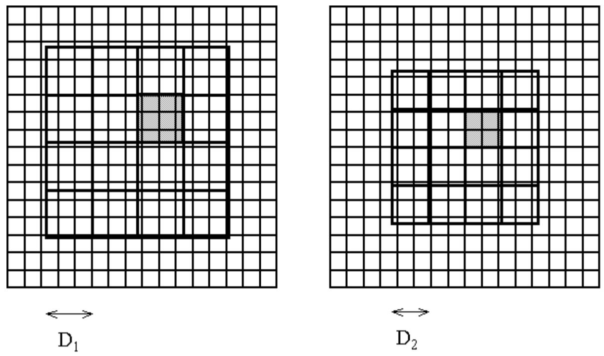


Figure 2.3: Image pixels from different low-resolution cameras mapped onto a high-resolution pixel grid.

amples are plotted in Fig. 2.5(a) for two imagers with pixel widths of $D_1 = 1.0$ and $D_2 = 0.8$. Note that the spectral nulls of these transfer functions occur at different frequencies. We will assume that the detector measurement is corrupted by additive white Gaussian noise (AWGN) with variance, σ^2. The first benefit that can be obtained from multi-aperture imaging is increased SNR. For example, averaging K measurements of the same object can be used to reduce the noise variance by K [42]. A more important benefit of multi-aperture imaging is channel diversity. Consider the frequency domain analysis of image averaging to visualize the benefit of diversity. Let $S(w)$ represent the Fourier transform of the object. Ignoring blurring due to optical elements, the spectrum of the k^{th} measured image can be expressed as $R_k(w) = S(w) H_k(w) + \sigma^2$. The signal and the noise components of $R_k(w)$ are $S(w) H_k(w)$ and σ^2 respectively. The SNR, which can be defined as the ratio of signal power to noise power, of the k^{th} measured image is $SNR_k(w) = 10 \log_{10} \frac{S^2(w) H_k^2(w)}{\sigma^2}$. The SNR of an image corrupted only by additive noise of the same strength is $SNR_O(w) = 10 \log_{10} \frac{S^2(w)}{\sigma^2}$. The reduction in SNR of the k^{th} measured image due to pixel blur is $L_k(w) = SNR_k(w) - SNR_O(w) = 10 \log_{10} H_k^2(w)$ and is plotted in Fig. 2.5(b) for the two pixel widths considered above. The SNR is reduced significantly at and around the spectral nulls for each measured spectrum. The object information at frequencies corresponding to these spectral nulls is completely lost and can not be recovered by just one measurement. However, if we measure two images with different pixel sizes or different camera magnifications such that the spectral nulls don't coincide, the object information at the spectral nulls of the first imager can be recovered from the second measurement and vice-versa. Extending this idea, if there are K different imagers, recovering the object information at the spectral nulls is relatively simple as long as the nulls



(a) (b)

Figure 2.4: Detector structures in two LR imagers with pixel width of (a) D_1 and (b) D_2. $D_1 > D_2$

don't coincide with each other as depicted in Fig. 2.5(c). This idea of spatial diversity in multiple channels is also exploited in blind identification problems in digital communication systems [43, 44].

2.1.1 Imaging model and diversity

Before discussing image processing techniques to combine multiple image frames we describe the image formation model used in this dissertation. We refer to an imager as a forward imaging channel whenever appropriate. Fig. 2.6 illustrates several degradations that can occur to a HR scene in a single channel of a multi-aperture imaging system. A HR image undergoes a geometric transformation, defined by the imager orientation and location with respect to an arbitrary reference imager, optical blurring due to the lens, pixel blurring and spatial sampling at the detector, and degradation by AWGN. Let $\{y_k\,(i,j)\,; k = 0, ..., K-1\}$ represent a set of K LR images. The k^{th} forward imaging channel can be described mathematically as follows:

$$y_k(i,j) = s(h_p(x,y) * h_a(x,y) * t_k(f(x,y))) + n_k(i,j) \tag{2.1}$$

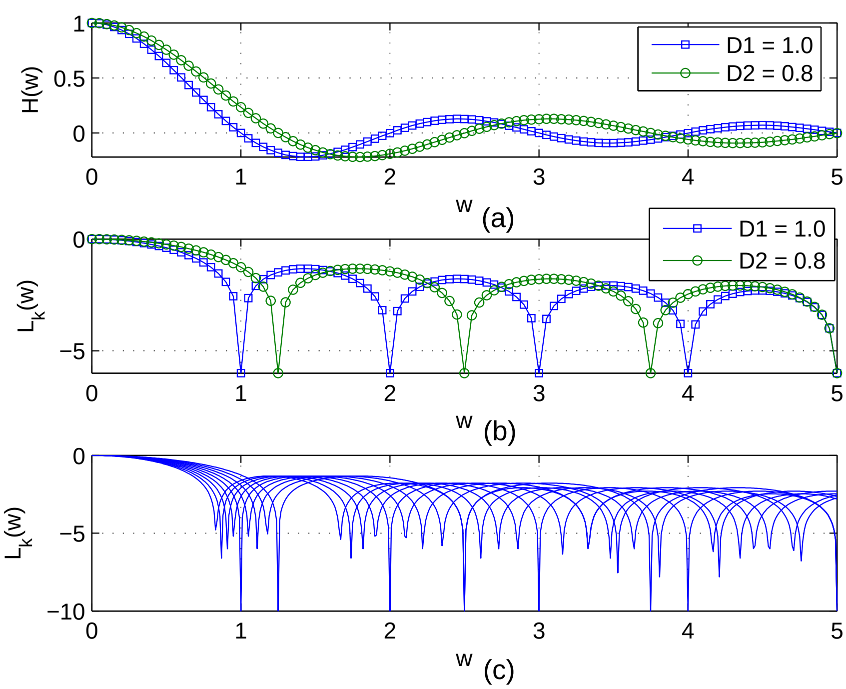


Figure 2.5: (a) Pixel transfer functions for two pixels of different size, (b) reduction in SNR for two different imagers. (c) Reduction in SNR for several imagers.

where $f(x, y)$ is the HR scene, $h_a(x, y)$ is a linear space-invariant optical point-spread function (PSF), $h_p(x, y)$ is the pixel PSF, t_k is the geometric transformation in the k^{th} channel, s is the sampling function mapping (x, y) to (i, j), and $n_k(i, j)$ is the additive noise in the k^{th} channel. The noise is assumed to be zero-mean white Gaussian distributed and independent of the measured intensity. We note that all operations occurring in Eq. 2.1 are linear, and the k^{th} imaging channel can be expressed by a system of linear equation as

$$y_k = SH_{ccd}H_{opt}T_k f + n_k. \tag{2.2}$$

The vector y_k of size M represents the measured pixels of the k^{th} LR image un-rastered into 1D. The vector f of size N represents the unknown HR object pixels. The matrix T_k represents a geometric transformation of the object pixels due to the position and orientation of the k^{th} camera with respect to a reference coordinate system. It can be characterized by four parameters namely scale, rotation, and translations in two image coordinates. The matrices H_{opt} and H_{ccd} represent the degradations (e.g., blur) due to optics and CCD detectors respectively. The matrix S is the under-sampling operator. The vector n_k of size M represents additive white Gaussian noise samples with variance σ^2. We make use of the formulation in Eq. 2.1 in the chapter 3 and the formulation in Eq. 2.2 in chapters 4 and 5.

We assume that all cameras observe the 2D scene from a common plane. The optical axis of each camera is shifted in both image coordinates by some (possibly sub-pixel) distance with respect to the rest of the cameras. Also, each camera has a different orientation and magnification. These assumptions allow us to model the geometric warping by an affine transformation. A point (x, y) in the scene is transformed to a new point (x', y') in the image by T_k according to the equation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = s_k \begin{bmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_k \\ b_k \end{bmatrix} \tag{2.3}$$

where s_k is the magnification, θ_k is the rotation angle, and, a_k and b_k are the translations in two object coordinates. The set of "channel parameters" $\{s_k,\ \theta_k,\ a_k,\ b_k : k = 1, ..., K\}$ define the diversity in the multi-aperture imaging system.

The blur introduced by a finite aperture in an incoherent optical system can be modeled as a low-pass linear-shift-invariant channel. We assume a rectangular aperture with the diffraction-limited transfer function $H(w_1, w_2)$

$$H(w_1, w_2) = \Lambda(\frac{w_1}{W_c})\Lambda(\frac{w_2}{W_c}) \tag{2.4}$$

where the triangle function, $\Lambda(.)$, is defined as

$$\begin{aligned} \Lambda(x) &= 1 - x \quad \text{if } \ |x| < 1 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

$w_1$and w_2 are spatial frequencies and W_c is the diffraction-limited cutoff frequency. The optical PSF, $h_a(x, y)$, in Eq. 2.1 is the inverse Fourier transform of $H(w_1, w_2)$. Each detector spatially integrates $D \times D$ HR pixels and each HR pixel is normalized to a unit length in the (x, y) plane. This operation can be expressed by a pixel PSF,

$h_p(x, y)$:

$$\begin{aligned} h_p(x, y) &= \frac{1}{D^2} \quad for \quad -\frac{D}{2} \leq x, y \leq \frac{D}{2} \\ &= 0 \quad otherwise \end{aligned}$$

The operation of the detector can be represented by a convolution with $h_p(x, y)$ followed by ideal sampling with a Dirac delta function with spacing D. The detector measurement is corrupted by AWGN with variance of σ^2.

The extent by which the diffraction limited blur spot covers the high-resolution pixel grid is used to normalize the optical cutoff frequency W_c. Fig. 2.7 is a 1D depiction of the blur spot and detector pixel for $W_c = 1.0$, $D = 2$ (Fig. 2.7(a)) and $D = 4$ (Fig. 2.7(b)). The blur spot size (d_b) can be defined as the separation between the first zeros of $h_a(x) = \text{sinc}^2(xW_c)$: $d_b = \frac{2}{W_c}$. For $W_c = 1.0$, $d_b = 2$, i.e., the optical blur covers 2 HR pixels. A detector array with $D = 2$ in Fig. 2.7(a) is matched to the optical blur spot. A detector array with $D = 4$ that covers twice the blur spot is shown in Fig. 2.7(b). Both of these systems are aliased, since the spatial sampling rate in both cases is smaller than the Nyquist rate. We will discuss in the next chapter how the IBP algorithm uses knowledge of the forward imaging channels to reconstruct a HR image, and analyze the effect of channel diversity on the reconstruction fidelity.

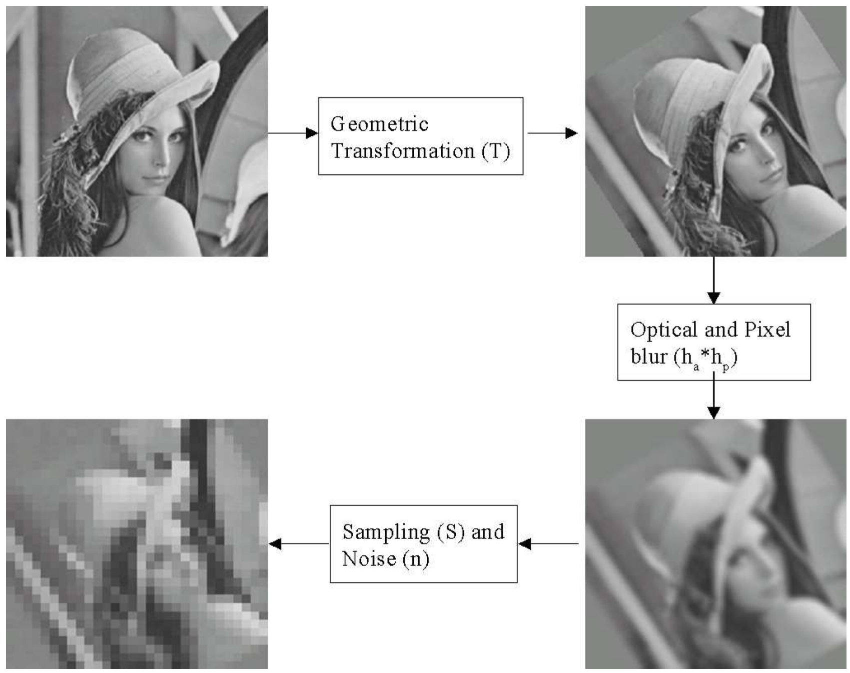


Figure 2.6: An illustration of degradations in the imaging model

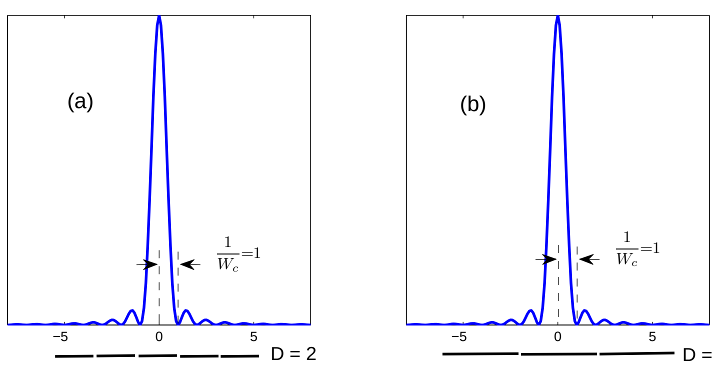


Figure 2.7: Comparison of diffraction limited blur spot and pixel size for (a) $W_c = 1.0$, $D = 2$, and (b) $W_c = 1.0$, $D = 4$. Blur spot $= \frac{2}{W_c}$ and pixel size $= D$.

CHAPTER 3

Diversity Analysis in Multi-Aperture Imaging Systems

3.1 Introduction

Multi-aperture or distributed imaging is analogous to "multiframe" superresolution systems in which a sequence of temporal frames from a camera are used to enhance the spatial resolution of the image frames. In our system, images from multiple cameras are combined to reconstruct a high resolution image. Image superresolution can be viewed as a method by which excess SNR is exploited to extract image details beyond the diffraction limit [45, 46]. As explained in the previous chapter multi-aperture imaging achieves resolution enhancement in two ways: (a) simple averaging of multiple measurements can yield excess SNR that can be exploited as in conventional super-resolution and (b) channel diversity provides additional object "information" by mitigating idiosyncracies like nulls and/or fades in the optical transfer function.

In this chapter, we consider multiple LR images that are shifted, rotated, and magnified with respect to each other. Using a modified IBP algorithm to reconstruct the HR image, we study the effect of channel diversity in achieving higher resolution. We study the trade-off between the detector noise and the pixel size, and the sensitivity of the system to imperfect knowledge of the channel parameters. We quantify the performance in terms of the root mean square error (RMSE) between the original and reconstructed images. The chapter is organized as follows: in section 3.2, we describe the imaging system model as well as the types and amounts of diversity involved in the system. We also briefly describe the modified IBP algorithm used in this work. In section 3.3, we present the simulation results and quantify the trade-off between diversity and system performance. In section 3.4, we introduce an experimental realization of a multi-aperture imaging system and discuss the practical issues involved in developing such a system.

3.2 Multi-Aperture Imaging System

3.2.1 Iterative Back Projection

The IBP algorithm was first proposed by Irani and Peleg [4]. The algorithm starts with an initial estimate $f^{(0)}$ for the HR image. The imaging process is simulated using Eq. 2.1 to obtain a set of LR images $\left\{y_k^{(0)}\right\}$ corresponding to the measured LR images $\{y_k\}\,,\ \ k=1,...,K.$

$$y_k^{(0)}\left(i,j\right)=s\left(h_p\left(x,y\right)*h_a\left(x,y\right)*t_k\left(f^{(0)}\left(x,y\right)\right)\right)$$

If $f^{(0)}$ had been the correct HR image, then the simulated images $\left\{y_k^{(0)}\right\}$ should be the same as the observed images $\{y_k\}$ in the absence of noise. The difference images $\left\{y_k-y_k^{(0)}\right\}$ are then calculated and "back projected" to improve the current estimate of the HR image. This process is repeated in order to minimize the residual error between the observed LR images and the simulated LR images. This residual RMSE, $e^{(n)}$, at the n^{th} iteration is defined as $e^{(n)}=\sqrt{\frac{1}{KM}\sum\limits_k\sum\limits_{i,j}\left(y_k\left(i,j\right)-y_k^{(n)}\left(i,j\right)\right)^2}$ where the LR image size is $M=L\times L$ pixels. The process of IBP is depicted schematically in Fig. 3.1. The back projecting of the difference image involves interpolation, inverse filtering and un-warping the geometric transformation introduced by the forward imaging channel. The first step (S^{-1}) is to interpolate the difference image to make it the same size as the HR image. Several methods can be used for interpolation including simple pixel replication, bilinear interpolation [42] , or more sophisticated methods like bi-cubic spline [47]. In our simulation studies, bilinear interpolation is employed. The second step is to apply a back-projection kernel (h_b) to the interpolated image. There are many choices for the back-projection filter to weigh the influence of LR pixels to HR pixels. For deblurring a single image without increasing the resolution, Irani and Peleg [4] use a filter $h_b(i,j) = \frac{h_a(i,j)}{c}$ where c is a constant. If $c = \sum_{i,j} h_a^2(i,j)$, this filter represents a pseudo-inverse of $h_a(i,j)$. We use $h_b(i,j) = \frac{1}{c}$ in the back projection. We notice that this is also a pseudo-inverse filter if the constant $c = \sum_{i,j} h_a(i,j)$. This method of excluding an explicit back-projection filter has been studied to reduce the complexity of algorithm without sacrificing reconstruction fidelity [48, 49]. The third step $\left(T_k^{-1}\right)$ inverts the geometric warping due to the camera orientation that is defined by equation 2.3. When all channel parameters are known precisely, this operation is a perfect inverse affine transformation. These un-warped difference images are weighted by a relaxation parameter λ $(0 < \lambda < 1)$ and are added to the current estimate of the HR image. The iterative process is stopped either when a certain number of iterations are performed or when the residual error $e^{(n)}$ does not vary significantly between two successive iterations. The relaxation parameter controls the convergence rate of the iterative algorithm: with a high value of λ the algorithm converges faster to a solution; however, the solution may be a local minimum. With a small value of λ the algorithm converges slowly to a smoother HR image. In the next section, we will present simulation results of systems with varying amount of diversity, different numbers of imagers, and different pixel sizes.

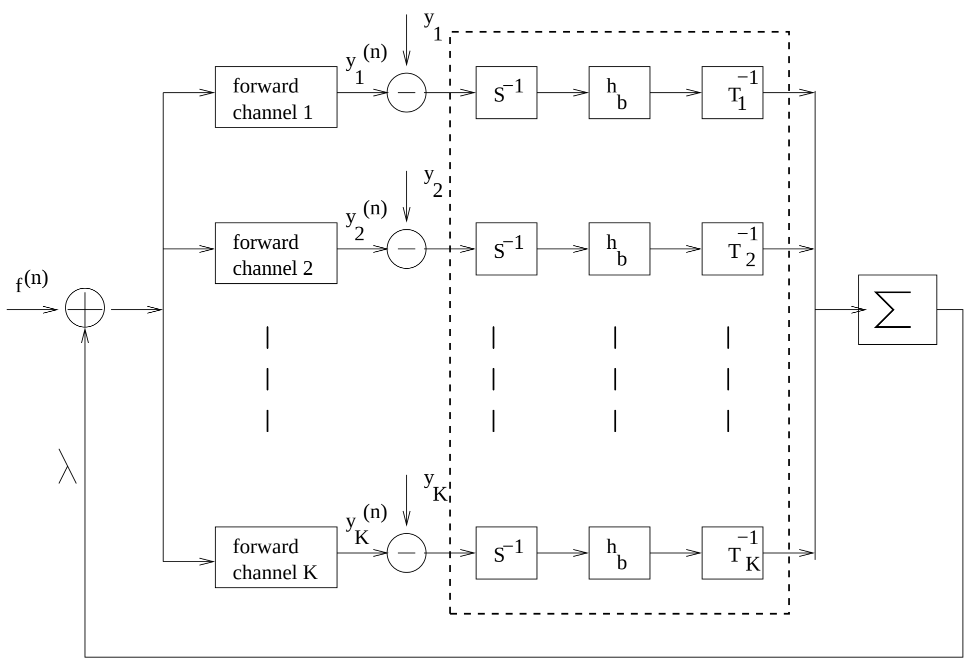


Figure 3.1: Illustration of IBP

3.3 Simulation Results

In this section we will quantify the diversity in the system, give numerical details of the parameters in the algorithm and present the simulations results. We describe the implementation details of an experimental multi-aperture imaging system in the next section. The HR image used for simulation purpose is a spatially discrete gray scale image of size $\sqrt{N} \times \sqrt{N}$ shown in Fig. 3.2. Fig. 3.3 shows sample images captured with $W_c = 0.25$ and, $D = 2$, 4 and 8 at SNR = 54dB. The SNR is defined as $SNR = \frac{\frac{1}{N}\sum\limits_{i,j} f^2(i,j)}{\sigma^2}$. Initial estimate of HR image, $f^{(0)}$, is formed from averaging all LR images after interpolation and un-warping the geometric transformation. Fig. 3.4 illustrates how the relaxation parameter, λ, affects the convergence of the iterative process in the reconstruction from 4 LR images with $D = 2$ at $SNR = 54dB$. The fidelity of the reconstructed image is quantified by the RMSE between the reconstructed image, $\widehat{f}$, and the original image, f, which is

Figure 3.2: Original Image

defined as $\sqrt{\frac{1}{N_C}\sum_{(i,j)\in C}\left(f\left(i,j\right)-\widehat{f}\left(i,j\right)\right)^2}$,where C is the set of pixels seen by all cameras and N_c is the total number of those pixels. Fig. 3.4(a) shows the residual RMSE versus the number of iterations, while (b) shows the reconstruction RMSE versus the number of iterations. A large value of λ converges faster but does not converge to the minimum reconstruction RMSE. A smaller value of λ converges much slower but to a better solution. In the simulation results that follow we use $\lambda = 0.5$ unless otherwise mentioned.

We study multi-aperture systems for two different diversity configurations. In the low-diversity case, we select the magnification in each LR imager uniformly from the range 0.95 and 1.05, the rotation between -5 and $+5$ degrees and the translations in both image coordinates between $-D$ and D HR pixels. In the high-diversity case, these ranges are 0.8 and 1.2, -15 and $+15$ degrees and $-1.5D$ and $1.5D$ HR pixels respectively. We compare these two cases with the shift-only case in which pixels are translated between $-D$ and $+D$ pixels. An example reconstruction using

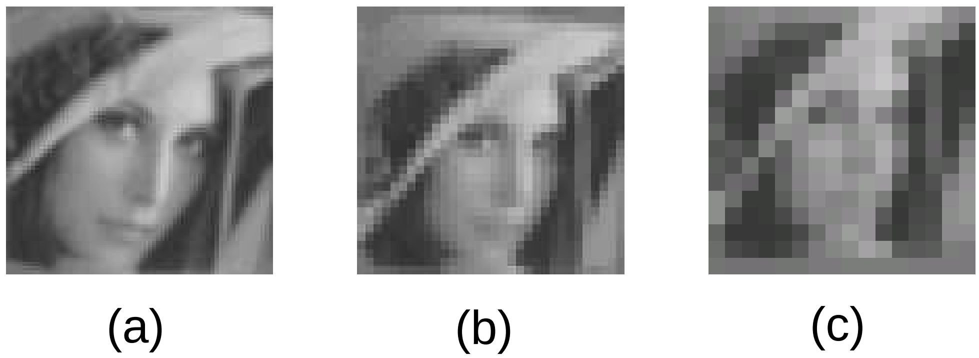


Figure 3.3: Sample measured LR images (a) $D = 2$, (b) $D = 4$, and (b) $D = 8$ at $SNR = 54dB$ and $W_c = 0.25$

these three configuration can be found in Fig. 3.5. These images are reconstructed at $SNR = 46dB$, $W_c = 1.0$ using 32 cameras with (a) shift-only, (b) low-diversity, and (c) high-diversity configurations. Each image shows only those pixels that are seen by all cameras. Because of the large variation in channel parameters for the high-diversity case, fewer pixels are seen by all cameras as compared to the low-diversity and shift-only diversity cases. Although all three images look visually pleasing compared to one of the low resolution image in Fig. 3.3(b), the edges in the high-diversity image 3.5(c) are recovered more accurately. To quantify the reconstruction performance, we use Monte-Carlo simulation to estimate the RMSE between the original and reconstructed images. For a given number of cameras, we simulate 10 different realizations of each diversity configuration. Fig. 3.6 plots RMSE versus the number of cameras for different pixel sizes, optical blur levels and SNR. Fig. 3.6(a) and (b) are for $W_c = 0.5$ and SNR = 54dB with $D = 2$ and $D = 4$ respectively; whereas (c) and (d) show the performance for $W_c = 1.0$ and

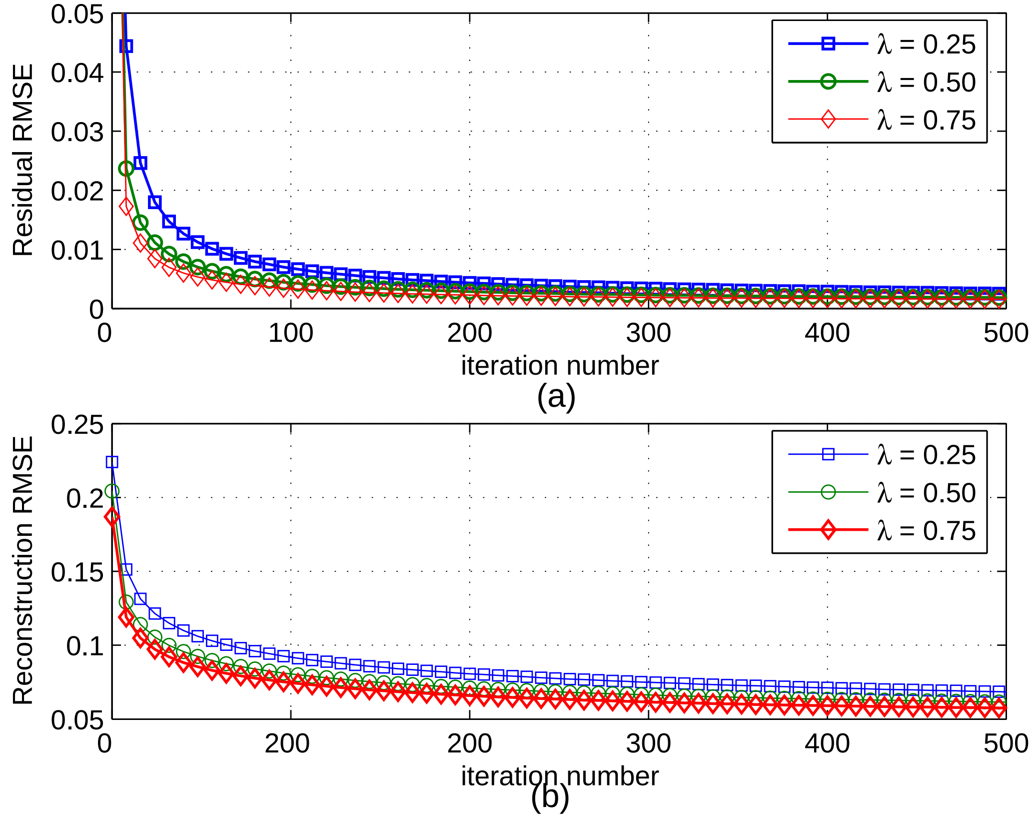


Figure 3.4: (a) Residual RMSE versus iteration number and (b) Reconstruction RMSE versus iteration number with different relaxation parameters for one sample realization of the multi-aperture system. $D = 2$, $K = 4$, $W_c = 1.0$, $SNR = 54dB$

$SNR = 46dB$ with $D = 2$ and $D = 4$ respectively. For the cases in Fig. 3.6(c) and (d), we increased the optical cutoff frequency by reducing the diffraction-limited blur spot and we reduced the SNR by increasing the detector noise level. The RMSE reduces with increasing number of cameras as expected in all three diversity configurations. For any given W_c, D and SNR, the RMSE decreases with increasing diversity. The RMSE for the high-diversity case is smaller than that for the low-diversity case which in turn is smaller than that for the shift-only case. For example, the images in Fig. 3.5 have RMSE values of (a) 0.13, (b) 0.125, and (c) 0.105.

When there is large number of cameras, high-diversity helps to obtain distinct

(a) (b) (c)

Figure 3.5: Reconstructed images at $SNR = 46dB$, $W_c = 1.0$, $D = 4$, and $K = 32$ with (a) shift-only, (b) low-diversity, and (c) high-diversity configurations. These images have RMSE values of 0.13, 0.125, and 0.105 respectively.

views of the scene from each camera. Although the RMSE is almost the same for all three cases with a small number of cameras, it is consistently smaller for low-diversity than for high-diversity. We get more information about the scene by keeping a few cameras closer. We can use this observation for the deployment of cameras in multi-aperture system. We should employ high-diversity when we have many cameras and employ low-diversity when we have fewer. We also note that for the same number of raw measured pixels $\left(\frac{KN}{D^2}\right)$, the reconstruction fidelity is higher for $D = 2$ and $K = 4$ than for $D = 4$ and $K = 16$. A large pixel size introduces more blur, more aliasing, more spectral nulls in the system transfer function and reduces the SNR at high frequencies. Hence, it requires a larger number of cameras to achieve the same reduction in RMSE. Fig. 3.7 shows the reconstructed images at $SNR = 54dB$ and $W_c = 0.25$ with (a) $D = 2$, $K = 4$, RMSE = 0.154 and (b) $D = 4$, $K = 16$, RMSE = 0.186. Table 3.1 summarizes these results in terms

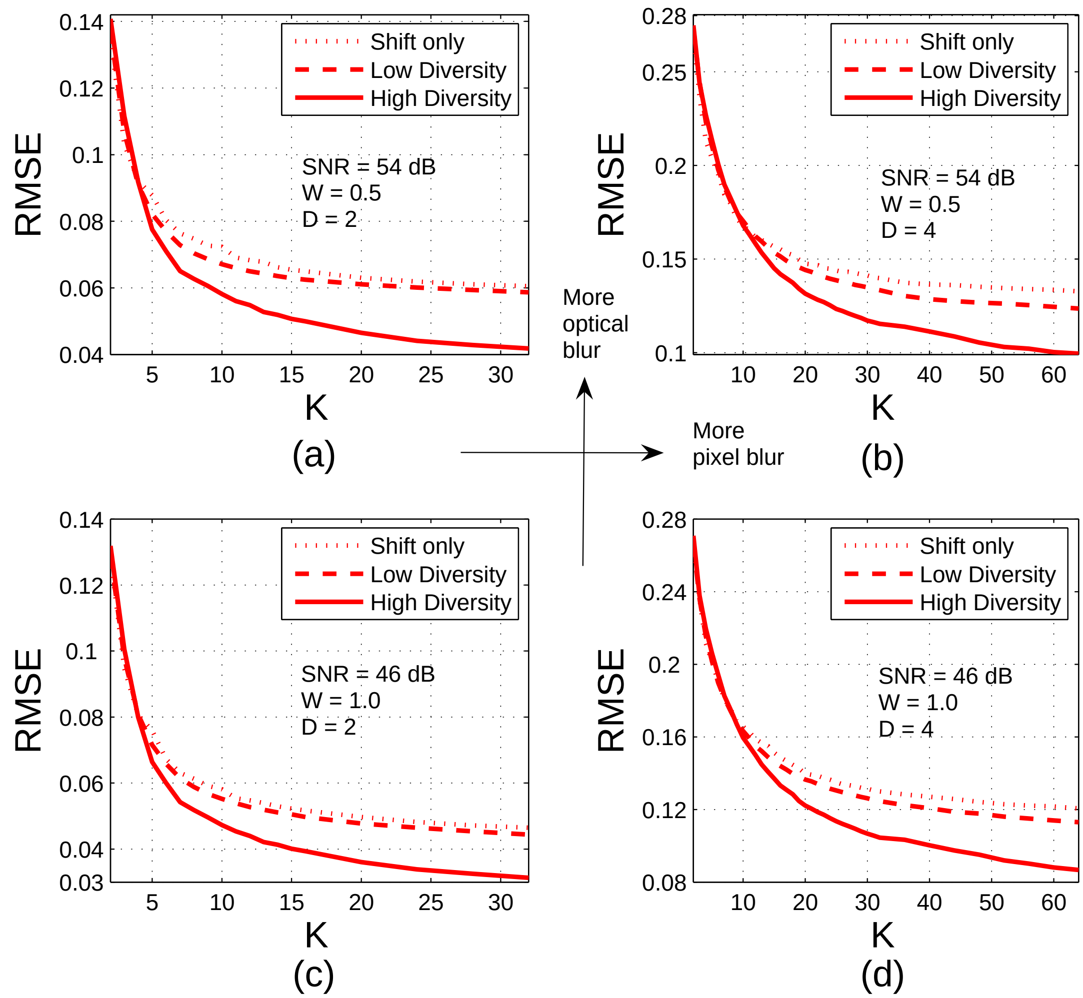


Figure 3.6: Reconstruction RMSE versus number of LR images demonstrates the benefit of measurement diversity for (a) $D = 2$, (b) $D = 4$ at $SNR = 54dB$ and $W_c = 0.5$ and (c) $D = 2$, (d) $D = 4$ at $SNR = 46dB$ and $W_c = 1.0$.

Table 3.1: Percentage RMSE reduction with different W_c, diversity and SNR. For $D = 2$, the reduction is computed from $K = 1$ to $K = 32$ and for $D = 4$, the reduction is computed from $K = 1$ to $K = 64$.

	Percentage Reduction in RMSE			
	$W_c = 0.5$	$SNR = 54dB$	$W_c = 1.0$	$SNR = 46dB$
	$D = 2$	$D = 4$	$D = 2$	$D = 4$
Shift-only	218	141	300	162
Low	228	159	319	180
High	360	222	494	265

of RMSE reduction. The table presents the quantity $\frac{(RMSE_{K_1} - RMSE_{K_2}) * 100\%}{RMSE_{K_2}}$ with $K_1 = 1$ and $K_2 = 32$ cameras in the case of $D = 2$ and with $K_1 = 1$ and $K_2 = 64$ in the case of $D = 4$. We notice a significant RMSE reduction with high-diversity in all four configurations. We see more improvement in reconstruction fidelity for $D = 2$ than for $D = 4$.

Table 3.2 provides an indication about how the optical cutoff frequency (W_c) affects the reconstruction fidelity. Once again we show the percentage improvement in the reconstruction RMSE. We expect that as the optical cutoff frequency is decreased for a given SNR, the reconstruction fidelity reduces as the high-frequency components of the object spectra are not passed through the system. Table 3.2 shows that the improvement in RMSE also reduces with decreasing optical bandwidth. For $SNR = 54dB$ and $D = 2$, the RMSE reduction degrades from 570% at $W_c = 1.0$ to 45% at $W_c = 0.25$. We need larger number of cameras to obtain the same improvement in the reconstruction fidelity for a small W_c. For example, it requires 64 cameras to reduce the RMSE by 222% with $W_c = 0.5$, $SNR = 54dB$, and $D = 4$ as compared to 23 cameras with $W_c = 1.0$.

Now we analyze spectral recovery using multi-aperture imaging with different diversity configurations. Fig. 3.8 shows the spectra of original and reconstructed

(a) (b)

Figure 3.7: Reconstructed images obtained from the same number of measured pixels $[\frac{MN^2}{D^2}]$ (a) $D = 2$, $K = 4$, (b) $D = 4$, $K = 16$ at $SNR = 54dB$ and $W_c = 0.25$.

Table 3.2: Percentage reduction in RMSE with different W_c and diversity at (a) noise-free, (b) $SNR = 54dB$ and (c) $SNR = 46dB$. For $D = 2$, the reduction is computed from $K = 1$ to $K = 16$ and for $D = 4$, the reduction is computed from $K = 1$ to $K = 32$.

	$D = 2$			$D = 4$		
	$W_c = 1.0$	$W_c = 0.5$	$W_c = 0.25$	$W_c = 1.0$	$W_c = 0.5$	$W_c = 0.25$
Shift-only	340	223	36	173	138	70
Low	385	241	41	200	148	73
High	758	392	51	300	209	80

(a)

	$D = 2$			$D = 4$		
	$W_c = 1.0$	$W_c = 0.5$	$W_c = 0.25$	$W_c = 1.0$	$W_c = 0.5$	$W_c = 0.25$
Shift-only	318	197	34	165	129	68
Low	350	209	37	185	140	70
High	570	286	45	266	177	77

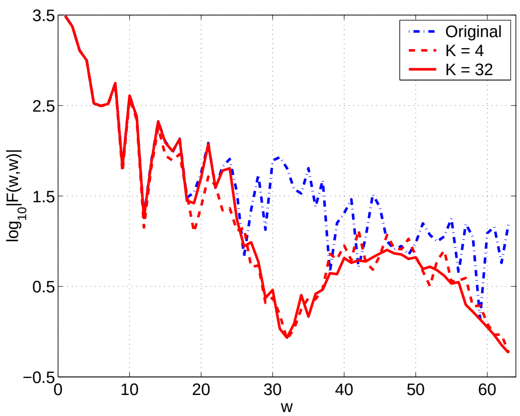
(b)

	$D = 2$			$D = 4$		
	$W_c = 1.0$	$W_c = 0.5$	$W_c = 0.25$	$W_c = 1.0$	$W_c = 0.5$	$W_c = 0.25$
Shift-only	260	154	32	143	109	61
Low	274	161	33	154	113	62
High	293	190	39	203	134	71

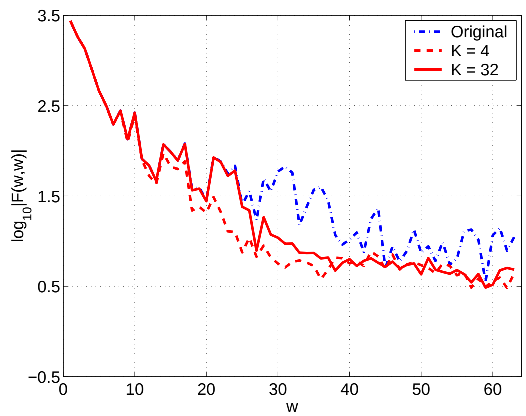
(c)

images with $W_c = 1.0$, $D = 4$, and $SNR = 46dB$ for $K = 4$ and $K = 32$. Each curve is generated using Monte-Carlo simulation and shows only the diagonal spectral components, $|F(w, w)|$, of the two-dimensional Fourier transform, $F(w_1, w_2)$, of the original or the reconstructed images. To compare the reconstruction spectra for different K, we consider only those pixels that are seen by all 32 cameras for the calculation of image spectra. The original spectra vary slightly in three different cases as the common pixels seen by all 32 cameras differ with the diversity configuration. As we see in Fig. 3.8, the reconstructed spectrum with $K = 32$ resembles more closely the original spectrum than with $K = 4$. The reconstructed spectral components match exactly the original spectrum at low frequencies. We see that increasing the number of cameras causes the reconstructed spectra to match out to higher frequencies. We recall that multiple measurements with magnification diversity can be used to recover frequency components around the spectral nulls. Since there is no magnification diversity for the shift-only case, the SNR of spectral components near the nulls are very low as seen in Fig. 3.8(a). It is noticeable in Fig. 3.8(b) and (c) that the SNR at those frequencies are boosted up in both low- and high-diversity cases. To compare the three different spectra, we plot them together in Fig. 3.9 for $SNR = 60dB$ and $K = 32$. Since the common pixels that are seen by all 32 cameras are different in each case, we consider only those pixels that are common in all three case for the estimation of spectra. As we can see the reconstructed spectrum for the high-diversity case best matches the original.

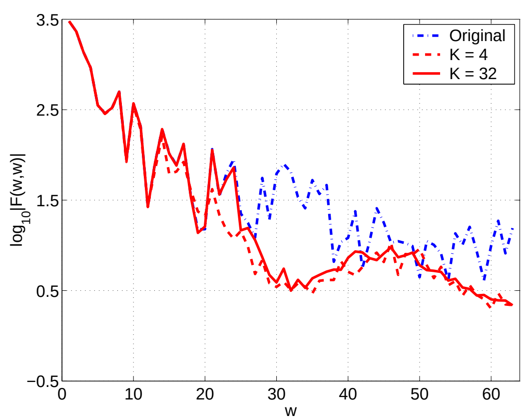
In a multi-aperture system we can mitigate any amount of additive noise by using a sufficiently large number of cameras. The reconstruction RMSE due to the additive noise alone would reduce as $\frac{1}{\sqrt{K}}$. In Fig. 3.10, we plot reconstruction RMSE versus the number of cameras at different values of SNR for (a) low-diversity and



(a)



(b)



(c)

Figure 3.8: Comparison of reconstructed spectra with $D = 4$, $W_c = 1.0$, $SNR = 46dB$ for (a) shift-only, (b) low-diversity, and (c) high-diversity cases. The reconstructed spectra in high-diversity case are closer to the original spectrum. Also notice that the SNR at frequencies around spectral nulls are higher in low- and high-diversity cases than shift-only case.

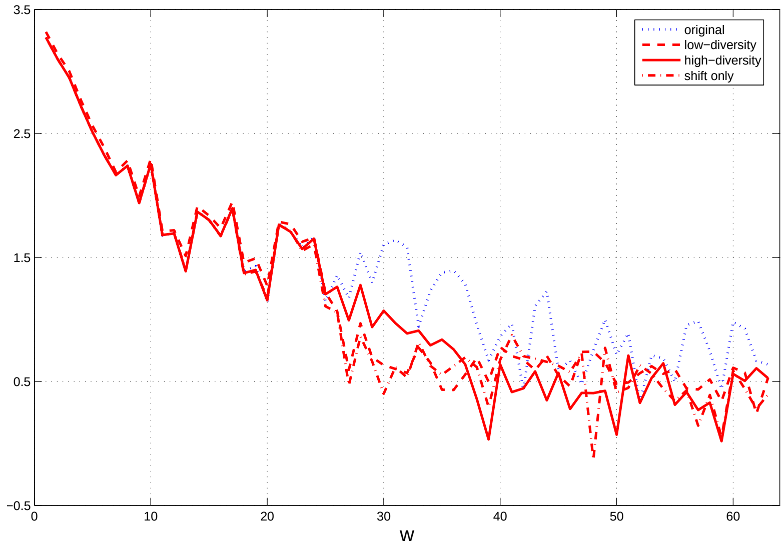


Figure 3.9: Comparison of reconstructed spectra for $K = 32$, $D = 4$, $W_c = 1.0$, $SNR = 60dB$ for different diversity configurations. Spectra are obtained from only those pixels that are seen by all the cameras and in all configurations.

(b) hi-diversity cases. The difference between the RMSE curves for the noise-free (square symbols) and 26dB (dimond symbol) cases reduces with increasing number of cameras in both cases. Fig. 3.11 shows a series of reconstructed images at high-diversity configuration, $SNR = 54dB$, $D = 4$, $W_c = 0.5$ with (a) $K = 2$, (b) $K = 4$, (c) $K = 8$, and (d) $K = 32$. The visual quality of images improves with increasing K. The image in 3.11(d) looks less noisy and the edges are restored better than in the rest of the images. These reconstructed images have RMSE values of 0.28, 0.23, 0.18, and 0.12 respectively.

The error in the reconstruction when using a larger number of cameras is mainly due to blurring in the optical system, the residual pixel blur and the interpolation step in the reconstruction. This hints to another diversity that we can exploit, namely, pixel fill-factor. Because of the spatial integration of the incident intensity on the active detection area, a large pixel introduces more blur. By reducing the

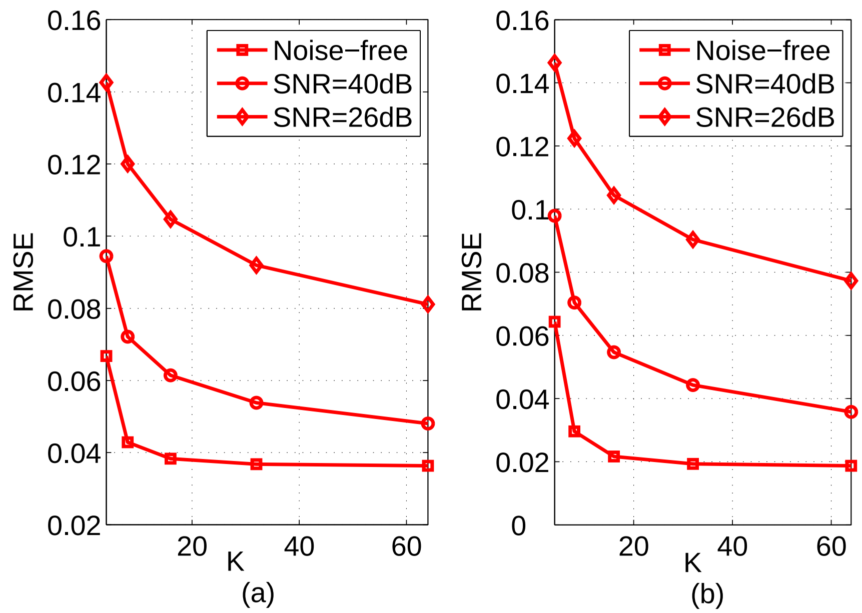


Figure 3.10: Reconstruction RMSE versus number of LR images at different SNR for $D = 2$ with $W_c = 1.0$ at (a) low and (b) high diversity.

pixel fill-factor, we can reduce the amount of pixel blur. However a small pixel collects fewer photons and in the presence of additive noise, this reduces the SNR of the measurement. Hence, at small fill-factor the additive noise dominates the reconstruction RMSE, whereas, at large fill-factor mainly the pixel blur contributes to the RMSE. There is an optimum value of fill-factor that trades off pixel blur and SNR. We plot the RMSE versus the linear pixel fill-factor in Fig. 3.12 for the high-diversity configuration with $D = 8$, $W_c = 0.25$, and $K = 32$ for different SNR values. The linear fill-factor of a square detector is defined as the ratio of the active pixel width to the inter-pixel spacing. For the examples shown, the optimum values of fill-factor are 30%, 60%, 60%, and 60% for SNR values of 60dB, 54dB, 46dB, and 40dB respectively.

All previous results have been obtained under the assumption that the imaging channel parameters, $\{s_k,\ \theta_k,\ a_k,\ b_k : k = 1, ..., K\}$, are known precisely. There are

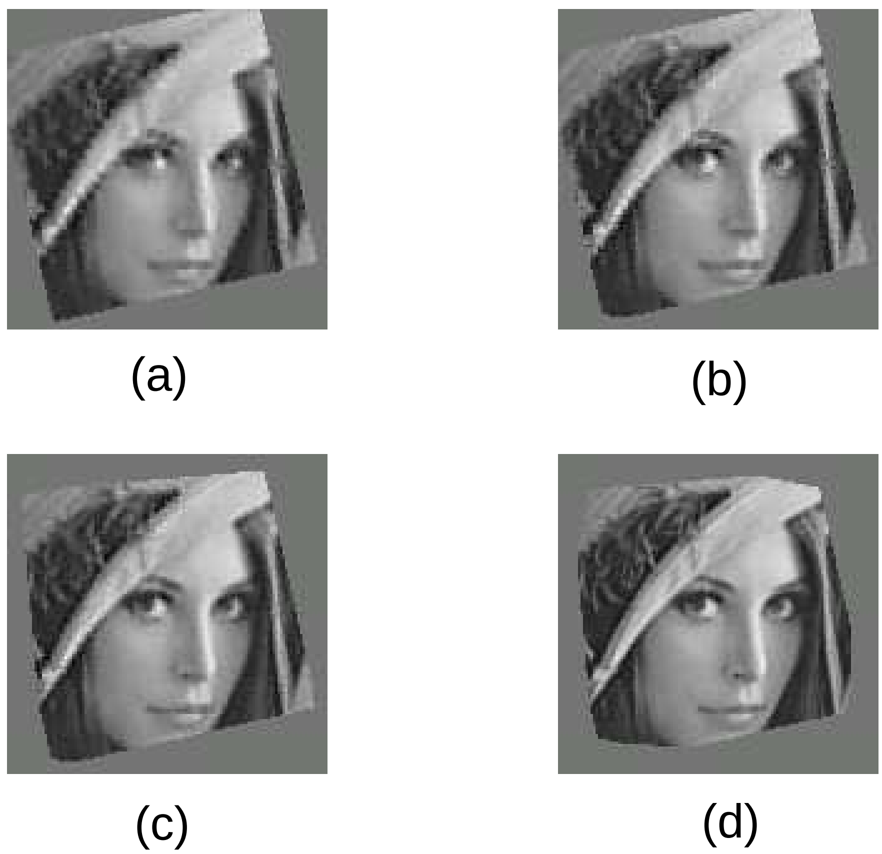


Figure 3.11: Reconstructed images at $SNR = 54dB$, $W_c = 0.5$ and $D = 4$ with (a) $K = 2$, (b) $K = 4$, (c) $K = 8$, and (d) $K = 32$.

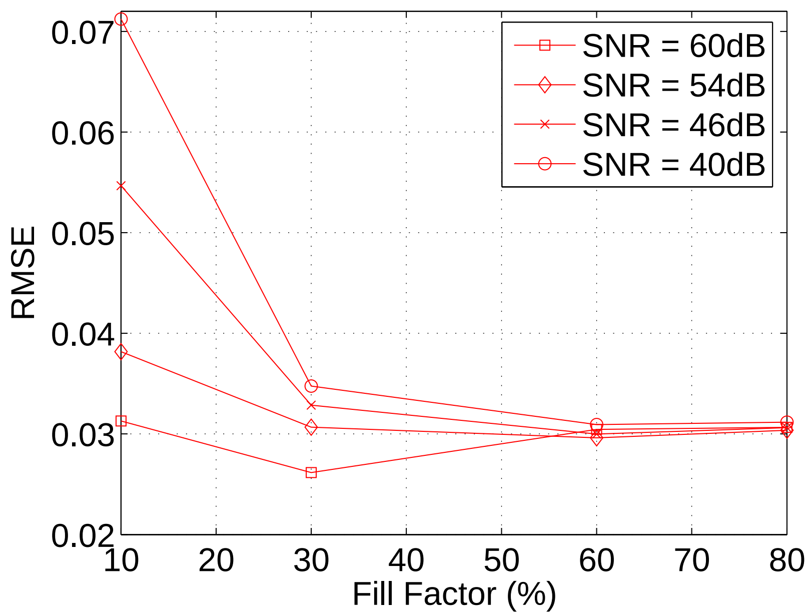


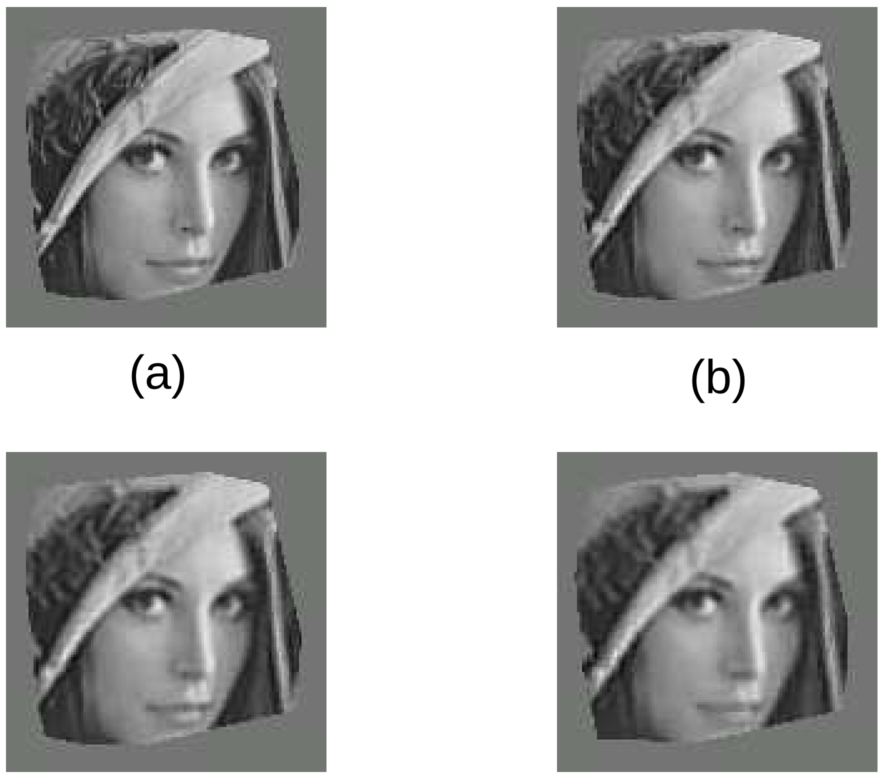
Figure 3.12: Reconstruction RMSE versus fill-factor for $D = 8$, $W_c = 0.25$ and $K = 32$ for different SNR.

many reasons why this assumption may not be valid in a practical situation (e.g., poor calibration, environmental fluctuations, etc.). We usually have only a coarse knowledge of the channel parameters, and in some cases we may need to estimate these parameters. The success of the SR reconstruction depends crucially on the accuracy of the channel parameters. We can expect a degradation in the image reconstruction quality when these estimates are not accurate. A few example reconstructed images are shown in Fig. 3.13 for $D = 4$, $W_c = 1.0$, and $K = 32$ at $SNR = 54dB$. The image in (a) is reconstructed using exact knowledge of the channel parameters and has RMSE = 0.086. The images shown in Fig. 3.13(b), (c), and (d) are reconstructed in the presence of imperfect channel knowledge. Specifically, we assume that the channel parameter errors are uniformly distributed random variables with standard deviations of $\pm 1\%$, $\pm 2\%$, and $\pm 5\%$ of their actual values respectively. These three reconstructions have RMSE of 0.206, 0.207, and 0.213 respectively. Because the pixels from different low-resolution images are not combined properly, the reconstructed images in Fig. 3.13 look smeared and lack details. We plot the reconstruction RMSE versus the channel parameters error in Fig. 3.14 for different numbers of cameras. Fig. 3.14(a) depicts the case of $D = 2$ and $K = 4$, 8, 12, and 16 whereas 3.14(b) depicts the case of $D = 4$ and $K = 8$, 16, 24, and 32. For the case of $D = 4$, we see that the reconstruction error increases by more than 2X for only $\pm 1\%$ error in the channel parameters. To make a multi-aperture system robust, it is very important to have good tolerance to variations in the channel parameters. This calls for a joint-estimation of the channel and HR image when the exact channel parameters are unknown. The activity is an element of our ongoing work.

3.4 Experiment

In this section, we present the sensor platform used to investigate multi-aperture imaging. The system design will be described and practical issues involved in implementing the algorithm will be discussed. Enhanced resolution images generated using a data set of low resolution images will be shown. Finally, we will propose improvements that will increase system performance.

The sensor platform is composed of a camera array connected to a workstation running the IBP algorithm. A front view of the camera array is shown in Fig. 3.15. The camera array supports 129 OmniVision OV6120 monochrome CMOS image sensors. (Note: One image sensor provides the system master clock synchronization and its image data is typically neglected.) Each sensor acquires a 352 x 288 pixel image with rectangular pixels measuring 9.0 by 8.2 microns. A 6.5mm focal length multi-element glass lens with an $F/\#$ of 2 is permanently attached to the coverglass



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

Figure 3.13: Reconstructed images with $D = 4$, $W_c = 1.0$, $K = 32$, $SNR = 54dB$ with (a) 0, (b) $\pm 1\%$, (c) $\pm 3\%$, and (d) $\pm 5\%$ error in the channel parameters.

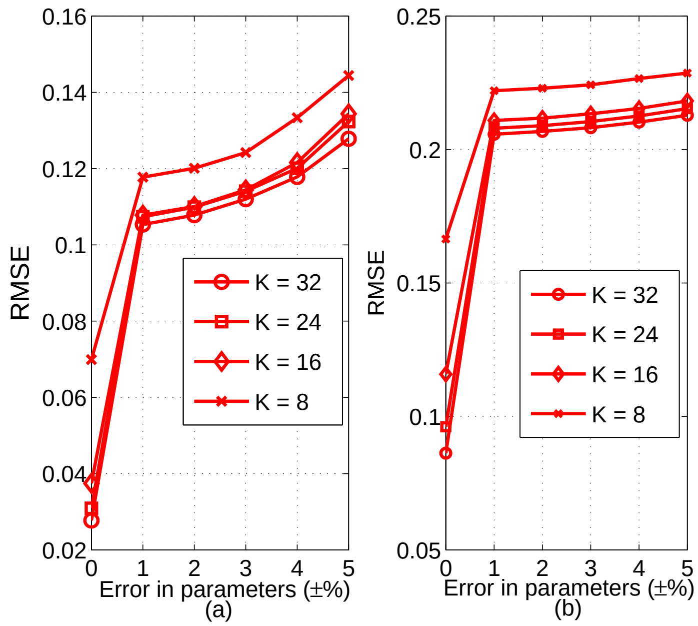


Figure 3.14: Reconstruction RMSE versus error in the channel parameters at $SNR = 54dB$, $W_c = 1.0$, and (a) $D = 2$, and (b) $D = 4$.

of 56 sensors. The lenses are permanently focused to a surface at a distance of 50 feet and the image centers are collectively aligned to a common point at infinite range. The remainder of the sensors will be outfitted with suitable focal length lenses pending the results of this initial investigation.

Custom designed electronics boards were built to operate the sensors. Four sensors are attached to each camera module and a unique rotation provided to each sensor's bonding area. By using two board designs and staggering the orientation of the boards in the mechanical framework, a large diversity of sensor rotations was achieved. Each sensor is digitally accessible and provides 8-bit monochrome raw image data. Control and data signals are routed between a set of eight sensor boards and a central PCI interface using flexible multiwire cable. The PCI interface services four cables and provides a direct connection to a PC104 computer. The

PC104 computer is responsible for controlling sensor array operation, acquiring image data sets, and communicating with the remote workstation via ethernet. The remote workstation provides the multi-aperture image combining and demonstration interface.

The sensor modules are attached to a welded Invar framework measuring 19-1/2 inches square. Invar was chosen for both mechanical and thermal stability in order to maintain long term subpixel registration of the images. The 32 camera modules are arranged on a six by six grid where the central four sites are allocated to the PCI interface and PC104 computer.

Although the camera array could have been assembled using commercially available cameras, there were advantages to developing a custom system. By doing so we were assured that the image data did not include compression artifacts, that image acquisition could be accurately synchronized, and that the sensors could be operated consistently with one another. Finally, although USB/Firewire bus specifications indicate large connection numbers were feasible, it was not apparent that such a dense interconnection scheme could be managed successfully.

Building a system implementing the multi-aperture imaging algorithm tests the practicality for working with real-world data. It also involves resolving issues not anticipated through modeling. For example, the lateral displacement of the apertures leads to range dependent parallax effects. The sensitivity to this parallax increases at short ranges and as reconstruction resolution increases. Also, the process of establishing subpixel registration of the images requires development of techniques exceeding simple mechanical placement of sensors. Finally, operating variations of the individual sensors and optics lead to additional complexities.

The most critical aspect of operating the multi-aperture imaging system is establishing subpixel registration of the sensor images. Although careful attention to mechanical alignment during lens attachment produced overall pixel-scale positioning uniformity, it is necessary to supplement this technique with additional image analysis. The difficulty of this task is further compounded by the lateral separation of the cameras which produces range dependent parallax effects.

The parameters describing the sensor geometry are:

- PDX and PDY give the lateral translation of the image center point relative to a reference point located at infinite range.
- THETA describes the rotation of the sensor.
- MAG gives the relative magnification to match a reference field of view.
- DX and DY are the position of the sensor relative to the central axis of the camera array and are necessary for determining range dependent variations. Coarse values of these parameters are determined from the designs.
- Z0 gives the operating range of the camera array for the data set.

The rectangular nature of the sensor pixels must also be incorporated into the analysis, however, optical distortion is currently neglected. It should be noted that the observed distortion in this system will interfere with the quality of the image reconstruction as the analysis region includes more pixels toward the extreme edges of the images.

The current process for calibrating camera geometry parameters is to maximize the cross-correlation of each sensor image against a reference image. A collection of images are acquired at a specific range. An automated program extracts a region of interest, resamples the region to a data array 16 times larger along each dimension using the design values for magnification and rotation. The test image is cross-correlated against the reference image using Fourier techniques to determine the optimal lateral translations. Magnification and rotation are further refined by exhaustive search over a grid of values near the design values.

In the sensor platform, the multiple-aperture software was modified to acquire images via the camera array rather than using the simulated data sets. The code was further streamlined to provide rapid resolution enhancement in order to demonstrate the technique in near real time. One portion of this performance improvement is achieved by restricting the evaluation region. The enhanced region or regions is then superimposed on one sensor image to produce a foveated image. This technique is shown in Fig. 3.16 where the two regions of enhanced resolution are outlined. The artifact in the bottom region is due to incomplete overlap of the image sets.

A demonstration of enhanced resolution images generated using the camera array data is illustrated in Fig. 3.17 through 3.19. Fig. 3.17 shows the standard 352x288 image acquired by one of the sensors of a target at a range of 17-1/2 feet. Fig. 3.18 shows a closeup of the resolution chart extracted from the previous image. An enhanced image generated by the multi-aperture algorithm is shown in Fig. 3.19. Forty sensor images were used to produce this image having four times as many pixels along each dimension compared to the original region. It is evident from the images that the multi-aperture platform is able to resolve finer detail than an individual sensor.

From our initial experience with the multiple-aperture sensor platform, we have identified aspects that require additional attention to optimize performance. The more notable recommendations are: improving camera geometry calibration techniques, compensating for optical distortion, shrinking the platform scale to reduce range dependencies, and integrating image resampling and transformation into higher performance processors or FPGAs.

Since registering images to subpixel accuracy is crucial to generating enhanced resolution images, developing a calibration technique for determining the camera array geometry is paramount. An automated calibration procedure, potentially employing a variation of the algorithm matching acquired images against a well determined reference target, is necessary as the number of sensors increases. The misregistration of images becomes more evident at close ranges (less than 20 feet in this platform) where the overlapping analysis regions differ between sensors. Most of the contribution is likely due to optical distortion, which we have estimated to be as large as one or two pixels at the extreme edge of the image. Thus, incorporating additional features of the optics into the image transformation stage would improve registration uniformity across the image.

Camera parallax produces range dependent image displacement affecting the subpixel registration required by the algorithm. The algorithm is least sensitive at long range operation, and most sensitive at close ranges or when non-planar targets are used. (Alternatively, this range dependent effect could be highly valuable if the platform were adapted to interpret three dimensional information from the scene.) One means of reducing the effect in the current system is to use a subset of closely located cameras. A redesigned system would substantially benefit by reducing the 3 inch sensor spacing to about 1 inch.

Although the current system requires a few seconds to generate an enhanced resolution image, we are able to envision video rate operation. Strategies for increased performance rate include streamlining the algorithm code and potentially implementing FPGA processing for image alignment and back-projection, and reducing

Figure 3.15: Front view of the camera array. Each square holds an electronics board supporting four sensors. Fifty-six of the 129 sensors have lenses attached. The central area is devoted to interface electronics and a PC104 computer.

data memory requirement by selectively acquiring images from camera batches as needed.

3.5 Conclusion

In this work we have analyzed the effects of channel diversity on the reconstruction of a high-resolution scene from multiple shifted, rotated and magnified images using a nonlinear iterative back projection method. Imaging channels with different amounts of diversity are considered. It was found that a high-diversity is favorable in the system with a large number of cameras. A large variation in channel parameters help to improve the SNR at frequencies near spectral nulls within the optical passband, thus improving the image reconstruction. The reconstruction RMSE can be reduced by 570% and 266% using 16 and 32 cameras with $W_c = 1.0$,

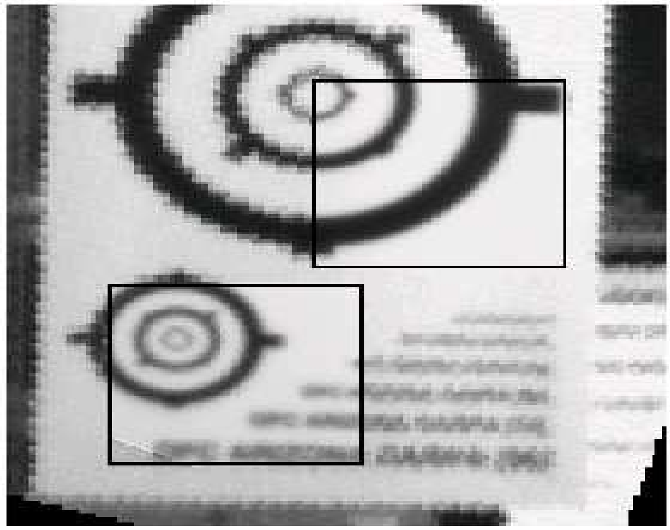


Figure 3.16: Demonstration of foveation showing two enhanced resolution regions.

Figure 3.17: Full image from single sensor. Note multi-line chart to the left of the circular target.

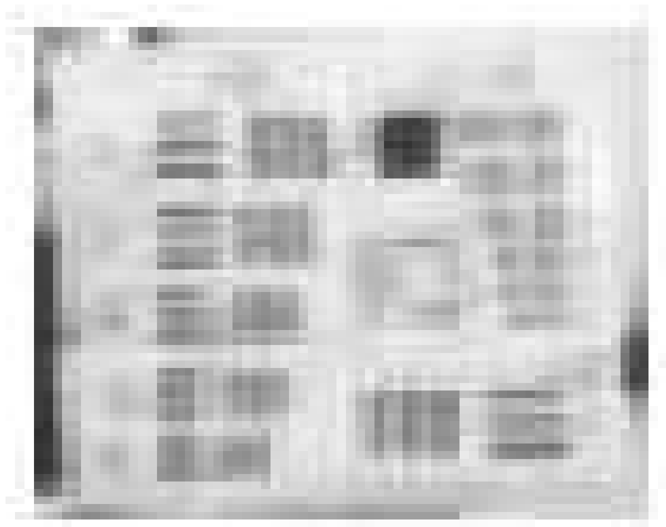


Figure 3.18: Individual sensor with closeup of the resolution chart.

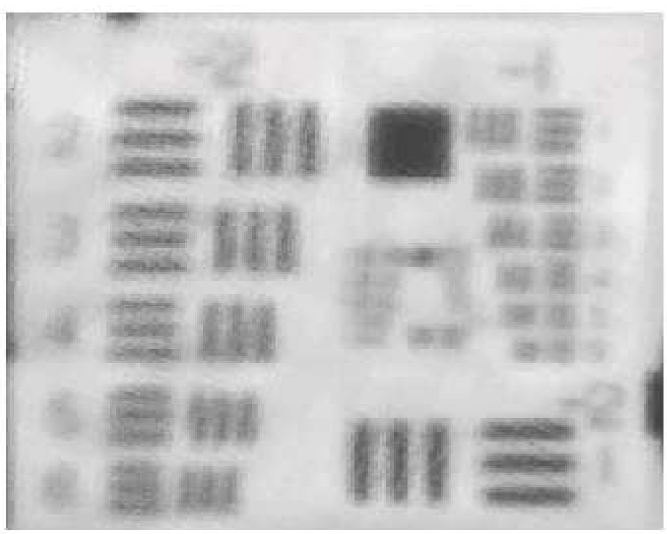


Figure 3.19: Estimate formed by multi-aperture imaging showing closeup of resolution chart.

$SNR = 54dB$ for $D = 2$ and $D = 4$ respectively. Any amount of additive noise in pixels can be eliminated by employing a sufficiently large number of imagers. The difference in RMSE between noise-free and 26dB case reduces by 33% from 4 to 64 cameras for $D = 2$ and $W_c = 1.0$. We showed the existence of a trade-off between the pixel size and the pixel blur, and an optimal value of pixel fill-factor can be found by balancing the additive noise and the blur. The optimum values of fill-factors vary from 30% to 60% for different SNR values considered for $D = 8$ and $W_c = 0.25$. The iterative algorithm is very sensitive to the accuracy of the channel parameters. Any inaccuracy in these parameters results in poorly reconstructed images. As an example, the reconstruction error increases by 69% and 105% at $\pm 1\%$ and $\pm 5\%$ channel parameter errors respectively for $D = 2$, $K = 4$, $SNR = 54dB$ and $W_c = 1.0$. The multi-aperture algorithm implemented on our sensor platform demonstrates the qualitative enhancement in resolution as produced in the simulation results. Although calibrating the sensor geometry, eliminating parallax effects, as well as simultaneously operating a large camera set and managing a large data stream are considerable challenges; we expect that the camera performance will steadily improve to match the simulation results. The effort is under development to simultaneously estimate the channel parameters and reconstruct a high-resolution image.

CHAPTER 4

Multiframe Superresolution of Binary Images

4.1 Introduction

It is well-known that any available prior knowledge about the object can be used to regularize the restoration process. Examples of prior knowledge include non-negative object pixels and finite object support. The multiframe Wiener filter[20] uses knowledge of the object power spectrum to restore the HR image from multiple shifted LR images. Molina et al. [50] modeled the HR image using a Markov chain in their restoration methods. Laplacian [25] or Gibbs [26] prior models can also be used to obtain a maximum a-posterior (MAP) estimate of the HR image. Baker et.al.[33] employed a recognition-based prior to identify relevant features in the LR images to improve the reconstruction.

Another form of object prior knowledge is the finite alphabet constraint. Given a N-pixel digital image that uses Q-levels for each pixel, only Q^N images are possible. The special case of binary-valued imagery ($Q = 2$) is important. Each pixel in a binary object (e.g., text, barcode, etc) is represented by either zero or one, and there are only 2^N possible such images. Communication theoretic superresolution methods like the Viterbi algorithm have been applied to restore a 2D binary object from a single blurry and noisy image [51, 52]. These algorithms exploit the fact that the channel response function is local in nature. A reduced-complexity parallel algorithm known as two-dimensional distributed data detection (2D4) [53] has also been used to restore single frame binary imagery. This technique is used in conjunction with complexity reduction procedures to handle a larger blur support and for restoring gray-scale images. We propose a scheme that employs 2D4 to restore a HR binary image from multiple blurry, under-sampled and noisy LR images. We also introduce a novel complexity reduction technique for imaging systems with large channel support. This chapter is organized as follows: in section 4.2 we present the multiframe imaging system and model the degradations that occur to a HR object or scene. We discuss two "traditional" multiframe image restoration methods: linear minimum mean square error (LMMSE) estimation and IBP. Section 4.3 deals with the 2D4 algorithm and complexity reduction techniques. In section 4.4 we present simulation results for the algorithms in different imaging systems, and analyze their computational complexities We conclude the chapter in section 4.5.

4.2 Multiframe Imaging

4.2.1 Imaging model

The k^{th} imaging channel can be expressed by a linear system of equation as given in Eq. 2.2. The degradations (imaging operators) in a single imaging channel are presented in Fig. 4.1 again for convenience. We can express Eq. 2.2 more concisely as

$$y_k = H_k f + n_k \tag{4.1}$$

where $H_k = SH_{ccd}H_{opt}T_k$. The set of measurement vectors, noise vectors, and the system matrices can be stacked to obtain the multiframe imaging equation:

$$y = Hf + n \tag{4.2}$$

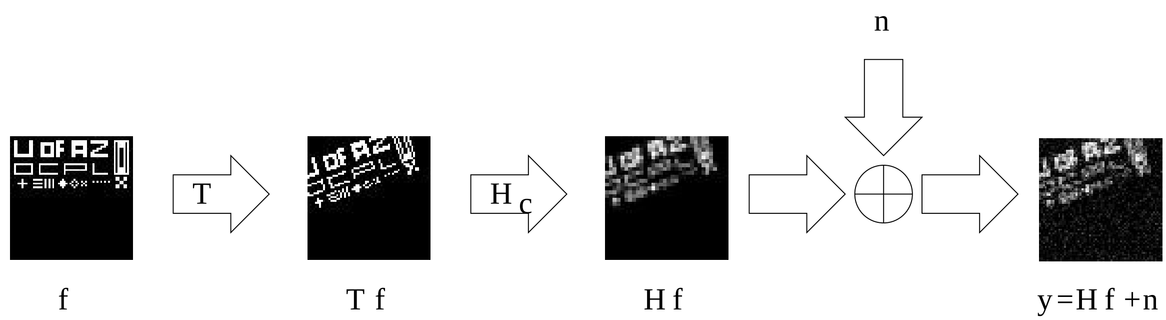


Figure 4.1: Schematic diagram of an imaging channel showing degradations in a typical multiframe imaging system. $H_c = SH_{ccd}H_{opt}$ and $H = H_cT$.

where $y = [y_0^T y_1^T ... y_{K-1}^T]^T$ and $n = [n_0^T n_1^T ... n_{K-1}^T]^T$ are vectors of length $P = KM$, and $H = [H_0^T H_1^T ... H_{K-1}^T]^T$ is the overall system matrix of size $P \times N$. Superresolution techniques offer solutions to this multiframe image restoration problem, and estimate/recover the HR object, f.

4.2.2 Wiener and IBP

The Wiener operator [20] is the LMMSE solution for estimating f from the measurements described in Eq.4.2. The Wiener operator, W_f, is defined in the spatial domain as

$$W_f = R_f H^T [H R_f H^T + \sigma^2 I]^{-1} \tag{4.3}$$

where the $N \times N$ matrix R_f is the autocorrelation function of the object, and I is the identity matrix of size $P \times P$. R_f can be calculated analytically (e.g., for random binary images) or obtained by training on several images from the relevant object class [54]. The LMMSE estimate of the object is given by $\widehat{f}_{LMMSE} = W_f y$. The Wiener is the optimal linear solution in the mean square error sense. It uses object prior knowledge in the form of second order statistics.

The Wiener operator provides a single-step estimator for f: IBP in contrast is an iterative algorithm. Additional object prior information (e.g., nonnegativity of object pixels) can be easily incorporated within such an iterative framework. IBP starts with an initial estimate $\widehat{f_0}$, and updates it at the $(j+1)^{th}$ iteration by back-projecting the difference between the actual measurements and the predicted measurements at the j^{th} iteration. The estimate of the object at the $(j+1)^{th}$ iteration can be expressed as

$$\widehat{f_{j+1}} = \widehat{f_j} + \alpha B[y - H\widehat{f_j}] \tag{4.4}$$

where $\widehat{f_j}$ is the estimate of the object at the j^{th} iteration, B is the back-projection operator, and $0 \leq \alpha \leq 1$ is the relaxation parameter. There are several choices for the back projection operator including the adjoint of H and the pseudo-inverse of H.[4] We use the Wiener operator $B = W_f$ in our implementation of IBP to regularize the reconstruction procedure using the object autocorrelation prior knowledge. The non-negativity constraint is also enforced on the estimate at each iteration.

4.3 2D4

The 2D4 algorithm was first proposed by Chen et al. [55] for removing the inter-symbol interference (ISI) in holographic memories. Later it was employed for restoration of an object from a blurred and noisy image in a shift invariant imaging system [53]. Here we describe how 2D4 can be employed to restore a binary object from degraded measurements in a more general linear shift-variant system.

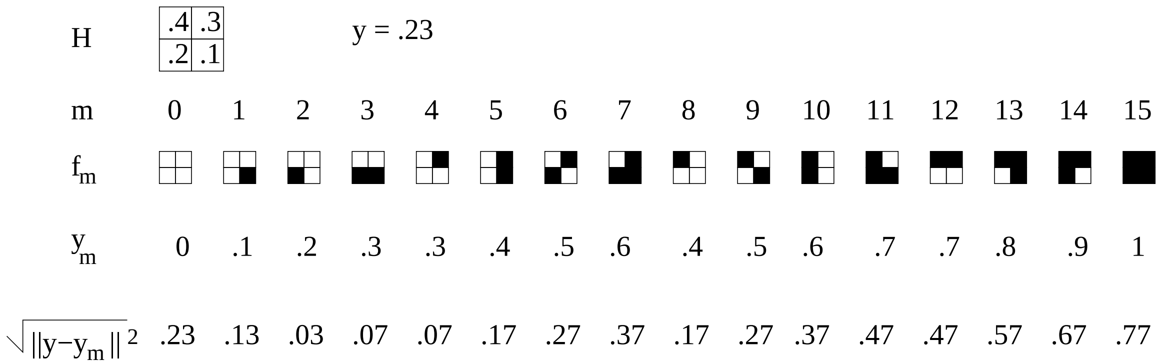


Figure 4.2: An example describing maximum likelihood estimation, and complexity reduction. Noise standard deviation $\sigma = .02$

4.3.1 The algorithm

One can envision a conceptually simple scheme to find the maximum likelihood (ML) estimate of a binary object via exhaustive search [56]. We start by listing all $f_m : m = 0, ..., 2^N - 1$ possible binary candidate objects of N pixels. The noise-free output of the imaging channel is calculated for each of these possible objects as $y_m = Hf_m : m = 0, ..., 2^N - 1$. Because the noise is AWGN, the ML object estimate corresponds to the y_m that minimizes $\|y - y_m\|^2$. In Fig. 4.2, for example, $N = 4$ and there are 16 possible candidate objects. The predicted measurements, and their distances to the measured value are listed below the object patterns. The predicted measurement from the candidate object $m = 2$ is closest to the measured value and hence is the ML estimate. However this conceptually simple procedure to find the ML estimate is computationally prohibitive for images of any practical size.

We will develop the 2D4 algorithm by first considering measurements from a single LR image, $y(i) : i = 0, ..., M - 1$. We remove the subscript k for readability of equations. From Eq. 4.1, the i^{th} measurement pixel can be expressed as

$$y(i) = \sum_{j \epsilon \ell_i} H(i,j) f(j) + n(i) \tag{4.5}$$

where ℓ_i is the set of object pixel locations that contribute to the i^{th} measurement pixel. It defines the shape and size, $L_i = |\ell_i|$, of the point response function (PRF) of the i^{th} detector [21]. The PRF of the i^{th} detector is obtained from the i^{th} row of H as $h_i = \{H(i,j), j \in \ell_i\}$.

Because the object pixels are binary-valued, we are interested in their likelihoods of being 0 or 1, $\{p_l(0), p_l(1) : l = 1, ..., N\}$. The measurement $y(i)$ could have resulted from any one of the 2^{L_i} binary local patterns at object pixel locations, $l \in \ell_i$. We denote these patterns by $\{P_{ij}(l) : j = 0, ..., 2^{L_i} - 1, l \in \ell_i\}$ where l indexes the object pixel locations that contribute to the i^{th} measurement. The j^{th} pattern $P_{ij}(l)$ contains the binary representation of j in the L_i-pixel locations given by $l \in \ell_i$. The likelihood of the i^{th} measurement conditioned on the j^{th} pattern can be calculated for AWGN as

$$C_{ij} \equiv p(y_i|P_{ij}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{[\sum_{l \in \ell_i} H(i,l) P_{ij}(l) - y(i)]^2}{2\sigma^2}} \tag{4.6}$$

where σ is the standard deviation of noise. The likelihood of the i^{th} measurement conditioned on the l^{th} object pixel being zero, $f(l) = 0$, $l \epsilon \ell_i$, is given by

$$p(y_i|f(l) = 0) = \sum_{\substack{j=0 \\ P_{ij}(l)=0}}^{2^{L_i}-1} C_{ij} \prod_{\substack{m \epsilon \ell_i \\ m \neq l}} p_m(P_{ij}(m)) \tag{4.7}$$

The summation is over only those object patterns that have a 0 in the l^{th} location. $P_{ij}(m)$ is either 0 or 1, and $p_m(P_{ij}(m))$ is the likelihood of the m^{th} object pixel being $P_{ij}(m)$. Here we borrow the idea of "extrinsic information" from the turbo decoding literature [57], and replace $p(y_i|f(l)=0)$ with $p_l(0)$ which yields

$$p_l(0) = \sum_{\substack{j=0 \\ P_{ij}(l)=0}}^{2^{L_i}-1} C_{ij} \prod_{\substack{m \epsilon \ell_i \\ m \neq l}} p_m(P_{ij}(m)) \tag{4.8}$$

In this equation the likelihood of an object pixel is seen to depend on the measured values (via C_{ij}) and the likelihoods of neighboring pixels (via $p_m(P_{ij}(m))$). Let I_l represent the set of measurements that are affected by the l^{th} object pixel. Then we can improve the estimation of the likelihoods of the l^{th} pixel by considering all measurements in the set I_l. We can define an iterative algorithm to exchange the pixel likelihoods among neighboring pixels in the restored image as follows.

$$pu_l^{q+1}(0) = \sum_{i \epsilon I_l} \sum_{\substack{j=0 \\ P_{ij}(l)=0}}^{2^{L_i}-1} C_{ij} \prod_{\substack{m \epsilon \ell_i \\ m \neq l}} p_m^q(P_{ij}(m)) \tag{4.9}$$

$$p_l^{q+1}(0) = (1-\beta)p_l^q(0) + \beta pu_l^{q+1}(0) \tag{4.10}$$

where q is the iteration index. $p_l^{q+1}(0)$ is the estimate of the likelihood that the l^{th} pixel takes the value zero, and $pu_l^{q+1}(0)$ is the magnitude of the update at iteration $q+1$. The likelihood $p_l^{q+1}(1)$ of the l^{th} pixel taking the value one and the magnitude of update, $pu_l^{q+1}(1)$, are calculated at iteration $q+1$ using a similar set of equations. This 2D4 algorithm is a two-step procedure at each iteration. First the likelihood updates for each pixel in the object are calculated based on the measurements and

.0001	.0026	.0058	.0026	.0001
.0026	.0537	.1190	.0537	.0026
.0058	.1190	.2641	.1190	.0058
.0026	.0537	.1190	.0537	.0026
.0001	.0026	.0058	.0026	.0001

(a)

.0562	.1246	.0562
.1246	.2765	.1246
.0562	.1246	.0562

(b)

Figure 4.3: Channel truncation approach to reduce the complexity. (a) Original coefficients and (b) truncated coefficients.

current values of the object pixel likelihoods. In the next step, a certain fraction ($0 < \beta < 1$) of the update is added to the current values of all pixel likelihoods. It is apparent from equations 4.9 and 4.10 that the likelihood of the l^{th} object pixel is affected by (1) all the measurements to which it contributes via its PRF, and (2) the likelihoods of all the neighboring pixels that determine those measurements. Owing to this exchange of likelihoods among neighboring pixels, the likelihood of each object pixel propagates throughout the object. This mechanism is analogous to belief propagation or message passing [58, 59].

4.3.2 Complexity reduction

The complexity of the 2D4 algorithm increases exponentially with the PRF support. If $L_i = 16$, there are $65,536$ local object patterns to be considered in order to update the pixel likelihoods based on the i^{th} measurement. The algorithm can therefore only be used for small blur support (e.g., 3×3 pixels for which $L_i = 9$). For imaging systems with larger channel support, 2D4 can be used in conjunction with complexity reduction techniques. Two such methods have been previously explored. The first involves truncating the PRF by setting the smallest coefficients to zero and renormalizing the approximate PRF. For example in Fig. 4.3 the original 5×5 PRF is truncated to obtain an approximate 3×3 PRF and then renormalized so that the coefficients sum to 1. The second approach to complexity reduction exploits the correlation structure among object pixels. Some of the 2^{L_i} possible local object patterns may never appear in a particular class of binary objects. A vector quantization technique has been employed in Ref. [53] to obtain a codebook, C, of typical local object patterns. The 2D4 algorithm then updates pixel likelihoods based on the $|C|$ patterns instead of 2^{L_i}.

The first method introduces errors in the calculation of conditional likelihoods expressed in Eq. 4.6. We can expect significant errors if the sum of all deleted coefficients is comparable to the standard deviation of noise. Because of the training process in the second method, the derived codebook becomes specific to a class of object. This may introduce errors due to the quantization. The codebook patterns are real-valued in general, and interpolation is required to obtain the likelihoods of binary pixels.

We present a new method for reducing the computational complexity of the algorithm. Our method involves adaptively selecting a subset of local object pixel patterns for each measurement based solely on the measured value and the noise statistics. If a measured value, $y(i)$, is close to zero at high signal-to-noise ratio (SNR), it is improbable that a local object pattern with large weight (i.e., containing many 1's) generated it. We define SNR as $SNR(dB) = 10 \ log_{10} \frac{.25}{\sigma^2}$. We adapted this definition as a random $(0, 1)$ binary image has variance of 0.25. However we use this definition for structured (non-random) images also. We select a confidence interval, δ, by which the predicted measurements can deviate from the actual measurement. Then for each $y(i)$ we find a subset, R_i, of allowable local object patterns according to

$$R_i = \{P_{ij} : |y(i) - \sum_{l \epsilon \ell_i} P_{ij}(l) h_i(l)| \leq \delta\} \tag{4.11}$$

The measurement $y(i)$ given a pattern, P_{ij}, is Gaussian distributed with mean $\sum_{l\epsilon\ell_i} P_{ij}(l)h_i(l)$ and variance σ^2. The size of the subset, $|R_i|$, is smaller than 2^{L_i} for measurements at high SNR (e.g., $SNR > 30dB$). $|R_i|$ depends on the measured value, $y(i)$, the standard deviation of noise, σ, and the confidence interval, δ. For a fixed $y(i)$ and σ, $|R_i|$ increases with increasing δ. And for a fixed $y(i)$ and δ, it increases with increasing σ. In Fig. 4.2 $L_i = 4$ and there are 16 local patterns, $y(i) = 0.23$, and $\sigma = 0.02$. If we choose a confidence interval $\delta = 6\sigma$, Eq. 4.11 defines patterns which yield predicted measurements in the range $(y(i) - \delta, y(i) + \delta) = (0.23 - 0.12, 0.23 + 0.12) = (0.11, 0.35)$. The size of the subset $R_i = \{P_{ij} : j = 2, 3, 4\}$ is 3 instead of the full 16. Notice that the ML estimate $(j = 2)$ is a member of this subset. The computational complexity of this modified 2D4 algorithm is directly dependent on $|R_i|$.

A brute force method for implementing Eq. 4.11 would require searching all 2^{L_i} patterns to find the subset, R_i, given $y(i)$ and δ. Because the patterns are binary, we can simplify this search using a tree-based procedure. Note that the predicted output for each P_{ij} is $\sum\limits_{l\in\ell_i} P_{ij}(l)h_i(l) = \sum\limits_{\substack{l\in\ell_i \\ P_{ij}(l)=1}} h(l)$. The construction of the tree is as follows: we place all 2^{L_i} possible local object patterns on as many nodes in a L_i-level tree. The nodes in the k^{th} level of the tree correspond to all patterns containing exactly k 1's. Nodes at the k^{th} level are originated from a parent node at the $(k-1)^{th}$ level. Each node is represented by an ordered pair; the first element is the pattern index, and the second element is the predicted channel output with that pattern as input. In each level all nodes with the same parent node are placed in

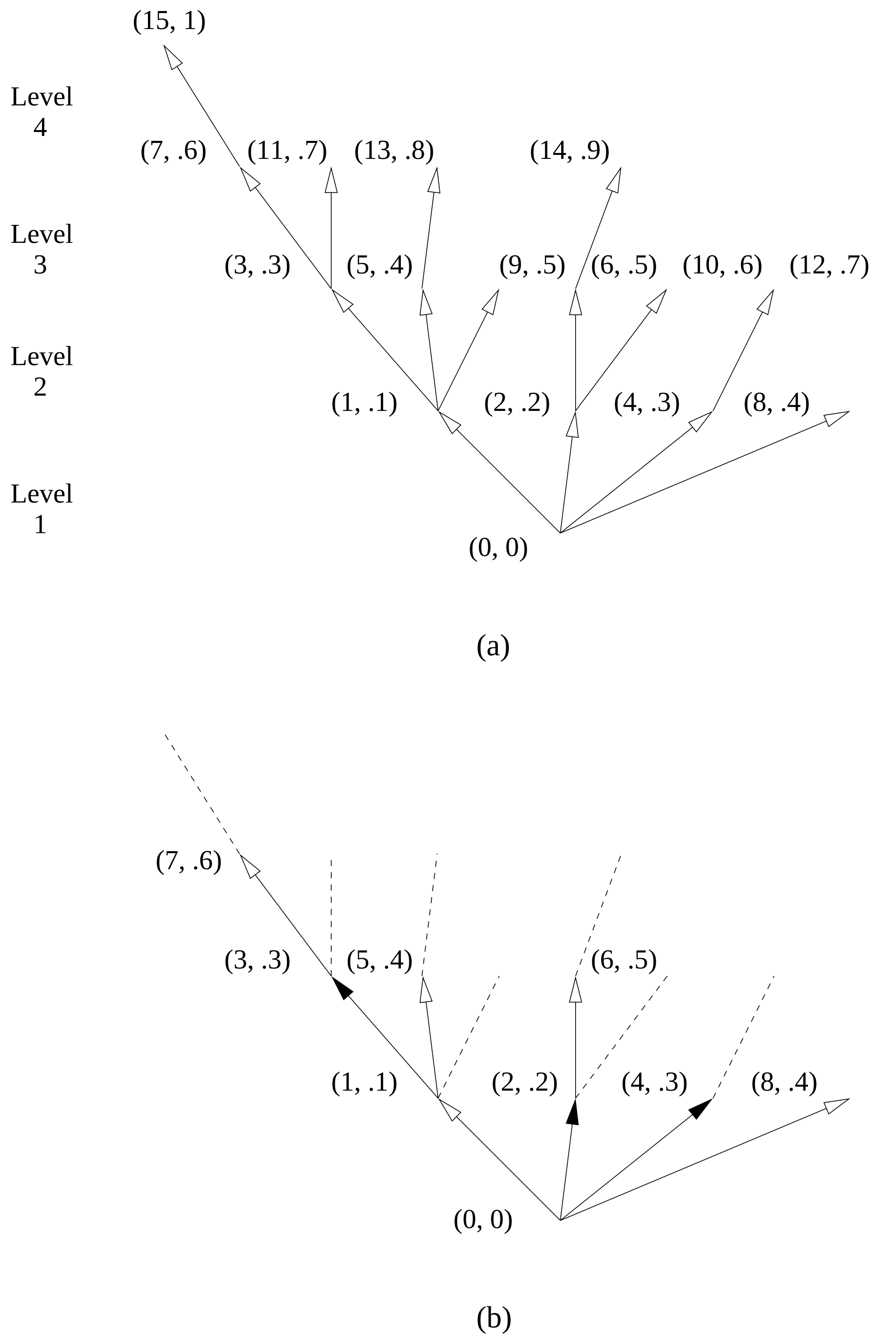


Figure 4.4: A tree-based search to find the subset of patterns, R_i, for the example in Fig. 4.2. (a) Full tree displaying all 16 nodes. (b) 9 nodes that are considered in the search. 3 nodes with dark-arrows form the subset R_i.

ascending order of their outputs. For example the tree in Fig. 4.4(a) is constructed for the channel PRF shown in Fig. 4.2. It has 4 levels and 16 nodes. The node in level 3 with the ordered pair $(5, .4)$ refers to the 5^{th} pattern ($m = 5$ in Fig. 4.2) corresponding to the output value of 0.4. Given a measurement pixel value $y(i)$, we begin the search from the node $(0, 0)$ and retain all nodes with outputs satisfying Eq. 4.11. If the output of a node at the k^{th} level exceeds $y(i) + \delta$, then the outputs of (1) all remaining nodes at the k^{th} level having the same parent node, and (2) all of its descendent nodes, exceed $y(i) + \delta$ as well. Hence those nodes (patterns) need not be considered in the search. We find the subset of patterns with outputs in the range $(0.11, 0.35)$ in Fig. 4.4(b) for the example mentioned above. Those patterns $R_i = \{P_{ij} : j = 2, 3, 4\}$ are represented by the nodes with dark arrows. The nodes with dashed lines in Fig. 4.4(b) represent all nodes that are ignored by the search routine. In this small example the number of patterns considered is 9 instead of 16. If N_i is the maximum number of 1's in any pattern that satisfies $\sum\limits_{l \in \ell_i} P_{ij}(l)h_i(l) \leq y(i) + \delta$, the number of patterns that are searched is less than $\sum\limits_{j=0}^{N_i+1} \binom{L_i}{j}$. We make another observation here that if the measured value is greater than 0.5, the nodes can be placed on the tree in descending order of their outputs. The search starts with the all 1 pattern instead of the all 0 pattern. In this case N_i is the maximum number of 0's in any pattern that satisfies $\sum\limits_{l \in \ell_i} P_{ij}(l)h_i(l) \geq y(i) - \delta$ and the same complexity is obtained. As a final note, this search procedure can be implemented using an efficient recursive programming method.

This method of complexity reduction may introduce an error when the noise sample, $n(i)$, corrupting a measurement is far away from its mean value of zero. The probability of that occurring depends on σ. For example if $\delta = 3\sigma$, $p(|n(i)| \leq 3\sigma) \geq 0.99$. As long as we choose a large enough δ (at the same time small enough to get a subset of manageable size), the error introduced by using the subset of patterns is small. The above method can be used with PRF support size as large as 16 at high SNR using a small δ. If the PRF size is larger than 16, we use this method in conjunction with the other published methods. For example if the PRF size is 20, we can ignore the 4 smallest coefficients to obtain an approximate PRF of size 16. However here the error is introduced both by the truncation of the PRF and by use of a partial list of possible patterns.

4.3.3 Processing of multiple images

Section 5.2.2 describes the 2D4 algorithm that can be used to find object pixel likelihoods based on a single measured image. Here we extend 2D4 to the case of multiframe image restoration. The idea is to estimate object pixel likelihoods based on multiple LR images in an iterative fashion to obtain a superresolved reconstruction. The schematic diagram for processing multiple images is shown in Fig. 4.5. The algorithm starts with an initial estimate of the object pixel likelihoods, $\{p_l^0(0), p_l^0(1)\}$. The likelihoods are updated based on the first LR image measurements. These updated likelihoods are passed on to process the second LR image and so on. One iteration of this scheme is complete when all K LR images are processed and the likelihoods are fed back to the first LR image. The algorithm can be stopped either after a certain number of iterations or when the tolerance on change in likelihoods is reached.

4.4 Simulation Results

In this section we present performance results for the algorithm described in the previous section. We consider an incoherent optical system with a rectangular aper-

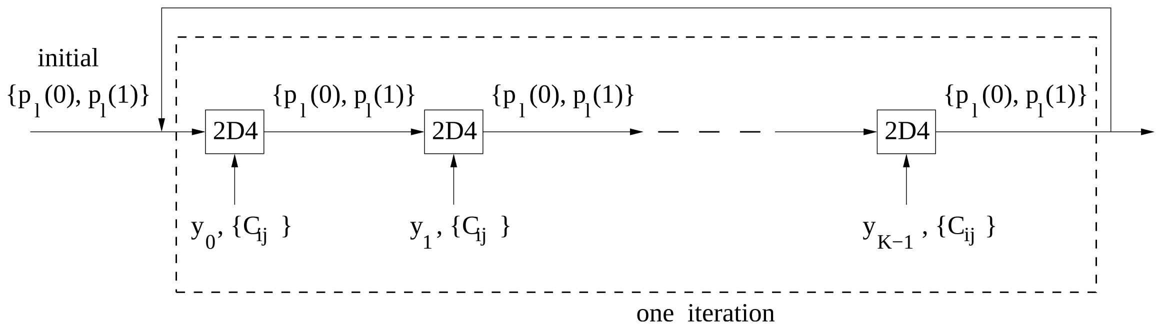


Figure 4.5: A scheme for processing multiple LR images with 2D4 algorithm.

ture. H_{opt} in the imaging equation Eq. 2.2 is determined by the optical point spread function, h_{opt}. The degree of optical blur w is defined as half-width of h_{opt}. It is equal to the separation between the peak and the first zero for h_{opt}. Fig. 4.6 is a 1D depiction of $h_{opt}(x) = \text{sinc}^2(\frac{x}{w})$ for $w = 0.5$. The higher the value of w, the larger is the amount of optical blur. The geometrical transformation operator, T_k, in Eq. 2.2 models the different views of the object in the k^{th} LR image, and is characterized by scale (s), rotation (θ), and translations (Δx, Δy) in two image coordinates. These parameters are chosen uniformly and at random from the ranges of values: $0.9 \leq s \leq 1.1$, $-15 \leq \theta \leq +15$ degrees and $-D \leq (\Delta x, \Delta y) \leq +D$ object pixels. The operators H_{ccd} and S are determined by the pitch of the measurement pixel, D, relative to the object pixel.

4.4.1 Reconstruction Fidelity

We use 10 random binary objects of size 32×32 pixels to quantify the reconstruction performances of all algorithms. Because the object is binary-valued, the reconstruction fidelity can be measured using the bit error rate (BER) between the original image, f, and the restored binary image $\widehat{f}$. BER is defined as $BER = \frac{\sum_{i \in C} |f(i) - \widehat{f}(i)|}{N_C}$ where C is the set of object pixels seen by all K cameras and N_C is the total number

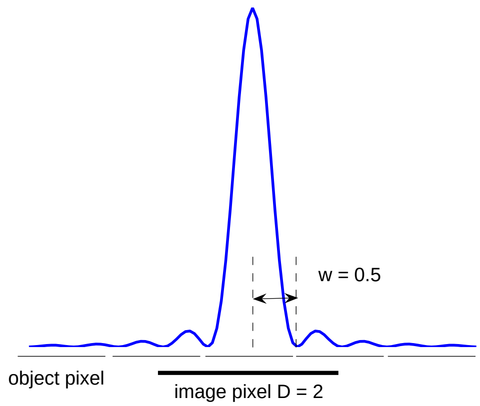


Figure 4.6: An example of the diffraction limited optical point spread function. The blur spot covers one object pixel. The measurement pixel width is twice the object pixel width.

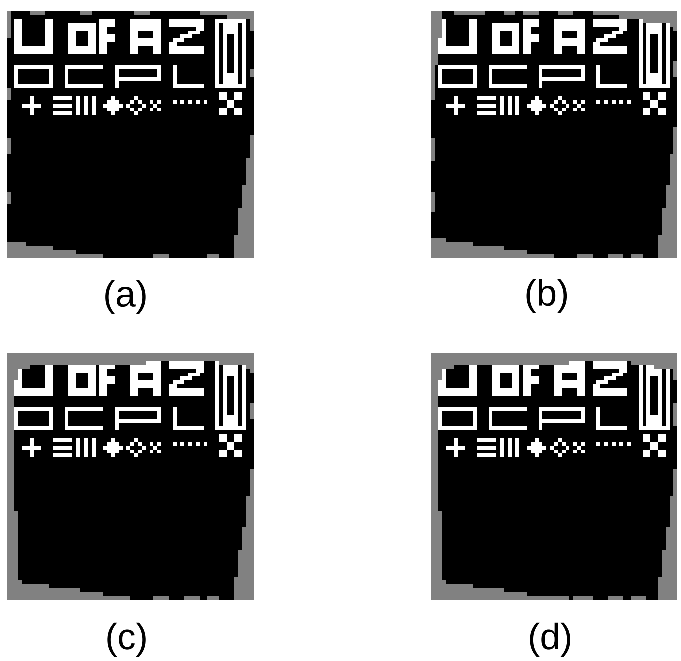


Figure 4.7: The set of common pixels, C, seen in a 64×64 object by all (a) $K = 1$, (b) $K = 2$, (c) $K = 3$ and (d) $K = 4$ cameras with $D = 2$. The total number of common pixels, N_C, are 3879, 3768, 3578, and 3561 respectively. The pixels in gray are not seen by all K cameras.

of those pixels. Fig. 4.7 shows the variation of the common set of pixels, C, in a 64×64 object with K for $D = 2$. The pixels that are not seen by all K cameras are shown in gray. N_C is non-increasing with K. The values of N_C are 3879, 3768, 3578, and 3561 for $K = 1, 2, 3$, and 4 respectively. The object autocorrelation required by the Wiener operator is modeled as a delta function for random imagery. Because we use the Wiener operator in the back projection for IBP, the reconstruction fidelity of the IBP is slightly better than, but follows that of the Wiener. Hence we show only the results of 2D4 and IBP to keep the plots readable. The relaxation parameters α in IBP and β in 2D4 are both set to 0.2. Generally, the smaller the relaxation parameter is, the slower is the convergence of the algorithm, and the better is the reconstruction fidelity [60, 53]. We use a Monte-Carlo simulation method to analyze the performances of the algorithms.

First we consider a system with no optical blur and $D = 2$ pixel down-sampling. In 2D4, we set $L_i = 16$ and we select a subset of patterns using a confidence interval, $\delta = 6\sigma$. Fig 4.8 plots the BER performances at different SNR values for different numbers of LR images, K. We make the following observations from this result. The reconstruction fidelity increases with K as expected for both IBP and 2D4. SNR values above which the BER performance of 2D4 is superior to that of IBP are listed in Table 4.1. At low SNR the knowledge of object autocorrelation function, R_f, regulates the reconstruction process in IBP. Also 2D4 suffers from propagation of pixel likelihood errors. Hence IBP outperforms 2D4 at low SNR. However the reconstruction BER using IBP reaches a floor after a certain SNR. The IBP BER with $K = 1$ remains practically the same for SNR from 20 to 50dB. The BER does not change significantly after 36dB and 46dB for $K = 2$ and $K = 3$ respectively. The performance of 2D4 improves steadily as SNR increases. In fact the object can

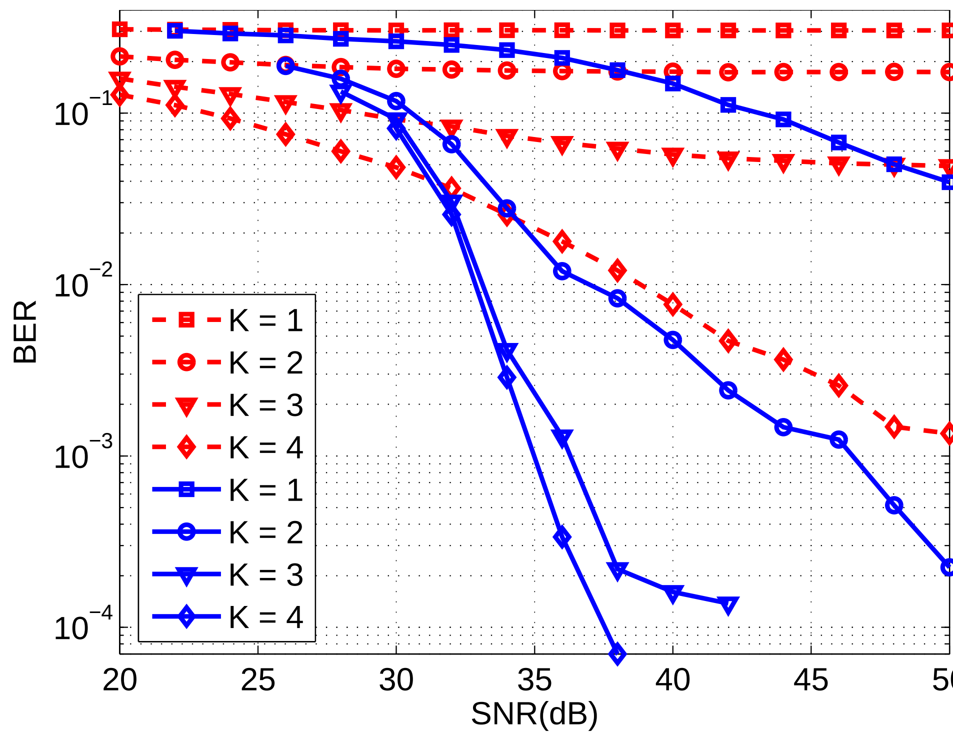


Figure 4.8: BER versus SNR for various K using an imaging system with $D = 2$ and no optical blur. The solid and dotted lines correspond to 2D4 and IBP respectively.

Table 4.1: SNR above which the BER performance of 2D4 is better than IBP with $D = 2$ pixel down-sampling.

	$K = 1$	$K = 2$	$K = 3$	$K = 4$
No optical blur	22	26	30	32
$w = 0.5$	20	26	31	34

be perfectly restored (i.e., less than 1 pixel error on average) at $SNR = 48$dB using just 2 LR images. The BER obtained at $SNR = 35$dB using 4 LR images by IBP can be obtained by using just 2 LR images via 2D4. BER $= 2 \times 10^{-3}$ is achieved at $SNR = 47$dB by IBP and at SNR $= 35$dB using $K = 4$ LR images. The SNR gains at BER $= 2 \times 10^{-3}$ offered by 2D4 over IBP with $K = 4$ LR images are listed in Table 4.2.

Now we consider an imaging system with $w = 0.5$ degree of optical blur and $D = 2$ pixel down-sampling. The optical blur spot covers one object pixel. The optical

Table 4.2: SNR gains at $BER = 2 \times 10^{-3}$ by 2D4 over IBP using $K = 4$ LR images and $D = 2$ pixel down-sampling.

	$K = 2$	$K = 3$	$K = 4$
No optical blur	4	11	12
$w = 0.5$	8	11	12

point spread function with $w = 0.5$ degree of blur is well-approximated by 3×3 image pixel support. In our discrete-to-discrete imaging system model, we use a bilinear interpolation method to implement T_k. Because of this, the measurement pixel PRF size may exceed 20. We truncate the PRF size to 16 using the method described in section 4.3.2. We select a subset of object local patterns of 16 pixels using a confidence interval $\delta = 6\sigma$ to further reduce the complexity. The reconstruction fidelity versus SNR is plotted at different SNR in Fig. 4.9 for random binary imagery. We see a similar trend in BER performance as before. SNR values above which the BER performance of 2D4 is superior to that of IBP is listed in Table 4.1. 2D4 provides significant BER improvement compared to IBP at $SNR > 35$dB . $BER = 2 \times 10^{-2}$ is obtained using IBP with $SNR = 36dB$ and $K = 4$ LR images; whereas only $K = 2$ LR images are required by 2D4. $BER = 5 \times 10^{-5}$ and $BER = 2 \times 10^{-2}$ are achieved by 2D4 and IBP respectively at $SNR = 38$dB. SNR gains at $BER = 2 \times 10^{-3}$ obtained using 2D4 over IBP with $K = 4$ LR images are listed in Table 4.2.

We use the 64×64 pixel image shown Fig. 4.10(a) as the object in order to inspect the visual quality of reconstruction. An imaging system with $w = 0.5$ degree of optical blur, $D = 2$ pixel down-sampling, and $SNR = 40$dB is considered. The autocorrelation function required by Wiener operator and IBP is extracted from the object itself. The 2D object is unrastered into 1D image. The autocorrelation function of this 1D image is plotted in Fig. 4.11(a). We fit this function by using the

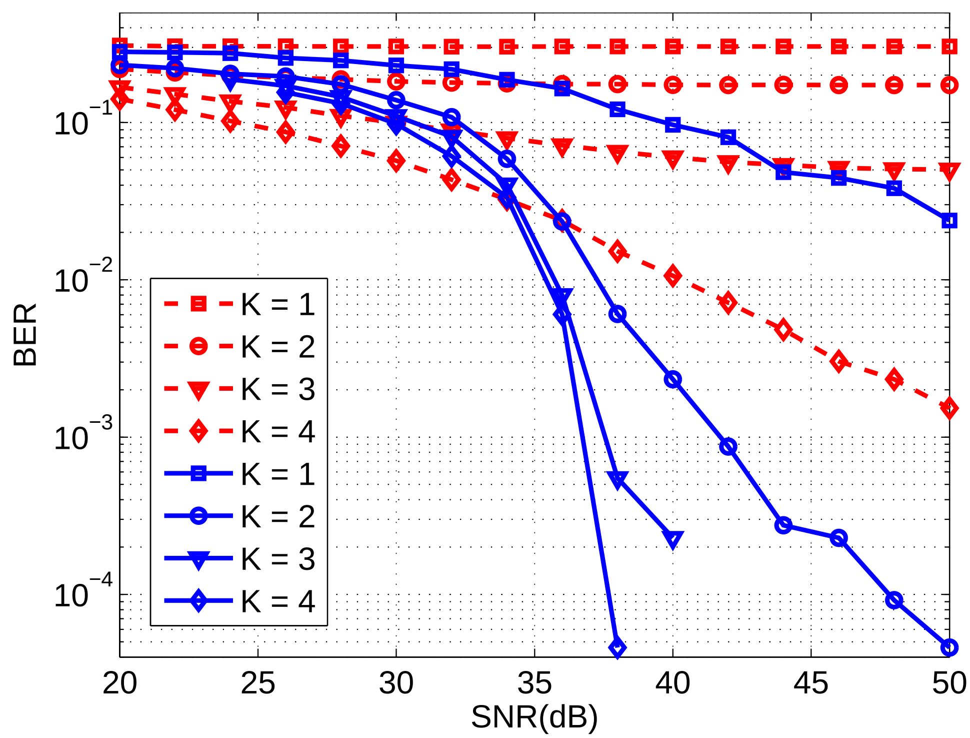


Figure 4.9: BER versus SNR for various K using an imaging system with D = 2, w = 0.5. The solid and dotted lines correspond to 2D4 and IBP respectively.

two generalized Laplacian functions of the form $e^{-|\frac{d}{\lambda}|^{\gamma}}$. The first function models the decay of the train of spikes corresponding to position indices $0, 64, 128, ..., 4032$ in Fig. 4.11(a). The parameters $\lambda = 8.29$ and $\gamma = 0.436$ give the best fit in mean square error sense. The model and the actual values are plotted in Fig. 4.11(b). This model is then used to obtain the best fit for all remaining variations (e.g., for position indices $64, 65, ..., 127$) in Fig. 4.11(a). The parameters of the second function are $\lambda = 7.60$ and $\gamma = 0.616$. An example of the model and the actual values are plotted in Fig. 4.11(c). The solid curves are the actual and the dashed curves are the modeled functions. The autocorrelation matrix, R_f, is completely determined by these four parameters. We resort to this two-function model as the correlation distances of the 2D object in two directions are different: it is 25 along vertical and 61 along horizontal. This model gave us a closer fit than other models studied. Fig. 4.10(b) shows a measured 32×32 LR image that has been interpolated to the original object size using pixel replication. The fine details in the objects are lost in this LR image. Fig. 4.10 (c), (e), and (g) are restored images using IBP after 50 iterations and $K = 1$, 2, 3 respectively. These images have 188, 29, and 2 pixel errors respectively. Images reconstructed using 2D4 after 50 iterations are shown in Fig. 4.10 (d), (f), and (h) and they have errors in 65, 17, and 1 pixels. We note that the pixel errors are counted only in the region that is seen in all K LR images. Images restored using 2D4 are visually better than IBP. Fine details in the object that correspond to high frequencies are well-restored by 2D4; while IBP fails to restore them.

We use the 64×64 object shown in Fig. 4.10(a) to obtain BER versus SNR performance in an imaging system with larger pixel size. We consider $D = 3$ pixel down-sampling with no optical blur. Once again 2D4 uses a truncated PRF with

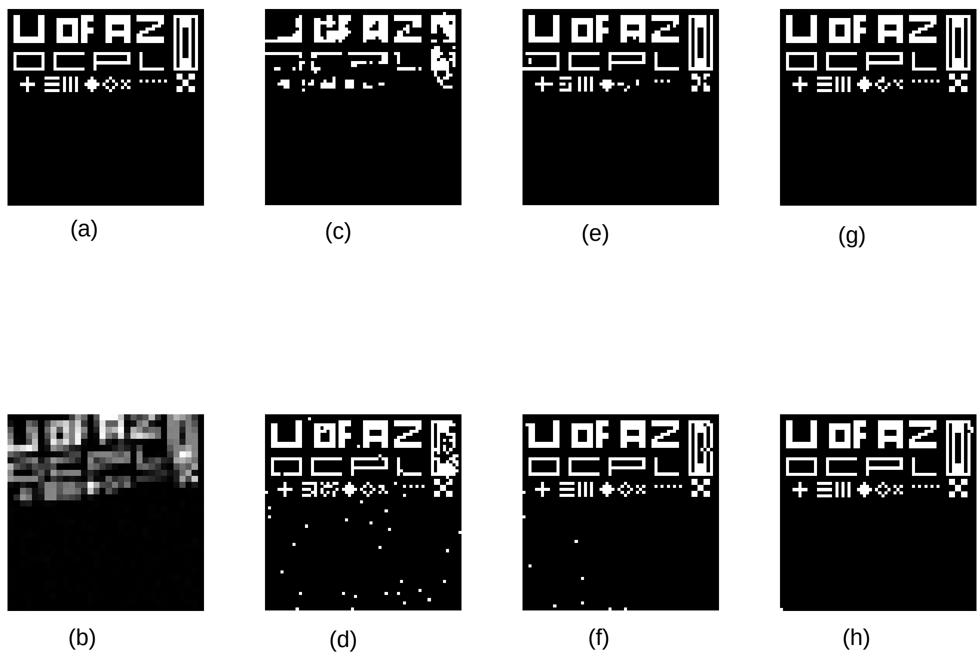


Figure 4.10: Example reconstructions of an object using IBP and 2D4 with $w = 0.5$, $D = 2$, and $SNR = 40dB$. (a) Original 64×64 object, (b) a measured image with $D = 2$ down-sampling that is interpolated using a pixel-replication method. Restored images by IBP (c) $K = 1$, (e) $K = 2$, and (g) $K = 3$ with 188, 29, and 2 pixel errors respectively, and by 2D4 (d) $K = 1$, (f) $K = 2$, and (h) $K = 3$ with 65, 17, and 1 errors respectively. Pixel errors are counted only in the region that is seen by all K LR images.

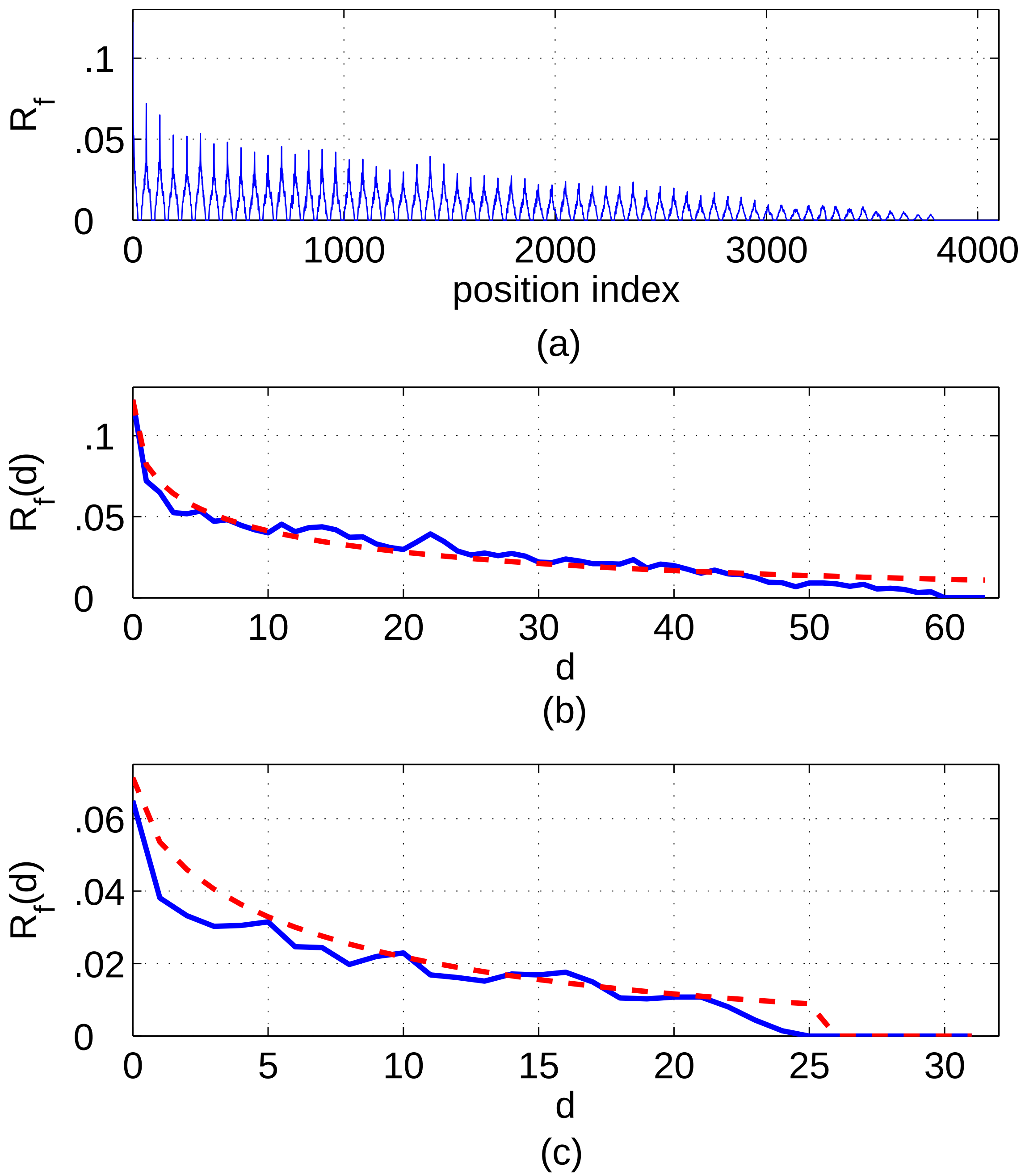


Figure 4.11: (a) Autocorrelation of the object in Fig. 4.10(a). (b) The first Laplacian function (dashed) models the object autocorrelation (solid) along horizontal direction. (c) The second function models the autocorrelation along vertical direction. The correlation distances along horizontal and vertical direction are 61 and 25 respectively.

Table 4.3: SNR above which the BER performance of 2D4 is better than IBP with $D = 3$ pixel down-sampling.

R_f	$K = 1$	$K = 3$	$K = 5$
Actual	55	53	53
Model	51	50	50

16 pixels. δ is set to 6σ. We use both the actual autocorrelation and the Laplacian model in the IBP. The BER versus SNR plots are shown in Fig. 4.12. SNR values above which the BER performance of 2D4 is superior to that of IBP are listed in Table 4.3. The IBP with the actual autocorrelation yields very good restoration performance. In fact the Wiener filter with the precise knowledge of object autocorrelation bounds the performance by any linear restoration method for this object. The BER performance of IBP with Laplacian model is only 1.3×, 2.0×, and 3.8× worse than IBP with perfect autocorrelation priors in the SNR range of $50 - 60$dB. At $SNR = 54dB$, BER achieved by 2D4 are 1.2×, 2.7×, and 12.1× lower than IBP with the Laplacian model at $K = 1$, 3, and 5 respectively. Example reconstructions of the object at $SNR = 54dB$ are shown in Fig. 4.13. Fig. 4.13(a) and (b) are two measured LR images. They are interpolated to the original size using a pixel replication method. Images restored using IBP with $K = 3$, 4, and 5 LR images are shown in Fig. 4.13(c), (e), and (g). They have 62, 35, and 16 pixel errors respectively. Fig. 4.13(d), (f), and (h) show restored images using 2D4 with $K = 3$, 4, and 5 LR images. They have 24, 4, and 2 pixel errors respectively. Also the 2D4 better-resolves the fine structures in the center of the object.

We consider here imaging systems with large detectors (e.g., $D = 4$). The channel support size becomes very large when there is a high down-sampling ratio. The numbers of local object patterns in the 2D4 grow exponentially. The complexity of 2D4 also depends on the standard deviation of AWGN. We study the computational

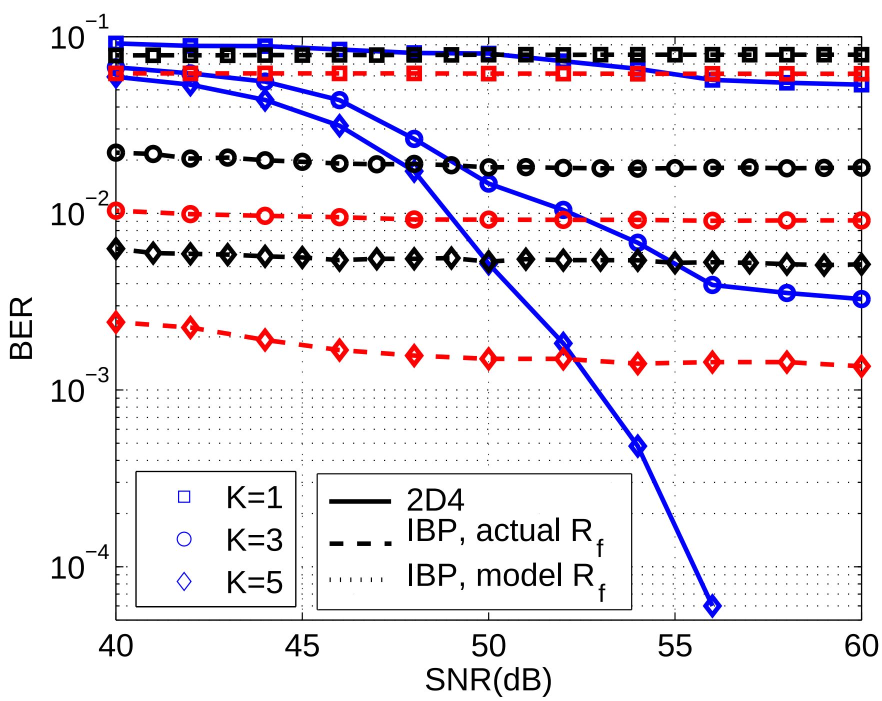


Figure 4.12: BER versus SNR plot for different K for an imaging system with $D = 3$ and no optical blur. The solid curve corresponds to 2D4. The dashed and dotted lines correspond to IBP using the actual and the model R_f respectively.

Figure 4.13: Example reconstructions of an object using multiple LR images by IBP and 2D4 at $SNR = 54dB$. (a) and (b) are measured images with $D = 3$ down-sampling. They are interpolated to the original size using a pixel-replication method. Restored images by IBP with (c) $K = 3$, (e) $K = 4$, and (g) $K = 5$ have 62, 35, and 16 pixel errors respectively, and by 2D4 with (d) $K = 3$, (f) $K = 4$, and (h) $K = 5$ have 24, 4, and 2 errors respectively. Pixel errors are counted only in the region that is seen by all K LR images.

complexity of the algorithms in the next section. Because of increased complexity, performance curves for systems with large detectors are time-consuming and memory-expensive to obtain. Instead, we present few example reconstructions for $D = 4$ case with no optical blur in Fig. 4.14 at a high $SNR = 70dB$. We form contrived channels by truncating the original channel supports (possibly more than 28 pixels) to 20 pixels. These contrived channels are treated as the "actual" PRFs in both IBP and 2D4. The confidence interval $\delta = 4\sigma$ is used in 2D4. Fig. 4.14(a) and (b) show two measured LR images that are interpolated as before. The object details are completely lost in these images. IBP uses the Laplacian-modeled object autocorrelation function. Images restored using IBP with $K = 2$, 4, and 8 LR images are shown in Fig. 4.14(c), (e), and (g). They have 310, 92, and 17 pixel errors respectively. Fig. 4.13(d), (f), and (h) show restored images using 2D4 with K = 2, 4, and 8 LR images. They have 141, 49, and 1 pixel errors respectively. Again, the 2D4 better resolves the fine object details than IBP. Some object details are restored by 2D4 using just $K = 2$ LR images. 2D4 achieves almost-perfect object reconstruction using $K = 8$ LR images.

4.4.2 Computational Complexity

In this section we analyze the complexity reduction achieved by the technique described in section 4.3.2. We also compare the computational complexities of IBP and 2D4. Two imaging systems are considered for the analysis of complexity reduction. One has $w = 0.5$ degree of optical blur, $D = 2$ pixel down-sampling, and $K = 4$ LR images. The second has $D = 3$ pixel down-sampling, and $K = 4$ LR images with no optical blur. The first case employs 10 random binary objects of size 32×32 pixels; whereas the 64×64 object in Fig. 4.10(a) is used in the second

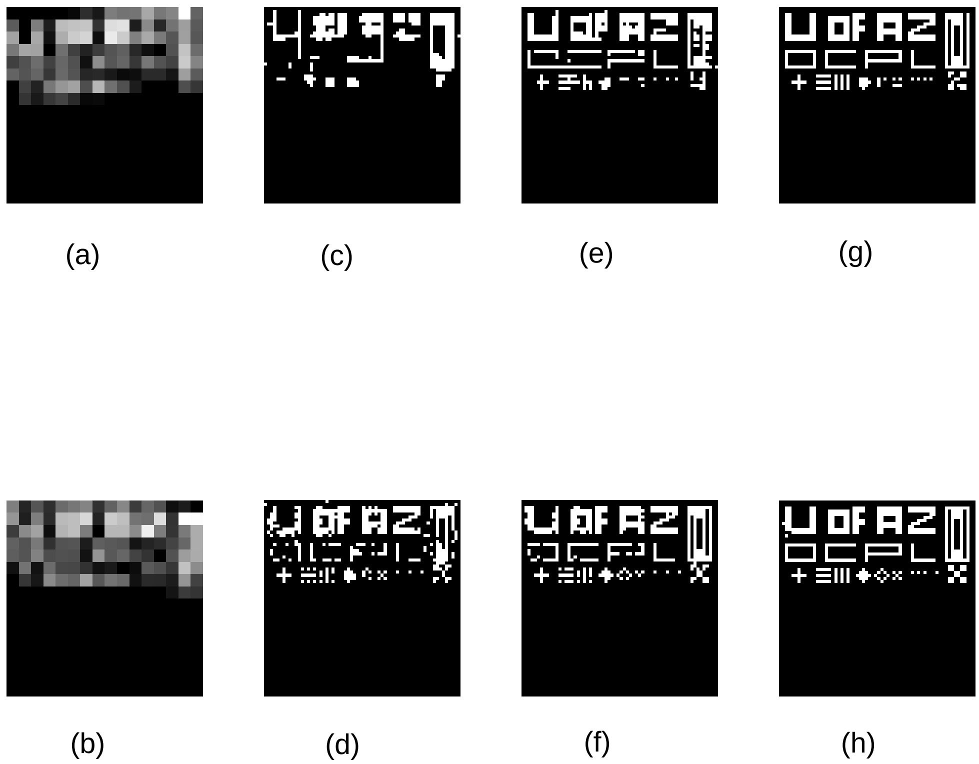


Figure 4.14: Example reconstructions of an object using multiple LR images by IBP and 2D4 at $SNR = 70dB$. (a) and (b) are measured images with $D = 4$ down-sampling. They are interpolated to the original size using a pixel-replication method. Restored images by IBP with (c)$K = 2$, (e)$K = 4$, and (g)$K = 8$ have 310, 92, and 17 pixel errors respectively, and by 2D4 with (d)$K = 2$, (f)$K = 4$, and (h)$K = 8$ have 141, 49, and 1 errors respectively. Pixel errors are counted only in the region that is seen by all K LR images.

case. In both cases we truncate the PRF to size $L_i = 16$ in 2D4. The total number of local object patterns is therefore $2^{L_i} = 65536$. We obtain subsets of local object patterns using $\delta = 6\sigma$. At $SNR = 30dB$, $\sigma = .016$. The predicted outputs of patterns, R_i, satisfying Eq. 4.11 fall in the ranges $(y(i) - .096, y(i) + .096)$. At a higher $SNR =$ 54dB, $\sigma = .001$. The predicted outputs of R_i fall in smaller ranges $(y(i) - .006, y(i) + .006)$. We can expect to obtain smaller subsets at $SNR =$ 54dB than 30dB. The size of the subset depends also on actual values of measurements. For example pixel measurements of a sparse binary object may yield smaller subsets. The complexity reduction achieved is quantified as $\frac{1}{P}\sum_{i=0}^{P-1}\frac{2^{L_i}}{|R_i|}$ where $P = KM$ is the total number of measurements. We calculate this quantity for both imaging systems. This reduction in the number of local object patterns is plotted versus SNR in Fig. 4.15. For fixed δ, the subset size of local object patterns reduce with increasing SNR as expected. 2D4 complexity is reduced by factor 3 at $SNR =$ 30dB and by factor 80 at $SNR =$ 60dB for random imagery. For the textured image the complexity reduces by factors 10 and 130 at $SNR = 30dB$ and 60dB respectively.

We study the computational cost of super-resolving an object using IBP and 2D4 in terms of the number of operations required for $D = 2$ pixel down-sampling and $w = 0.5$ degree of blur. Even though the optics is shift-invariant, the down-sampling by the detectors makes the imaging system shift-variant. In this case, the Wiener operator that is used in the back-projection kernel is calculated in the spatial domain. The system matrix, H, and the autocorrelation matrix, R_f, are of sizes $P \times N$ and $N \times N$ respectively. The calculation of the Wiener operator is given in Eq. 4.3. However we make use of the sparsity of H to simplify the matrix multiplications involving H. Let L represent the full support size of the PRFs of the measurements. Then calculation of the Wiener operator would take roughly $O(PNL + P^2L + P^3 + NPL + NP^2)$ operations. It is assumed here that it requires $O(P^3)$ operations to invert a matrix of size $P \times P$ (in Eq. 4.3). The number of operations in each iteration of IBP is roughly $O(NL + NP)$. For the complexity calculations of the 2D4 algorithm, the truncated PRF support size, L_i, is set to 16. Let N_i be the maximum number of either 1 or 0 in any pattern that satisfies the condition in Eq. 4.11 as explained in section 4.3.2. The total number of subsets of patterns, and hence the computations required is at most $\sum_{i=0}^{P-1} \sum_{j=0}^{N_i} \binom{L_i}{j}$. These computations involve only additions, and in this process the predicted output of these patterns are also calculated ($\sum_{l \in \ell_i} H(i,l)P_{ij}(l)$ in Eq.4.6). The remaining computational complexity of 2D4 algorithm depends on the subset sizes, σ, δ, actual measured values, and total number of iterations. We modify our restoration program to estimate the computational cost. We replace the actual steps of the algorithm by a counter that counts the numbers of operations along the program flow. Fig. 4.16 plots number of operations versus SNR for different numbers of LR images. The total number of iterations in both algorithms is 50, and 2D4 uses a confidence interval of 6σ. We see the general trend of increasing computational complexity with increasing number of LR images in both algorithms as expected. The complexity is invariant to the noise levels for IBP. However the complexity of 2D4 depends on the SNR; the size of R_i increases with increasing σ (decreasing SNR) for fixed δ and measured values as explained in section 4.3.2. Hence the complexity of 2D4 is much higher at lower SNR values. At SNR range of 50-60dB, the 2D4 is $1.5 - 5\times$ more complex than IBP for $K = 4$.

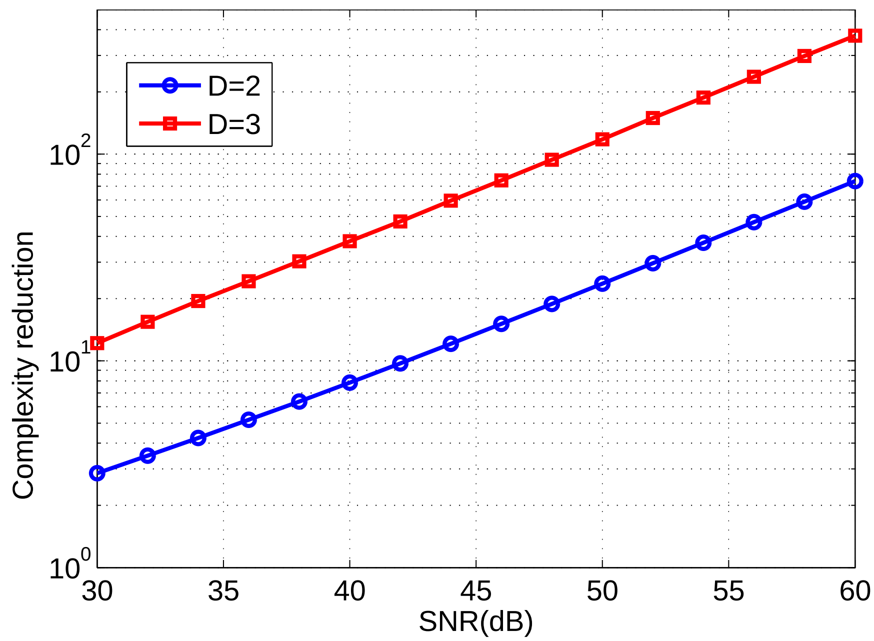


Figure 4.15: Reduction in number of local object patterns versus SNR.

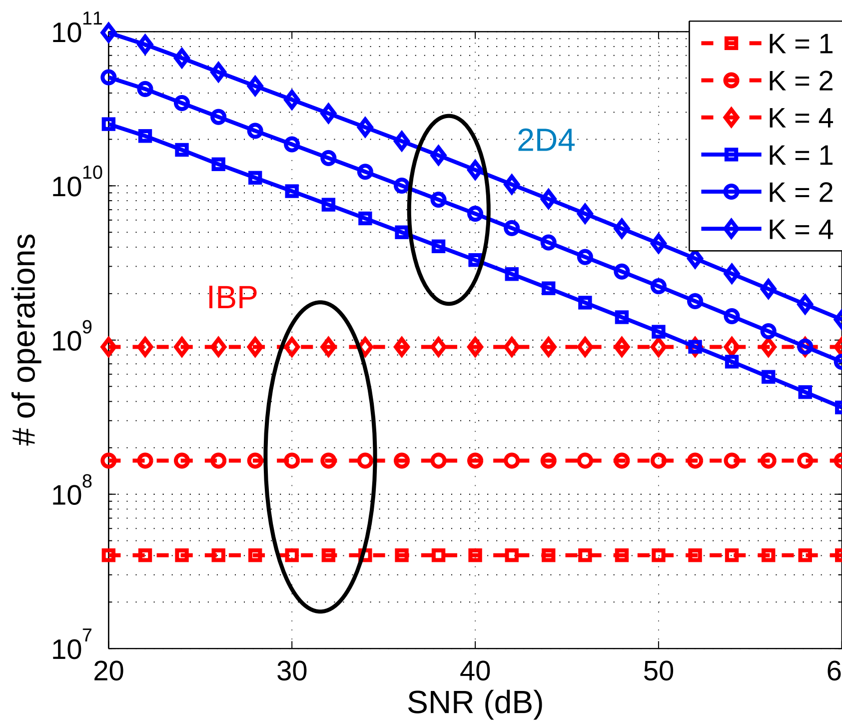


Figure 4.16: Computational complexity versus SNR for 2D4 and IBP.

4.5 Conclusions

We have described a two-dimensional likelihood based algorithm for restoring a high-resolution object from multiple degraded, shifted, rotated, scaled and under-sampled low-resolution images. The 2D4 algorithm estimates the object pixel likelihoods based on measurements, noise statistics, and exchange of neighboring pixel likelihoods using a belief propagation mechanism. A novel non-training based technique is presented to overcome one of the main hurdles of the algorithm, namely, computational complexity. We outlined an efficient method to search for the subset of object patterns to be used in the algorithm. The BER fidelity of 2D4 is much better than IBP at high SNR values. In an imaging system with $w = 0.5$ degree of optical blur and $D = 2$ pixel down-sampling, we achieve significant SNR gains of $\{8, 12\}$ dB by 2D4 with $K = \{2, 4\}$ LR images over IBP with $K = 4$ at a BER $= 2 \times 10^{-3}$. We also presented results for imaging systems with $D = 3$ and $D = 4$ pixel down-sampling. The images restored using 2D4 show finer object details than those using IBP. We showed that the computational complexity of 2D4 is less than $5\times$ that of IBP at high SNR values of 50-60dB.

Because of the exponentially growing complexity of the algorithm, the 2D4 is not suitable for restoring binary images from multiple heavily under-sampled LR images. For example, the PRF size with $D = 3$ down-sampled pixels and $w = 0.5$ degree of optical blur may well exceed 25 pixels. If it is truncated to 16 pixels, significant error is introduced in the calculations of likelihoods. The 2D4 algorithm with a subset of patterns may fail to estimate the object pixel likelihoods due to error propagation in the message passing mechanism. With some added complexity, say $L_i = 20$, we may improve the reconstruction. At $SNR = 60dB$, we obtain a complexity reduction by a factor of 700 by using the technique presented here.

However the computational cost and the time required to search for the subset of patterns limit the application of the algorithm for restoring large objects, and/or for processing in real time.

CHAPTER 5

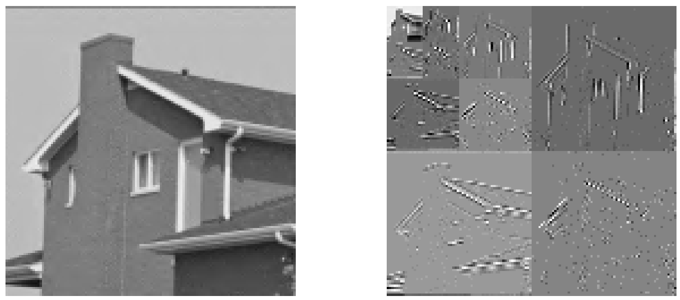
Sparsity Constrained Multiframe Image Restoration

5.1 Introduction

In many cases the objects that we are interested in imaging are sparse. For example an astronomical image might contain a twin-star at its center with dark background everywhere else. This object is considered sparse because it has relatively few bright pixels and a large number of background pixels. The idea of sparsity can be extended to objects that appear dense in the space-domain, but are sparse in some transform domain. For example the 128×128-pixels object shown in Fig. 5.1(a) is dense with 16384 non-zero pixels in the space-domain. If we compute a wavelet transform of this object using the Daubechies-4 filter, there are only 4079 non-zero wavelet coefficients. Hence this object is 25% sparse in the wavelet domain. The wavelet decomposition of the object at 3 resolution levels is shown in Fig. 5.1(b). Wavelet coefficients in each subband are scaled in order to enhance the contrast for display purpose. Recent enthusiasm concerning sparsity as a form of prior knowledge has fueled many new signal/image reconstruction algorithms [61, 55, 62, 63, 64, 65]. These algorithms operate by finding the most sparse signal that satisfies the measurement constraints. A precise metric of signal/image sparsity is the ℓ_0 norm which simply counts the number of non-zero elements in a vector. Signal reconstruction via minimizing the ℓ_0 norm however is a combinatorial problem which is NP-hard [61]. Hence other approximate metrics have been devised to measure sparsity. The ℓ_1 norm is the most popular of these approximations. The

ℓ_1 norm of a discrete signal, $\mathbf{z}$, of length N is defined as $||\mathbf{z}||_1 = \sum_{i=0}^{N-1} |z_i|$. A basis pursuit algorithm finds the optimal signal that minimizes the ℓ_1 norm by making use of recent advances in linear and quadratic programming methods [55]. ℓ_1 norm minimization techniques implemented using linear programs have been applied to recover sparse signals from random projections [62]. Daubechies et al. [63] used an iterative thresholding method to solve linear inverse problems. They combined iterative back-projection of an error signal with wavelet soft-thresholding [66] to obtain a sparse solution. A similar approach was developed in an expectation-maximization (EM) framework to restore an image in the wavelet domain [67]. Another EM algorithm is proposed in Ref. [68] to obtain a sparse representation of a signal in a sparse Bayesian learning framework.

In this chapter we consider incorporating the prior knowledge that many natural objects are sparse in the wavelet domain. Even though sparsity constrained algorithms have been recently used in reconstruction of signals from compressed measurements, this chapter describes a new algorithm for incorporating sparsity constraints to superresolve an object from multiple LR images. We present a novel regularized image restoration algorithm that minimizes the ℓ_1 norm of the object. In many regularization problems the fidelity of the estimate depends heavily on the regularization parameter and finding an optimal regularization parameter is quite challenging. We describe how a Krylov subspace based method, LSQR, [69, 70] can be used to efficiently find the optimal regularization parameter. We compare results from our method with the EM algorithm proposed by Figueiredo [67]. It is also well-known that natural objects can be modeled in the wavelet domain using generalized Gaussian densities (GGD) [71]. We use a set of natural images as training objects and learn the GGD parameters. Object models using these learned param-



(a) (b)

Figure 5.1: An example of a 128×128-pixels object in the (a) space domain with 16384 non-zero pixels, and (b) wavelet domain with 4079 non-zero wavelet coefficients. The wavelet decomposition is performed at 3 resolution levels using the Daubechies-4 filter. The wavelet coefficients in each subband are scaled to enhance image contrast for display purpose.

eters are used in the MAP setting to obtain a sparse object estimate. We compare reconstructions from these nonlinear techniques with linear minimum mean squared error (LMMSE) estimation. This chapter is organized as follows. In section 5.2 we briefly discuss sparse signal reconstruction techniques and present the proposed multiframe image restoration algorithms. Sparse object reconstructions obtained using different imaging system models and different algorithms are analyzed in section 5.3. The sparse reconstruction methods are applied to images gathered in a laboratory experiment in section 5.4. We conclude the chapter in section 5.5.

5.2 Sparsity Constrained Image Restoration

Before discussing various sparsity-motivated reconstruction algorithms we first express the multiframe imaging equation in terms of an object representation, z, in some transform domain, e.g., principal components or wavelets, as

$$y = Hf + n = HW^T W f = Qz + n \tag{5.1}$$

where W is a unitary transformation matrix, $Q = HW^T$, and $z = Wf$. If the object is sparse in the wavelet domain, then W represents the 2D discrete wavelet transformation matrix. The object estimate, $\widehat{z}$, is first obtained in the transform domain using a sparse image reconstruction method. The Daubechies-4 wavelet filter used in our work produces a sparse W. Hence Eq. 5.1 again represents a sparse linear system of equations and sparse matrix techniques can be used to solve it efficiently. The object estimate in the spatial domain, $\widehat{f}$, is then determined by $\widehat{f} = W^T \widehat{z}$.

5.2.1 Sparse Reconstruction Methods

As we mentioned in section 1 there are several techniques available for sparse image reconstruction. Many of these algorithms were developed for systems without measurement noise, $y = Hf$, and then were extended to include measurement noise. In the presence of AWGN, basis pursuit with inequality constraints [72] obtains a solution to Eq. 5.1 according to

$$\widehat{z} = \arg\min_{z} ||z||_1 \quad \text{s.t.} \quad ||Qz - y||_2^2 \leq \epsilon \tag{5.2}$$

where $\epsilon > 0$ is a small constant. This convex minimization problem can be recast as an instance of the second order cone problem and solved using efficient interior-point methods [55].

Another popular method of signal reconstruction is based on iterative wavelet shrinkage [63]. This technique attempts to solve the regularized ℓ_1 minimization

$$\widehat{z} = \arg\min_{z} [\ ||Qz - y||_2^2 + \lambda ||z||_1\ ] \tag{5.3}$$

where λ is the regularization parameter. This method obtains a solution to the above minimization problem in a two-step iterative process. Starting from an initial estimate, the estimate is updated at the $(k+1)^{\text{th}}$ iteration as follow:

$$\begin{aligned} \widehat{z}^{(k+1)} &= \widehat{z}^{(k)} + Q^T(y - Q\ \widehat{z}^{(k)}) \\ \widehat{z}^{(k+1)} &= sgn(\widehat{z}^{(k+1)}) \times \max(\ |\widehat{z}^{(k+1)}| - \tau\ ,\ 0\ ) \end{aligned} \tag{5.4}$$

The first step is the back projection of the difference signal, $y - Q\ \widehat{z}^{(k)}$, using the adjoint operator and the second step is soft-thresholding using the threshold τ. It is clear to see how this method promotes sparsity in the estimate: thresholding at each iteration removes small wavelet coefficients. However determining an optimal threshold value is not straightforward. The algorithm can be started with $\widehat{z}^{(0)}$ initialized to a zero-vector or $Q^T y$. The algorithm is stopped when the reduction in the objective function values in successive iterations falls below a certain fixed threshold. Solving the minimization problem in Eq. 5.3 by using an EM method also leads to the same iterative strategy of Eq. 5.4 [67]. Often a "debiasing" step is performed on the estimate obtained using the EM algorithm [73]. Debiasing attempts to determine the best fit estimate in terms of least squares on the support set (locations of non-zero elements) of the estimate from EM. This is often required to make sure that data consistency is not violated while minimizing the total (objective) cost function in Eq. 5.3. However it may not be desirable as it might amplify the measurement noise and undo the effects done by wavelet shrinkage. This is especially true when the threshold parameter is not close to the optimal value and/or the support set of the estimate is not a close representation to that of the actual object. Hence we avoid the "debiasing" step in our work.

5.2.2 The proposed algorithm

We first consider a general regularized least square problem with a regularization operator L as

$$\widehat{z}(y,\lambda) = \arg\min_{z} \left[\ ||Qz - y||_2^2 + \lambda ||Lz||_2^2\ \right] \tag{5.5}$$

where λ is the regularization parameter. Regularized least squares approach provides some flexibility in the minimization of an objective function with two costs. For example when the measurement noise variance is not accurately known, the regularization parameter can be used to trade-off data consistency and object prior costs. The regularization operator may simply be an identity operator (which refers to Tikhonov regularization [21]) or a Laplacian operator [25]. Below we describe a data-dependent L that promotes sparsity. Several methods have been proposed to estimate the optimal value of the regularization parameter [74, 75]. There are many efficient Krylov subspace based methods to solve Eq. 5.5, e.g., LSQR [69, 70]. LSQR is analytically equivalent to the conjugate gradient method and is numerically more stable. The possibility of efficiently solving Eq. 5.5 for several values of λ is a very attractive feature of LSQR as it facilitates an efficient determination of the optimal

λ. The quantities generated during the intermediate process in LSQR algorithm are independent of regularization parameter λ ([69] section 2.3). We modify the LSQR algorithm to solve Eq. 5.5 for multiple λ values with some increase in memory requirement and computational complexity. For detailed discussions of the Krylov subspace based methods, in particular LSQR, and their applications see references [69, 70, 75, 76, 77, 74].

Now we describe an doubly-iterative algorithm to solve Eq. 5.3. First we rewrite Eq. 5.3 in the form of Eq. 5.5. The operator L is the diagonal matrix whose i^{th} element on the diagonal $L_{ii} = \frac{1}{|z_i|^{\frac{1}{2}}}$. A similar technique was used to obtain the scaling matrix in an affine scaling methodology for best basis selection [78]. Any small or zero wavelet coefficients do not pose numerical problems as we explain in section 5.2.3. Note that in our case L depends on the object. We propose to solve this minimization problem in an iterative fashion. The solution to Eq. 5.5 is a function of both y and λ. We represent by $\widehat{z}^{(k)} \equiv \widehat{z}(y, \lambda^{(k)})$ the estimate obtained by solving Eq. 5.3 using the optimal regularization parameter, $\lambda^{(k)}$, at the k^{th} iteration. Given the estimate at the k^{th} iteration, we obtain an estimate at the next iteration as follows:

$$\widehat{z}^{(k+1)} = \arg\min_{z} [\ ||Qz - y||_2^2 + \lambda^{(k+1)} ||L^{(k)} z||_2^2\ ] \tag{5.6}$$

where $\widehat{z}^{(k+1)}$ is the estimate at the $(k+1)^{th}$ iteration, and $L^{(k)}$ is the diagonal regularization operator obtained using $\widehat{z}^{k}$. The i^{th} element of the diagonal of $L^{(k)}$ is $\frac{1}{|\widehat{z}_i^{(k)}|^{\frac{1}{2}}}$. And the optimal regularization parameter at the $(k+1)^{th}$ iteration is determined as

$$\lambda^{(k+1)} = \arg\min_{\lambda} ||L^{(k)} \widehat{z}(y, \lambda, k)||_2^2 \quad \text{s.t.} \quad ||Q\widehat{z}(y, \lambda, k) - y||_2^2 \leq P\sigma^2 \tag{5.7}$$

where P is the number of measurement pixels and $\widehat{z}(y,\lambda,k)$ is the solution to the minimization problem

$$\widehat{z}(y,\lambda,k) = \arg\min_{z} [\ ||Qz-y||_2^2 + \lambda||L^{(k)}z||_2^2\ ] \tag{5.8}$$

We refer to this doubly-iterative algorithm as "ℓ_1 penalty". This is doubly-iterative because the LSQR method used to solve Eq. 5.6 is itself iterative. Techniques such as the L-curve method can instead be used to determine the optimal parameter $\lambda^{(k+1)}$. The L-curve is a parametric plot of $||L^{(k)}\widehat{z}(y,\lambda,k)||_2^2$ versus $||Q\widehat{z}(y,\lambda,k)-y||_2^2$ with λ as a parameter. The $\lambda^{(k+1)}$ refers to the L-corner [75] which is the point with maximum curvature on the L-curve. However we observed that methods to determine the L-corner point are often not robust and not reliable for very ill-conditioned systems. Hence we compute the regularization parameter by solving Eq. 5.7. The objective cost function at the k^{th} iteration is

$$C^{(k)} = ||Q\widehat{z}^{(k)} - y||_2^2 + \lambda^{(k)}||L^{(k-1)}\widehat{z}^{(k)}||_2^2 \tag{5.9}$$

and the algorithm is stopped when the reduction in objective cost between two consecutive iterations is smaller than a threshold: i.e., when $\frac{C^{(k)}-C^{(k-1)}}{C^{(k)}} \leq \epsilon$ where $\epsilon \geq 0$.

5.2.3 Computational aspects

It may at first appear that it is necessary to solve a set of regularized equations Eq. 5.6, 5.7, and 5.8 at each iteration to obtain the final estimate. However the LSQR algorithm can be modified to generate solutions $\widehat{z}(y,\lambda,k)$ for multiple values of the regularization parameter with only a slight increase in computation and memory requirements [69, 70]. So Eq. 5.6-5.8 are solved in one pass of the LSQR algorithm. Also Eq. 5.6 is solved by first converting the problem into so-called "standard form" [75]

$$\begin{aligned}
\widehat{a}^{(k+1)} &= \arg\min_a ||Q[L^{(k)}]^{-1}a - y||_2^2 + \lambda^{(k+1)}||a||_2^2 \\
\widehat{z}^{(k+1)} &= [L^{(k)}]^{-1}\widehat{a}^{(k+1)}
\end{aligned}$$

The "standard form" refers to the Tikhonov regularized least square problem. $[L^{(k)}]^{-1}$ is the (pseudo) inverse matrix of the diagonal regularization operator $L^{(k)}$, and its i^{th} diagonal element is $([L^{(k)}]^{-1})_{ii} = |\widehat{z}_i^{(k)}|^{\frac{1}{2}}$. Hence there are no numerical problems in linearizing the problem even when many wavelet coefficients are equal to or close to zero. Also if $\widehat{z}_i^{(k)} = 0$, then $([L^{(k)}]^{-1})_{ii} = 0$ and $\widehat{z}_i^{(k+1)} = 0$. Alternatively one can create an augmented linear system of equations by removing all columns from Q for which $\widehat{z}_i^{(k)}$ is zero or smaller than a certain threshold value. This means that at the $(k+1)^{th}$ iteration only the non-zero wavelet coefficients in $\widehat{z}^{(k)}$ from the k^{th} iteration are updated.

5.2.4 Convergence proof

Assume that the starting point for the algorithm is calculated as a least square solution such that $||Q\widehat{z}^{(0)} - y||_2^2 = P\sigma^2$ where P is the length of the measurement vector y. At $k \geq 1$, $\widehat{z}^{(k+1)}$ is obtained by solving equations 5.6 and 5.7. According to theorem 1 in Ref. [79], $||Lz(y,\lambda)||_2^2$ is a monotonically decreasing function of $||Qz(y,\lambda) - y||_2^2$ with λ as parameter. This means that $\min_\lambda ||L^{(k)}\widehat{z}(y,\lambda,k)||_2^2$ satisfying the constraint in Eq. 5.7 occurs when $||Q\widehat{z}(y,\lambda,k) - y||_2^2$ is at the maximum value of $P\sigma^2$ for all k. It is assumed for this analysis that we can choose any regularization parameter λ in a continuum between 0 and ∞ even though in practice we only choose a finite set of discrete λ values. So in order to prove that $C^{(k+1)} \leq C^{(k)}$, it suffices to prove that $\lambda^{(k+1)}||L^{(k)}\widehat{z}^{(k+1)}||_2^2 \leq \lambda^{(k)}||L^{(k-1)}\widehat{z}^{(k)}||_2^2$. If $\lambda^{(k+1)} = 0$, then there is no further regularization and $\widehat{z}^{(k+1)} = \widehat{z}^{(k)} \implies$ convergence. Suppose $\lambda^{(k+1)} > 0$. Then

$$\begin{aligned} ||L^{(k)}\widehat{z}^{(k+1)}||_2^2 &= ||L^{(k)}\widehat{z}(y, \lambda^{(k+1)})||_2^2 \\ &\leq ||L^{(k)}\widehat{z}(y, 0)||_2^2 \\ &= ||L^{(k-1)}\widehat{z}^{(k)}||_2^2 \end{aligned}$$

The inequality in the second line above is again because of Theorem 1 in Ref. [79]. So, as long as we choose the optimal regularization parameter $\lambda^{(k+1)} \leq \lambda^{(k)}$, the objective cost function is decreasing with k. Intuitively it is appealing because as the iteration increases, the estimate is driven towards the actual object and hence less regularization is required.

5.2.5 Trained Object Prior

It has been observed that the wavelet coefficients of natural images are heavy-tailed and centered around zero [71]. This behavior is well-approximated using generalized Gaussian density (GGD) statistics and this prior knowledge has been exploited in image coding to minimize the quantization noise on the wavelet representation [80].

A GGD function is parameterized by α and β according to

$$p(z;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|z|/\alpha)^{\beta}} \tag{5.10}$$

where $\Gamma(.)$ is the Gamma function with $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt,\ x > 0$. The scale parameter α controls the width of the function and the shape parameter β controls its exponential decay rate. Examples of this function are shown in Fig. 5.2 for two different α and four different β values. Notice that $\beta = 2$ refers to a Gaussian density and $\beta = 1$ to a Laplacian density. Generally wavelet coefficients in each subband of the wavelet decomposition of an object are modeled by one GGD function. An example of such a model is depicted in Fig. 5.3(b) for one wavelet subband of the 128×128-pixels natural object in Fig. 5.3(a). The histogram of 1024 wavelet coefficients associated with the horizontal subband at the 2^{nd} highest decomposition scale is shown using 64 bins. The model fit to the histogram is estimated using the least square criterion. The estimated parameters $\alpha = 0.383$ and $\beta = 0.685$. A L-level wavelet decomposition of an object produces $3L + 1$ subbands. An object can thus be modeled in the wavelet domain using as many GGD functions: $3L + 1$ number of α parameters and $3L + 1$ number of β parameters. Several authors have studied the correlation of GGD model parameters across different decomposition levels and/or different orientations at a particular scale level [29]. In such cases an object can be represented by using fewer than $6L + 2$ parameters.

Suppose that the object is modeled in the wavelet domain using L-levels of wavelet decompositions. The wavelet subband in the l^{th} scale and the m^{th} orientation is characterized by a GGD function with parameters α_{lm} and β_{lm} where $l = 1, 2, ..., L$ and $m = 1, 2, 3 \equiv \{LH, HL, HH\}$. This scheme is represented pictorially in Fig.

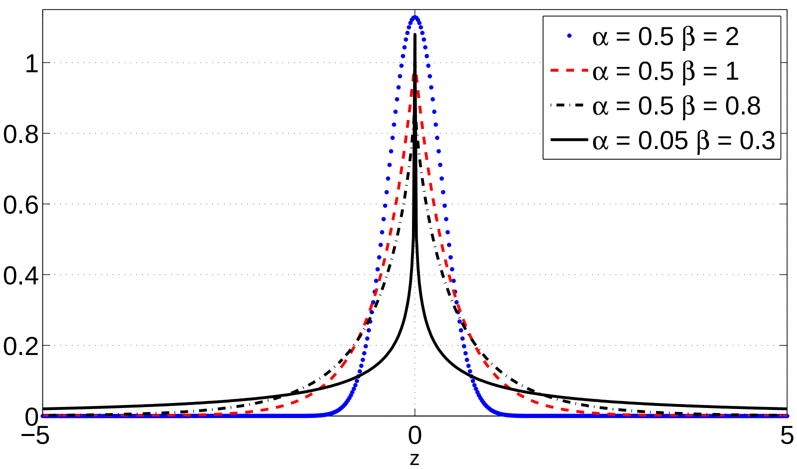


Figure 5.2: Examples of generalized Gaussian density functions. $\beta = 2$ refers to a Gaussian density and $\beta = 1$ refers to a Laplacian density.

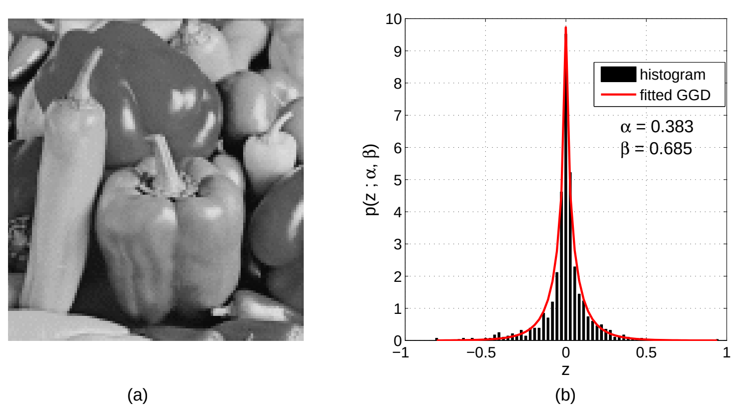


Figure 5.3: An example object model in the wavelet domain using a generalized Gaussian density. (a) A 128 $\times$ 128 -pixels natural object, and (b) histogram of wavelet coefficients from (a) in the LH subband of the 2^{nd} decomposition level along with a GGD fit. The fit parameters are $\alpha = 0.383$ and $\beta = 0.685$. The Daubechies-4 filter is used for computing the wavelet transform.

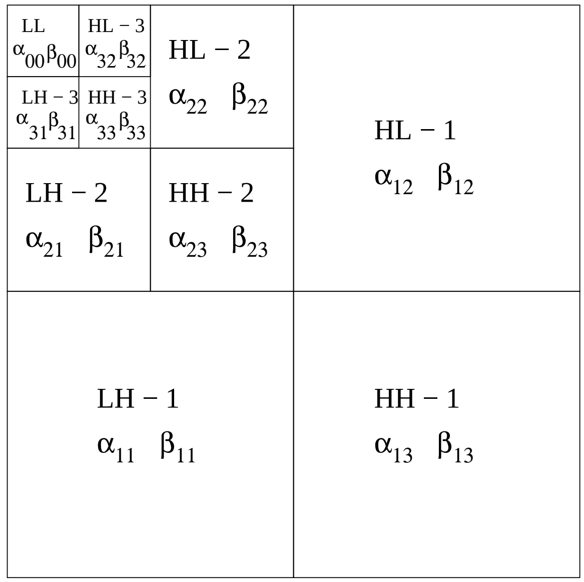


Figure 5.4: Schematic of a 3-level wavelet decomposition and the associated GGD model parameters for each subband.

5.4 for $L = 3$. The low pass LL band is also modeled by a GGD function with parameters α_{00} and β_{00}. Using the MAP formulation the object estimate is given by

$$\begin{aligned}\widehat{z} &= \max_z P(z|y) \\ &= \max_z \frac{P(y|z)P(z)}{P(y)} \quad \text{(Bayes rule)} \\ &= \max_z P(y|z)P(z) \quad \text{(y is fixed)}\end{aligned}$$

The probability of y conditioned on z is Gaussian with $P(y|z) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\{-\frac{||y-Qz||_2^2}{2\sigma^2}\}$. The wavelet coefficients are assumed to be independent of each other and the probability of z is given by

$$P(z) = \prod_{i=0}^{N-1} \frac{\widetilde{\beta}_i}{2\widetilde{\alpha}_i \Gamma(1/\widetilde{\beta}_i)} e^{-(|z_i|/\widetilde{\alpha}_i)^{\widetilde{\beta}_i}} \qquad (5.11)$$

where $\widetilde{\alpha_i}$ and $\widetilde{\beta_i}$ are determined by the $(l,m)^{th}$ subband to which the i^{th} coefficient belongs, i.e., $\widetilde{\alpha_i} = \alpha_{lm}$ and $\widetilde{\beta_i} = \beta_{lm}$. Taking the logarithm the MAP estimate is now obtained as a regularized estimate

$$\widehat{z} = \arg\min_z [\ ||y - Qz||_2^2 + \lambda \sum_{i=0}^{N-1} (|z_i|/\widetilde{\alpha_i})^{\widetilde{\beta_i}}\ ] \tag{5.12}$$

where λ is the regularization parameter that trades-off cost factors due to the data consistency, $||y - Qz||_2^2$, and the object prior, $\sum_{i=0}^{N-1} (|z_i|/\widetilde{\alpha_i})^{\widetilde{\beta_i}}$.

The GGD priors also promote sparsity. To elaborate this point, suppose that each wavelet subband is modeled using GGD functions with fixed $\beta = 1$. In this case the object estimate in Eq. 5.12 becomes $\widehat{z} = \min_z ||y - Qz||_2^2 + \lambda \sum_{i=0}^{N-1} \frac{|z_i|}{\widetilde{\alpha_i}}$. This is similar to Eq. 5.3 except that the i^{th} wavelet coefficient is now scaled by a factor $1/\widetilde{\alpha_i}$. The scaling factor is learned from a set of training images and depends on the subband of which the coefficient is a member. Because $\widetilde{\alpha_i}$'s are constants, the regularization cost reduces as the wavelet coefficients tend towards zero; thus yielding a sparse estimate. Furthermore the norm-like quantity $||z||_\beta = \sum_{i=0}^{N-1} |z_i|^\beta$ when $\beta < 1$ is also used as a metric to quantify the sparsity of an object [78, 81, 82]. $||z||_\beta$ is often referred to as a "diversity measure" or an "entropy-like measure" because it is not a true norm for $\beta < 1$ [78]. Minimization of cost functions of the form $\widehat{z} = \min_z ||y - Qz||_2^2 + \lambda ||z||_\beta$ for $\beta < 1$ results in sparse solutions as well [78]. The same idea can be extended to cases where each wavelet coefficient is scaled by a constant to yield a regularization cost in the form $\sum_{i=0}^{N-1} (|z_i|/\widetilde{\alpha_i})^\beta$. Notice that Eq. 5.12 can be considered as a further adaptation of the diversity measures to compute sparse solutions. We solve Eq. 5.12 using the same algorithm outlined in section 5.2.2. At the k^{th} iteration the algorithm solves Eq. 5.6, 5.7, and 5.8 to update the estimate. The resulting regularization operator at the k^{th} iteration, $L^{(k)}$, is again diagonal. The i^{th} diagonal element is equal to $(\alpha^{\beta/2}\ |z_i|^{1-\beta/2})^{-1}$. We refer to this algorithm as "GGD penalty." The convergence proof stated in section 5.2.4 holds for this object prior as well. In the next section we present the performance of the various algorithms described so far.

5.3 Simulation Results

We consider a distributed imaging system consisting of K low-resolution cameras. Each of them is assumed to be an incoherent optical system with a rectangular aperture. H_{opt} in the imaging equation Eq. 2.2 is then determined by the diffraction-limited optical point spread function, h_{opt}. The degree of optical blur w is defined as the half-width of h_{opt} which is the separation between the peak and the first zero for h_{opt}. Fig. 4.6 is a 1D depiction of $h_{opt}(x) = \mathrm{sinc}^2(\frac{x}{w})$ for $w = 0.5$. Larger w corresponds to larger diffraction-limited optical blur. The operators H_{ccd} and S are determined by the pitch of the measurement pixel, D, relative to the object pixel. The geometrical transformation operator, T_k, in Eq. 2.2 models the particular view of the object in the k^{th} LR image, and is characterized by scale (s), rotation (θ), and translations (Δx, Δy) in two image coordinates. These parameters are chosen uniformly and at random from the ranges of values: $0.9 \leq s \leq 1.1$, $-15 \leq \theta \leq +15$ degrees and $-D \leq (\Delta x, \Delta y) \leq +D$ object pixels. The measurement SNR is defined as $SNR(dB) = 10\log_{10}\frac{1}{\sigma^2}$ where σ^2 is the variance of the sensor AWGN and the object is gray-scale with pixels values normalized between 0 and 1. The reconstruction fidelity is quantified by the root mean square error (RMSE) that is defined as $RMSE = \sqrt{\frac{\sum_{i\in N_c}(f_i-\widehat{f}_i)^2}{|N_c|}}$ where $\widehat{f}$ is the object estimate, N_c is the set of pixels that are seen by all K cameras, and $|N_c|$ is the total number of such

Table 5.1: The GGD function parameters that are obtained by fitting the histograms of wavelet coefficients of training objects shown in Fig. 5.5. The value of $z_{min} = 0.0238$. The LL subband is modeled by using a GGD density having parameters $\alpha = 2.479$ and $\beta = 1$.

	LH		HL		HH	
	α	β	α	β	α	β
$l = 1$	4.65×10^{-4}	0.319	1.04×10^{-4}	0.270	1.64×10^{-3}	0.452
$l = 2$	2.45×10^{-3}	0.341	6.17×10^{-4}	0.273	1.21×10^{-3}	0.349
$l = 3$	0.125	0.613	3.47×10^{-2}	0.434	1.79×10^{-2}	0.432
$l = 4$	0.249	0.667	0.112	0.537	0.174	0.677

common pixels. We also analyze the reconstructions qualitatively using a visually weighted RMSE (VRMSE). We calculate the VRMSE by using a filter proposed in Ref. [83]. While calculating this metric, the filter weighs different frequencies by different amounts based on the empirical studies of human visual systems. VRMSE has been shown to provide a good estimate of subjective visual quality.

5.3.1 Training objects and model adequacy

To obtain prior knowledge about the object we use 32 natural images of size 128×128-pixels from the USC-SIPI database as training objects [84]. These images are often used for data compression related research activity. We pre-process these objects by compressing them to have a distortion root mean squared error of 1%. After pre-processing we obtain a set of moderately sparse, i.e., $||z||_0 \leq 50\%$, images. 26 such objects are shown in Fig. 5.5. We estimate the mean object vector and the covariance matrix required by the LMMSE operator using this training set. We assumed a correlation distance of 2 pixels: i.e., two positions that are 2 pixels apart in any direction are assumed to be uncorrelated.

For the GGD method we model wavelet subbands at four decomposition levels and 3 orientations (horizontal LH, vertical HL, and diagonal HH), and the low-pass

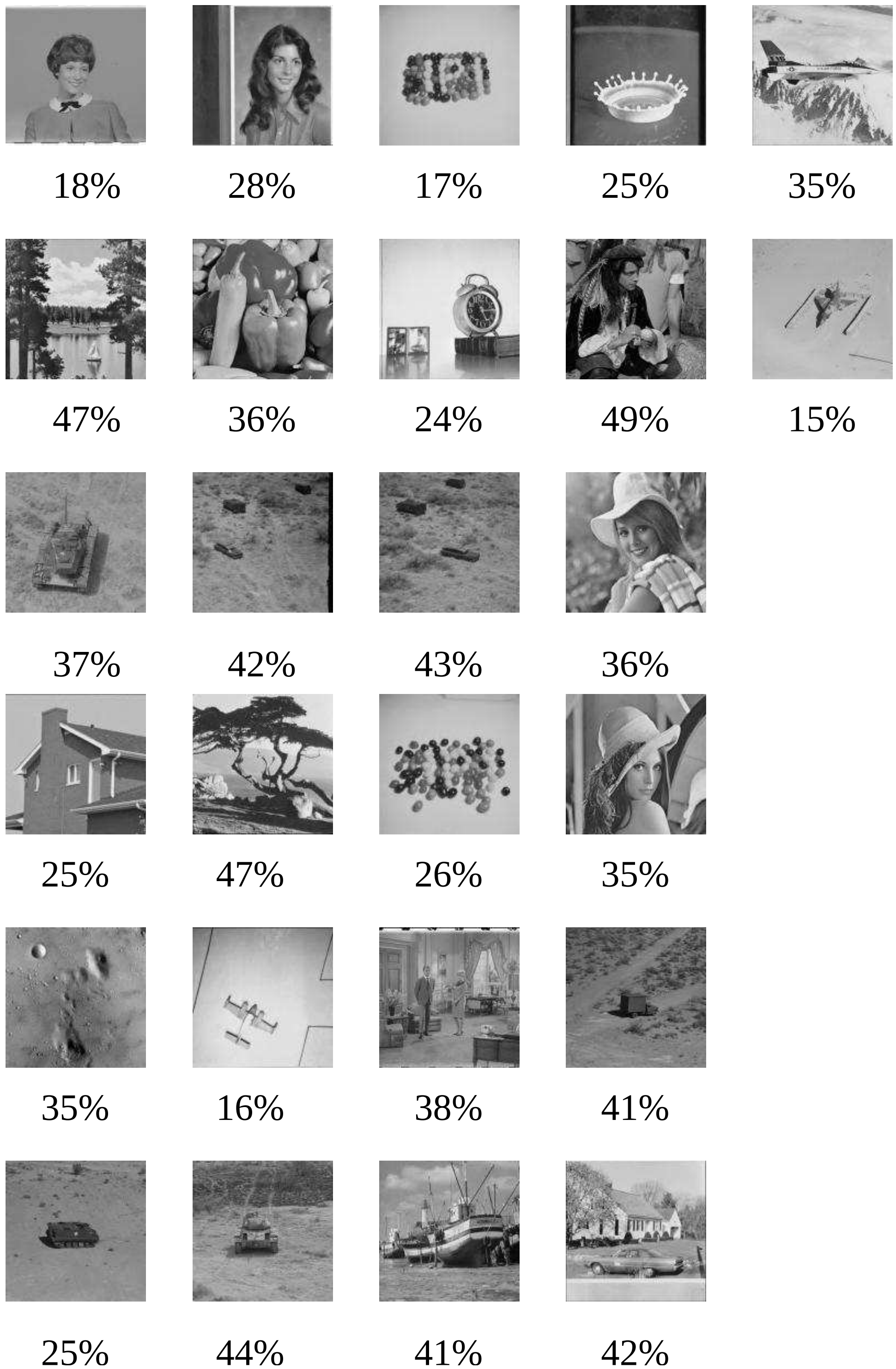


Figure 5.5: A set of 26 sparse objects of size 128×128 pixels used for training. The number below each object is the percentage of non-zero wavelet coefficients.

LL subband. The wavelet coefficients belonging to each subband are collected from all training objects. We then fit the histogram of wavelet coefficients with a GGD function. The parameters α and β are determined by minimizing the squared error between the histogram and the model. The values of α and β for all wavelet subbands are tabulated in Table 5.1. We notice that the variation of α with decomposition levels is consistent with previously published results [29]: the width of the GGD function generally increases as the decomposition level increases. In addition to that we observe increasing β values with increasing level number for a given orientation. When the objects are very sparse, e.g. $||z||_0 \leq 30\%$, the wavelet coefficients in subbands in the 1^{st} and 2^{nd} decomposition levels (at highest frequency scales) are mostly zeros with occasional spikes of non-zero elements. An example histogram of such a sparse subband is shown in Fig. 5.6. We observed that in such cases a GGD function is not a good fit to the data: the estimated parameter $\alpha \approx 0$. During reconstruction using Eq. 5.12, the presence of $\alpha = 0$ forces all coefficients in that particular subband of the estimate, $\widehat{z}$, to zero. Thus the estimate ignores any rare spikes that may correspond to strong edge details in the object. To avoid this problem we choose to model by fitting only the non-zero coefficients with a GGD function. We learn an additional parameter z_{min}, the smallest absolute wavelet coefficient in that subband. However instead of learning the z_{min} in every subband we choose the smallest wavelet coefficient of all training wavelet coefficients and use it as the common parameter. An example of a GGD function fitted to histogram in a sparser subband is shown in Fig. 5.6. We use these GGD parameters and z_{min} in the GGD penalty restoration algorithm. The regularization operator is calculated in a similar way as described in section 5.2.5.

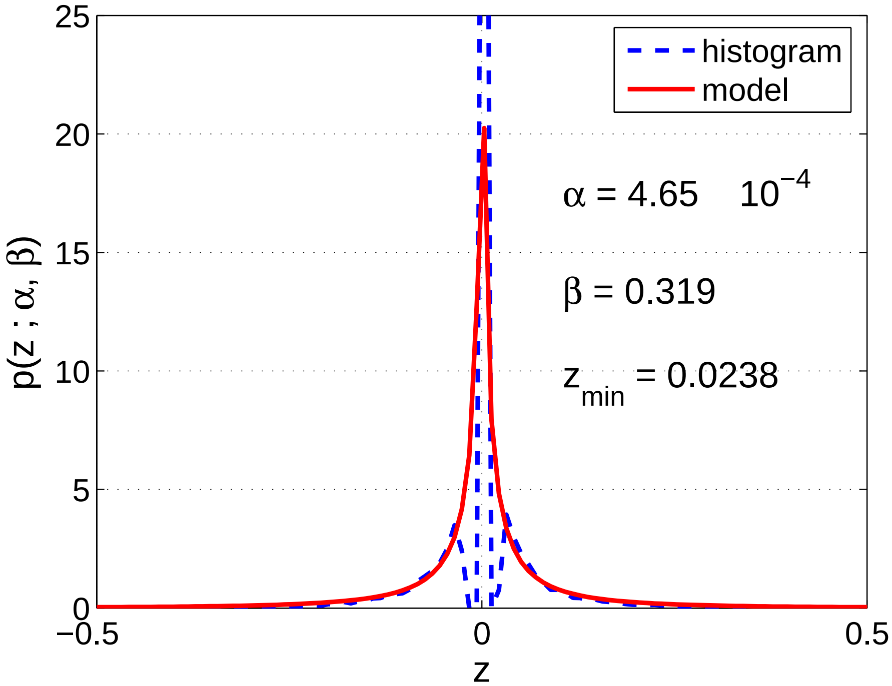


Figure 5.6: An example demonstrating GGD model inadequacy for a very sparse wavelet subband. Only non-zero wavelet coefficients whose absolute value is greater than or equal to z_{min} are used to obtain a GGD function fit. The knowledge of z_{min} is also used in the sparse reconstruction algorithm.

5.3.2 Reconstruction Fidelity

We employ Monte-Carlo simulations to test the performance of each algorithm. The threshold value in the EM algorithm is set to $\tau = 3\sigma$ as in Ref. [67]. In the first iteration of ℓ_1 penalty and GGD penalty the optimal regularization parameter is chosen from the set of λ's such that $\log_{10} \lambda = \{-3, -2.99, ..., -0.01, 0\}$. In the subsequent iterations a search for the optimal parameters is conducted in the vicinity of the optimal parameter of the previous iteration. Specifically $\lambda^{(k+1)}$ is found in the set of $\lambda \in (\frac{\lambda^{(k)}}{5},\ 5\lambda^{(k)})$. The convergence threshold is chosen to be $\epsilon = 10^{-2}$. All EM, ℓ_1 penalty and GGD penalty methods are initialized to the LMMSE estimate.

First we consider a distributed imaging system with $D = 2$ pixel down-sampling and $w = 0.5$ degree of optical blur. Fig. 5.7 shows example reconstructions of a sparse object using the algorithms outlined in the previous sections. The 128×128-pixels original sparse object is shown in Fig. 5.7(a). This object is 75% sparse in the wavelet domain: i.e., $\ell_0 = 25\%$. Low-resolution images from three cameras at SNR = 20dB are shown in Fig. 5.7(b), (c), and (d). These images are interpolated to original size using a pixel replication method. The second row of images in Fig. 5.7(e), (f), (g), and (h) are the reconstructions using $K = 2$ LR images using LMMSE, EM, ℓ_1 penalty, and GGD penalty methods respectively. The RMSE values of these reconstructions are 6.33%, 6.49%, 5.97%, and 6.08% respectively. Because we are incorporating the object sparsity constraints, we also quantify the reconstruction fidelity using the ℓ_0 norm of reconstructed images. The ℓ_0 norms of these reconstructions are 99.99%, 1.92%, 39.98%, and 9.31% respectively. The ℓ_0 norm is calculated by counting all wavelet coefficients that have an absolute value greater than 10^{-6}. As expected the LMMSE does not produce a sparse solution. All other methods provide sparse solutions with different levels of sparsity. The

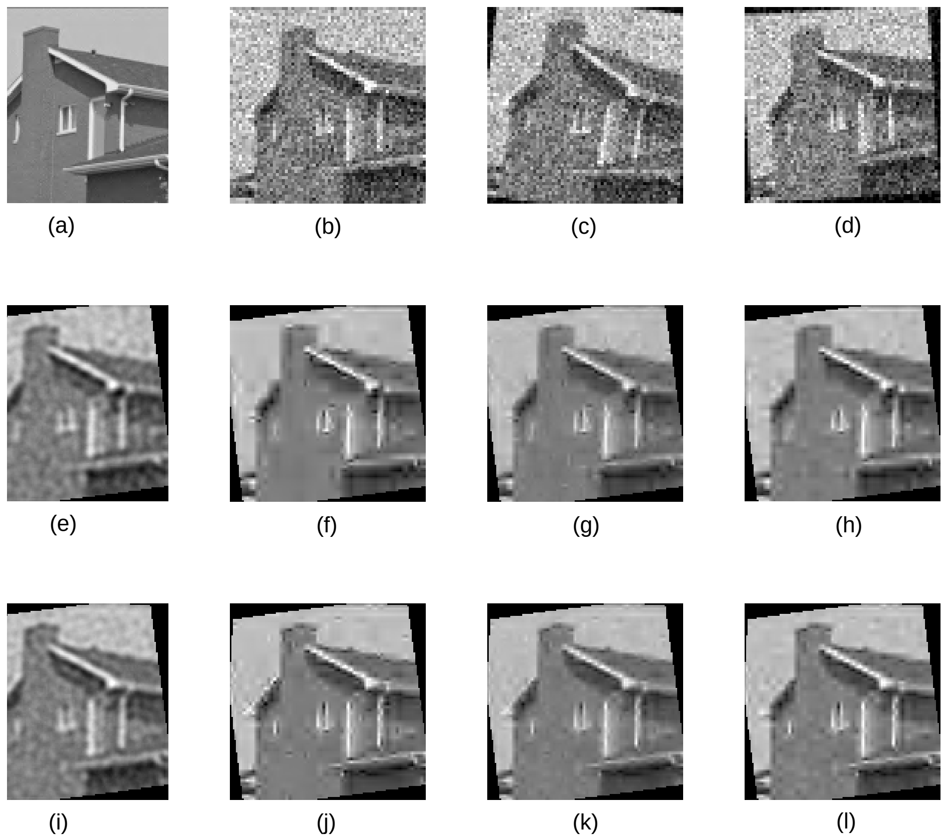


Figure 5.7: Sample restored images at $SNR = 20$dB by using algorithms described in this chapter. (a) A 128×128-pixel object, and (b)-(d) three LR images that degraded by optical blur with $w = 0.5$ and $D = 2$ pixel down-sampling. Images restored from $K = 2$ LR frames and using the (e) LMMSE, (f) EM, (g) ℓ_1 penalty, and (h) GGD penalty algorithms. They have RMSE values of 6.33%, 6.49%, 5.97%, and 6.08% respectively. Images in (i), (j), (k), and (l) are reconstructed from $K = 4$ LR frames using LMMSE, EM, ℓ_1 and GGD penalty respectively. They have RMSE values of 5.75%, 5.73%, 5.31%, and 5.43% respectively.

EM and GGD penalty methods produce sparser reconstructions than ℓ_1 penalty. The LMMSE estimate appears noisy whereas all sparsity-based approaches produce de-noised estimates. The visually weighted RMSE values are 0.70%, 0.71%, 0.69%, and 0.59% respectively. Object estimates restored by LMMSE, EM, ℓ_1, and GGD penalty methods using $K = 4$ LR images are shown in Fig. 5.7(i)-(l) respectively. Reconstruction fidelity of these images, both visually and quantitatively, are better than those in the second row. These restorations have RMSE values of 5.75%, 5.73%, 5.31%, and 5.43% respectively, and VRMSE values of 0.56%, 0.55%, 0.61%, and 0.50% respectively. Sparsity levels (ℓ_0 norms) of these reconstructions are 100%, 3.34%, 72.03%, and 11.46% respectively. All three wavelet-based methods produce de-noised reconstructions while the LMMSE estimate appears more noisy than other methods.

We illustrate the convergence of the doubly-iterative algorithm in Fig. 5.8 for the case of $D = 2$ pixel down-samping, $K = 2$ LR images, and $SNR = 20$dB. The ℓ_1 penalty method converges in 5 iterations and GGD penalty converges in 3. The convergence threshold $\epsilon = 0.01$ is used here. The total objective cost ($\sqrt{C^{(k)}}$), the regularization cost ($\lambda^{(k)}||L^{(k)}\widehat{z}^{(k)}||_2$), and the estimation error ($||z - \widehat{z}^{(k)}||_2$) are plotted in Fig. 5.8(a) for both methods. The objective cost and regularization cost monotonically reduce as the iterations evolve. This may not be true for the estimation error (and RMSE): it may increase after reaching a minimum value. Often the estimate deviates from the true object in mean square error sense while both the regularization cost and the total objective cost reduce. Fig. 5.8(b) plots the regularization parameter versus iteration number for both methods. The λ values reduce with iteration for both methods. The sparsity of the estimate, in terms of ℓ_0 norm, at each iteration is plotted in Fig. 5.8(c). As expected the sparsity of the

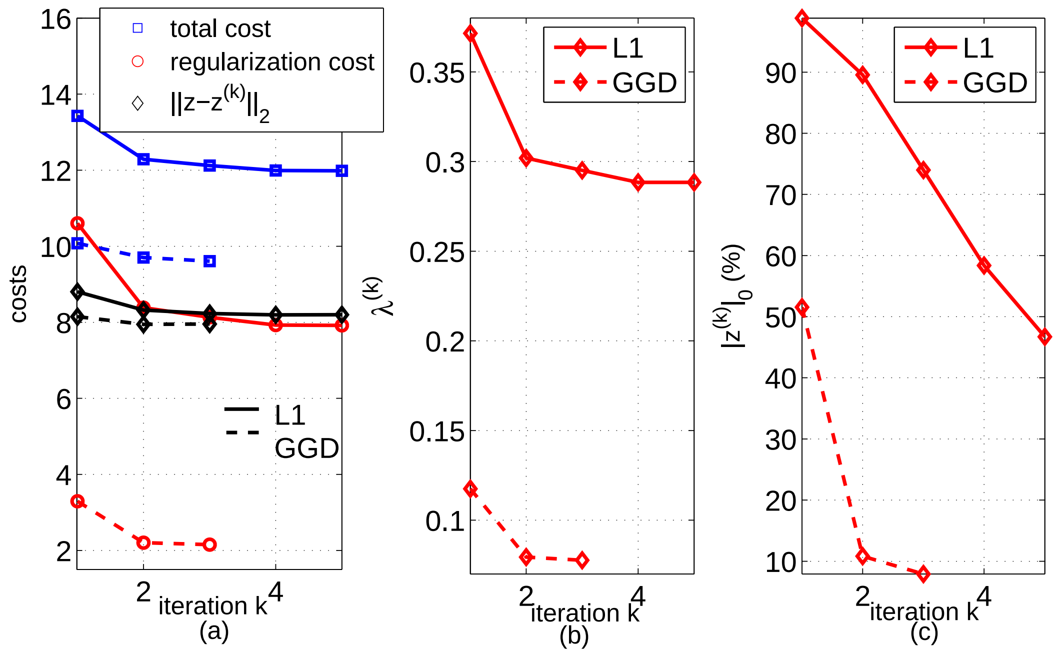


Figure 5.8: An example of the evolution of the doubly-iterative algorithm outlined section 5.2.2 for ℓ_1 penalty (solid line) and GGD penalty (dashed line). (a) Total objective cost $(\sqrt{C^{(k)}})$, regularization cost $(\lambda^{(k)}||L^{(k)}\widehat{z}^{(k)}||_2)$, and estimation error $(||z - \widehat{z}^{(k)}||_2)$ versus iteration number, k, (b) regularization parameters versus k, and (c) L_0 (%) norm of estimates versus k.

estimate increases with the iteration number: the GGD penalty method produces sparser object estimates than does the ℓ_1 penalty method.

The performance of these algorithms over a range of SNR values is quantified in Fig. 5.9. The reconstruction RMSE is plotted versus SNR for the case of $D = 2$ pixel down-sampling and $w = 0.5$ degree of optical blur. The results for $K = 1$, 2, 3, and 4 LR images are shown in Fig. 5.9(a), (b), (c), and (d) respectively. The objects from Fig. 5.5 that are at least 70% sparse, i.e., $\ell_0 \leq 30\%$, are used to generate these performance results. Basic observations are: (1) The RMSE of estimates produced by any algorithm reduce as measurement SNR increases. (2) At a given SNR the reconstruction RMSE reduces with the number of LR images. The ℓ_1 penalty and

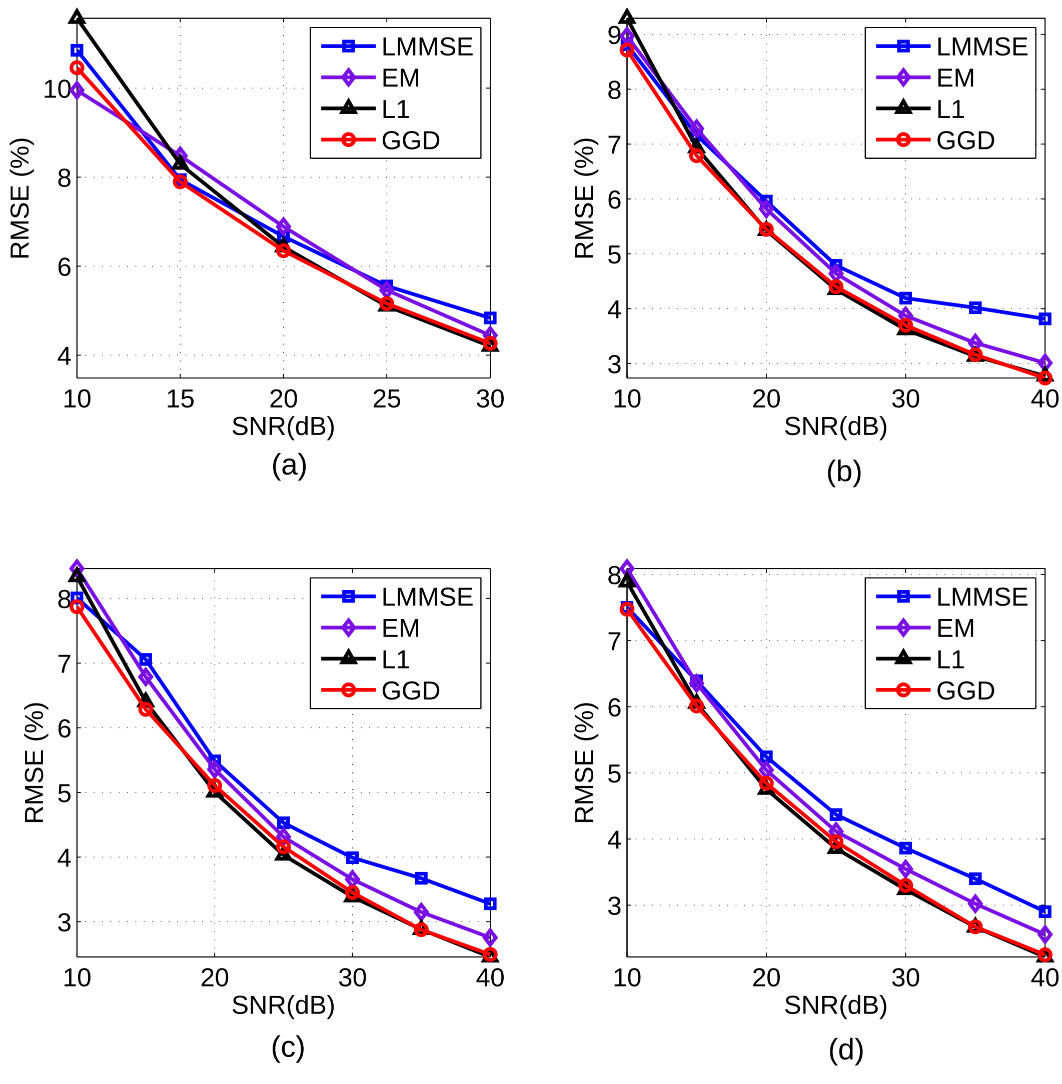


Figure 5.9: The plot of reconstruction RMSE versus SNR for the case $D = 2$ pixel down-sampling, $w = 0.5$ degree of optical blur, and (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, and (d) $K = 4$ low-resolution images. The RMSE is calculated only over the pixels that are seen in all K images.

GGD penalty methods produce reconstruction with smaller RMSE than the EM method and LMMSE method at SNR values greater than 18dB for all K. For example the reconstruction RMSE obtained by these algorithms is 10% smaller than that of the EM algorithm and 20% smaller than that of LMMSE at $SNR = 30$dB and $K = 4$ LR images. At low SNR values the GGD penalty method gains by the object-specific training. The RMSE of the GGD penalty algorithm is smaller than that of the ℓ_1 penalty method in the SNR range of $10 - 15$dB. We also analyzed the visually weighed RMSE of these reconstructions. VRMSE is plotted versus SNR for $K = 1$, 2, 3, and 4 in Fig. 5.10(a), (b), (c), and (d) respectively. The VRMSE is consistently the smallest for the GGD penalty method over a wide range of SNR and K. The LMMSE produces estimates having significantly higher VRMSE values for $K = 1$, 2, and 3. The ℓ_1 penalty and EM estimates have comparable VRMSE values. We also examined the sparsity of these reconstructions. The ℓ_0 norms of the reconstructions are plotted against SNR in Fig. 5.11(a)-(d) for $K =$ 1-4 LR images respectively. The spatial domain LMMSE method produces the least sparse estimates at all SNR levels and values of K. The average ℓ_0 norm of 10 objects considered here (whose $\ell_0 \leq 30\%$) is 21.60%. The estimates produced by the ℓ_1 penalty method for different K and SNR values considered here have ℓ_0 values in the range of 60-90%. The ℓ_0 norms of EM and GGD penalty estimates are smaller than the average value. The EM method produces the sparsest estimates. These two methods produce sparser estimates because they force wavelet coefficients towards zero explicitly either by soft-thresholding (τ in EM) or hard-thresholding (z_{min} and α's in GGD penalty). Another trend is that the sparsity levels in the estimates get closer to their actual values as the SNR increases.

Now we consider a distributed imaging system with a larger $D = 4$ pixel down-

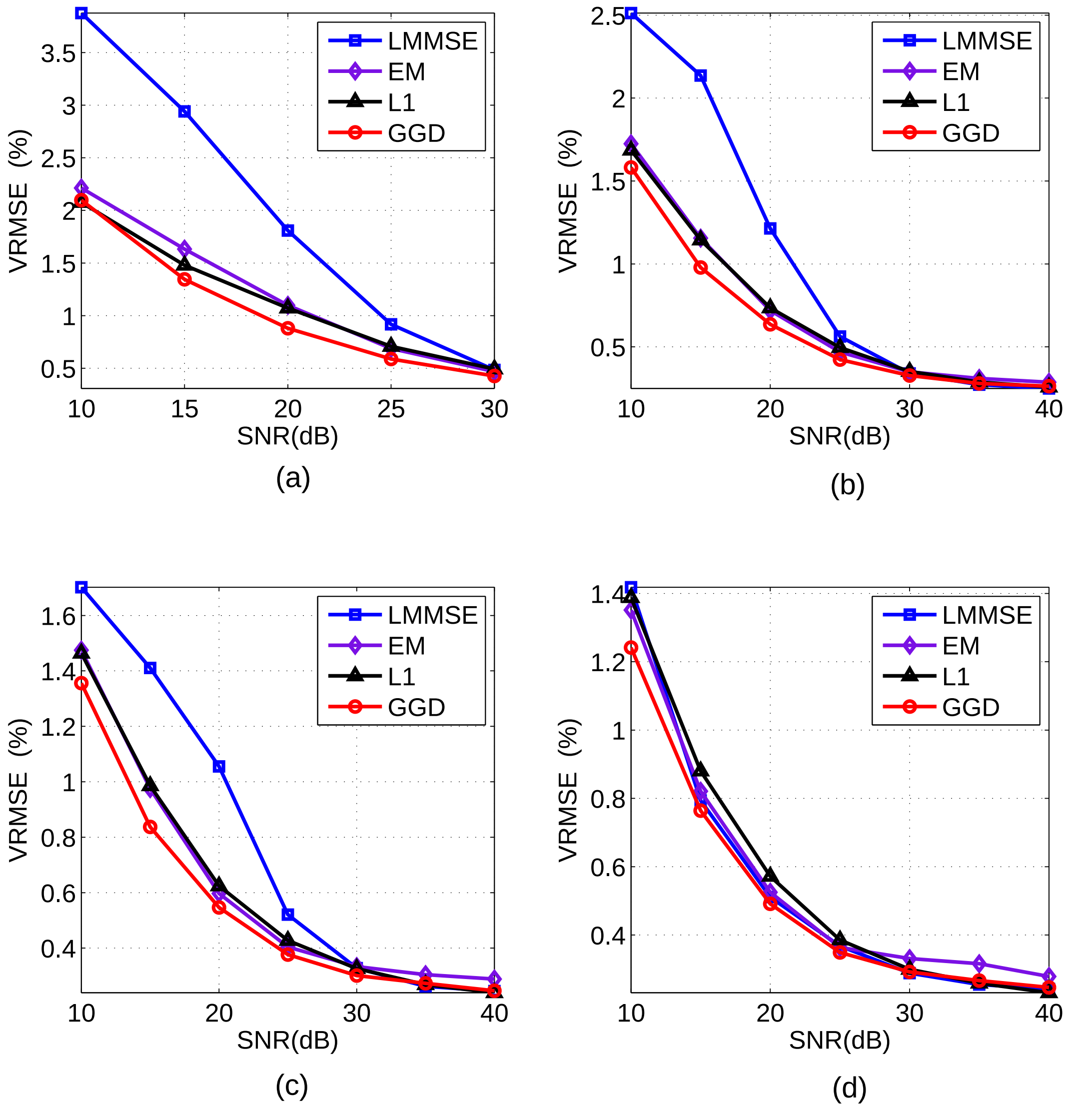


Figure 5.10: The plot of visually weighted RMSE versus SNR for the case $D = 2$ pixel down-sampling, $w = 0.5$ degree of optical blur, and (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, and (d) $K = 4$ low-resolution images.

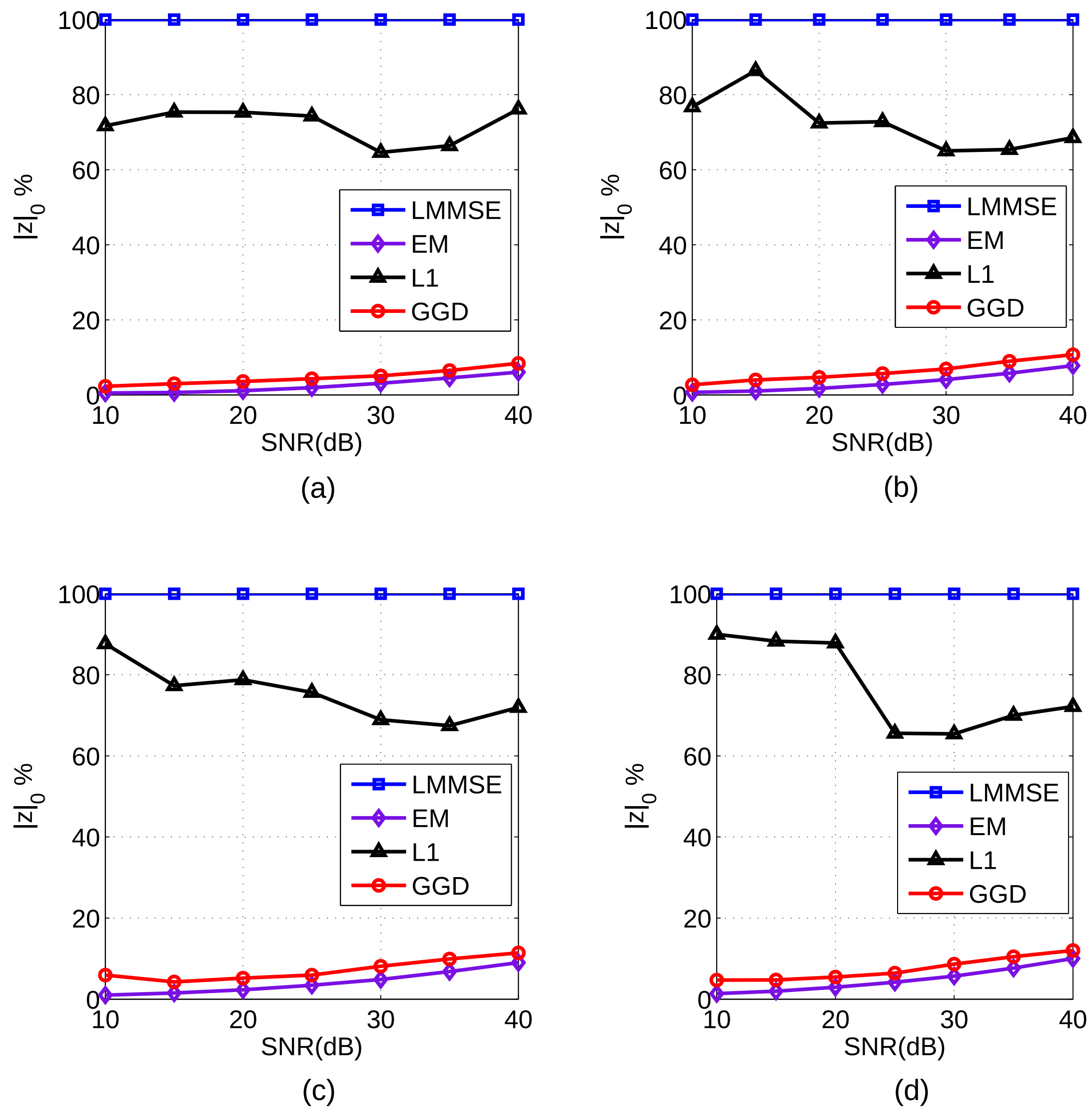


Figure 5.11: The plot of L_0 norm (%) of reconstructions versus SNR for the case $D = 2$ pixel down-sampling, $w = 0.5$ degree of optical blur, and (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, and (d) $K = 4$ low-resolution images.

sampling and $w = 1.0$ degree of optical blur. Fig. 5.12 shows example reconstructions of a sparse object. The 128×128-pixels original object and three low-resolution images measured at $SNR = 50$dB appear in Fig. 5.12(a) and (b)-(d) respectively. The object is 76% sparse: i.e., $\ell_0 = 24\%$. The second and the third rows display object estimates obtained by using $K = 4$ and $K = 16$ LR images respectively. The images in (e), (f), (g), and (h) are reconstructed from $K = 4$ LR images by using the LMMSE, EM, ℓ_1 penalty, and GGD penalty methods respectively. They have RMSE values of 5.81%, 7.06%, 4.92%, and 5.09% respectively and VRMSE values of 0.48%, 0.45%, 0.37%, and 0.39% respectively. The ℓ_0 norms of these reconstructions are 99.98%, 5.85%, 56.10%, and 9.59% respectively. The third row images in (i), (j), (k), and (l) respectively are reconstructed using the LMMSE, EM, ℓ_1 penalty, and GGD penalty methods and $K = 12$ LR images. The images in the third row ($K = 12$) have better reconstruction fidelity (both quantitative and visual) than those in the second row ($K = 4$). They have RMSE values of 4.50%, 5.15%, 3.71%, and 3.71% respectively and VRMSE values of 0.47%, 0.35%, 0.30%, and 0.30% respectively. The ℓ_0 norms of these reconstructions are 100%, 8.24%, 55.75%, and 13.42% respectively.

The performance of these algorithms over a range of SNR values is quantified in Fig. 5.13. The reconstruction RMSE is plotted versus SNR for the case of $D = 4$ pixel down-sampling and $w = 1.0$ degree of optical blur. The results for $K = 4$, 8, and 12 LR images are shown in Fig. 5.13(a), (b), and (c) respectively. Basic observations made for the case of $D = 2$ hold for the case of $D = 4$. The reconstruction RMSE of the ℓ_1 method is consistently the smallest for all K and SNR values considered. At high SNR values ($>$ 42dB for $K = 8$ and $>$ 40dB for $K = 12$) the performances of GGD penalty is similar to that of ℓ_1 penalty. For

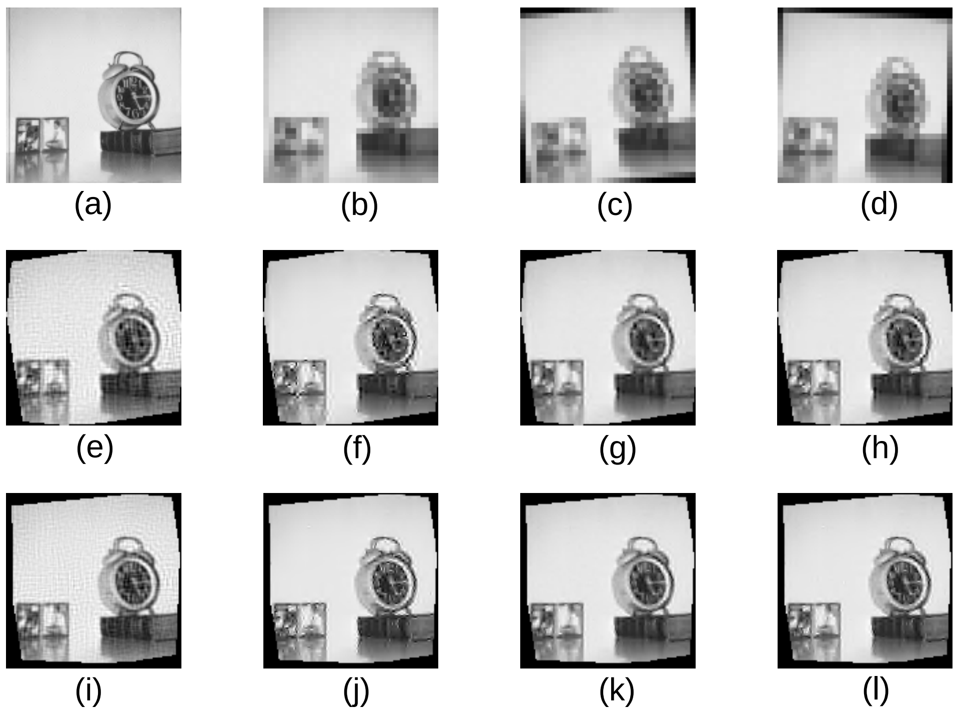


Figure 5.12: Sample restored images at $SNR = 50$dB by using algorithms described in this chapter. (a) A 128×128-pixel object, and (b)-(d) three LR images that are degraded by optical blur with $w = 1.0$ blur and $D = 4$ pixel down-sampling. Object estimates from $K = 4$ such LR images using the (e) LMMSE, (f) EM, (g) ℓ_1 penalty, and (h) GGD penalty algorithms. Images in (e), (f), (g), and (h) have RMSE values of 5.81%, 7.06%, 4.92%, and 5.09% respectively. The estimates obtained from $K = 12$ LR images by using the (i) LMMSE, (j) EM, (k) ℓ_1, and (l) GGD penalty methods. They have RMSE values of 4.50%, 5.15%, 3.71%, and 3.71% respectively.

example the reconstruction RMSE obtained by these methods is 5.5% smaller than that of the EM algorithm and 14.3% smaller than that of LMMSE at $SNR = 50dB$ and $K = 8$ LR images. VRMSE is plotted versus SNR for $K = 4$, 8, and 12 in Fig. 5.14(a), (b), and (d) respectively. The VRMSE is the smallest for the GGD penalty method over a wide range of SNR and K. At high SNR the VRMSE values obtained by all algorithms are almost the same. The ℓ_0 norms of the reconstructions from $K = 4$, 8, and 12 LR images are plotted against SNR in Fig. 5.15(a), (b) and (c) respectively. As in the case of $D = 2$ the spatial domain LMMSE method produces the least sparse estimates. The ℓ_1 penalty method produces estimates that have ℓ_0 values in the range of 30-70%. The ℓ_0 norms of the EM and GGD penalty estimates are smaller than the average value of 21.60%. As before the EM method produces the sparsest estimates. The sparsity levels of the estimates get closer to their true values as K and SNR increase. The performance of all wavelet-domain methods will depend on the quality of the initial estimate. This is more critical when thresholding is involved because it determines the support set (locations of non-zero wavelet coefficients) of the estimate in the next iteration. The initial estimate given by the LMMSE is closer to the object and has better reconstruction fidelity in the $D = 2$ cases. Hence we see significant improvements in reconstruction RMSE by ℓ_1 and GGD penalty methods in the $D = 2$ case than $D = 4$. In the next section we apply these algorithms to combine images captured in laboratory experiments.

5.4 Experiments

Though the main purpose of this chapter is to present a new multiframe superresolution algorithm we briefly demonstrate the application of the algorithm on real images. We employed an array of inexpensive Firewire-based cameras in an experi-

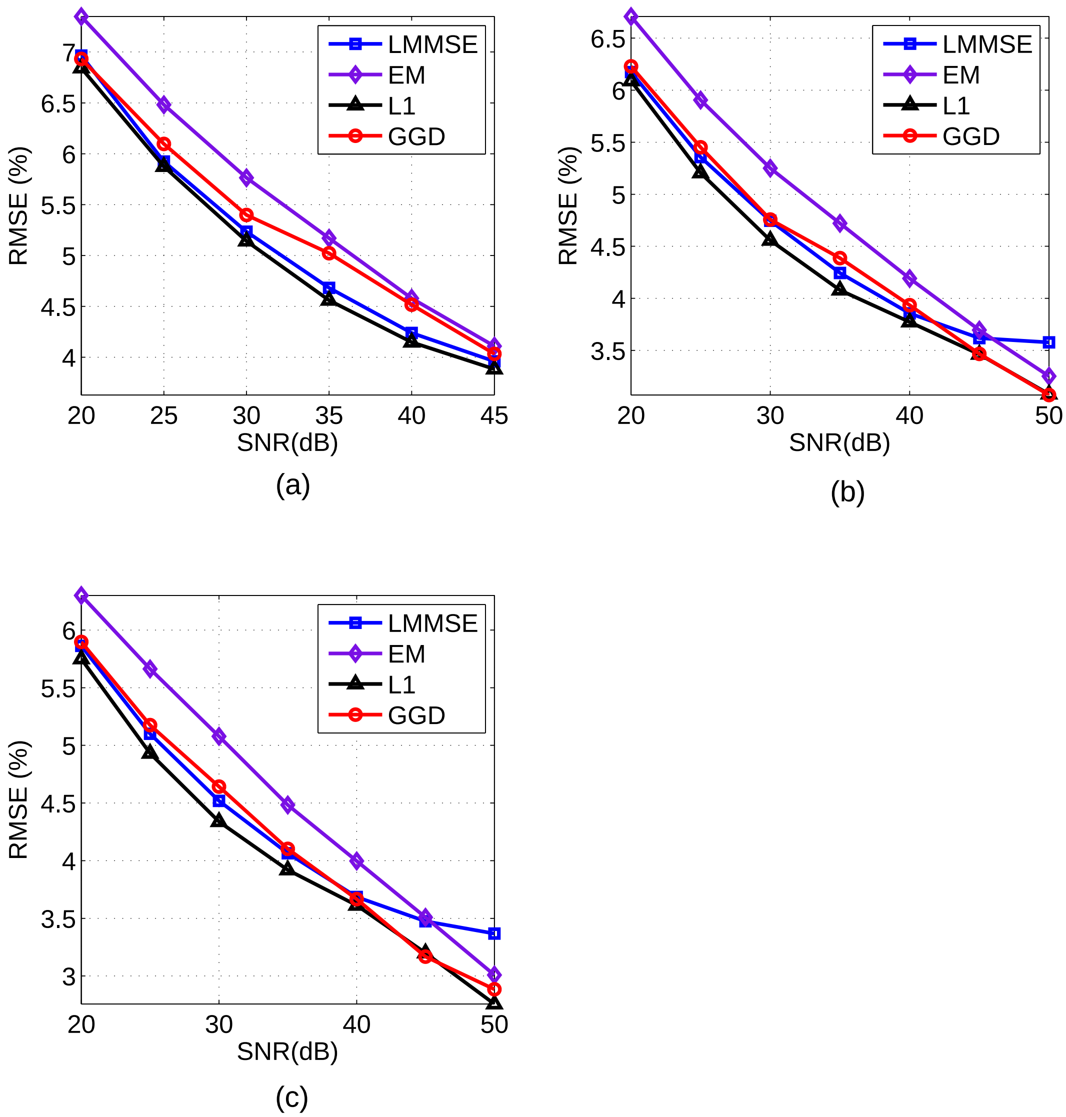


Figure 5.13: The plot of reconstruction RMSE versus SNR for the case $D = 4$ pixel down-sampling, $w = 1.0$ degree of optical blur, and (a) $K = 4$, (b) $K = 8$, and (c) $K = 12$ low-resolution images.

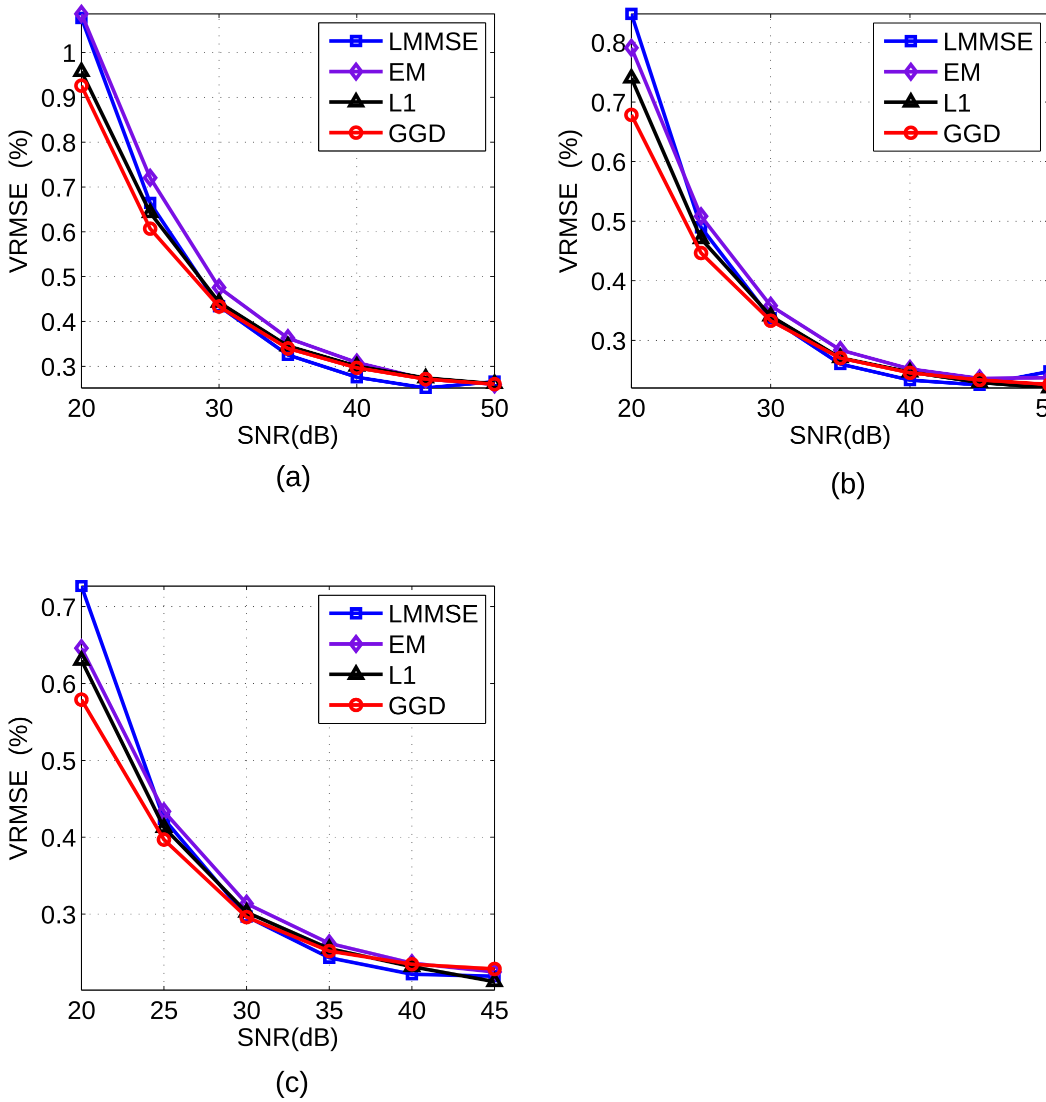


Figure 5.14: The plot of visually weighted RMSE versus SNR for the case $D = 4$ pixel down-sampling, $w = 1.0$ degree of optical blur, and (a) $K = 4$, (b) $K = 8$, and (c) $K = 12$ low-resolution images.

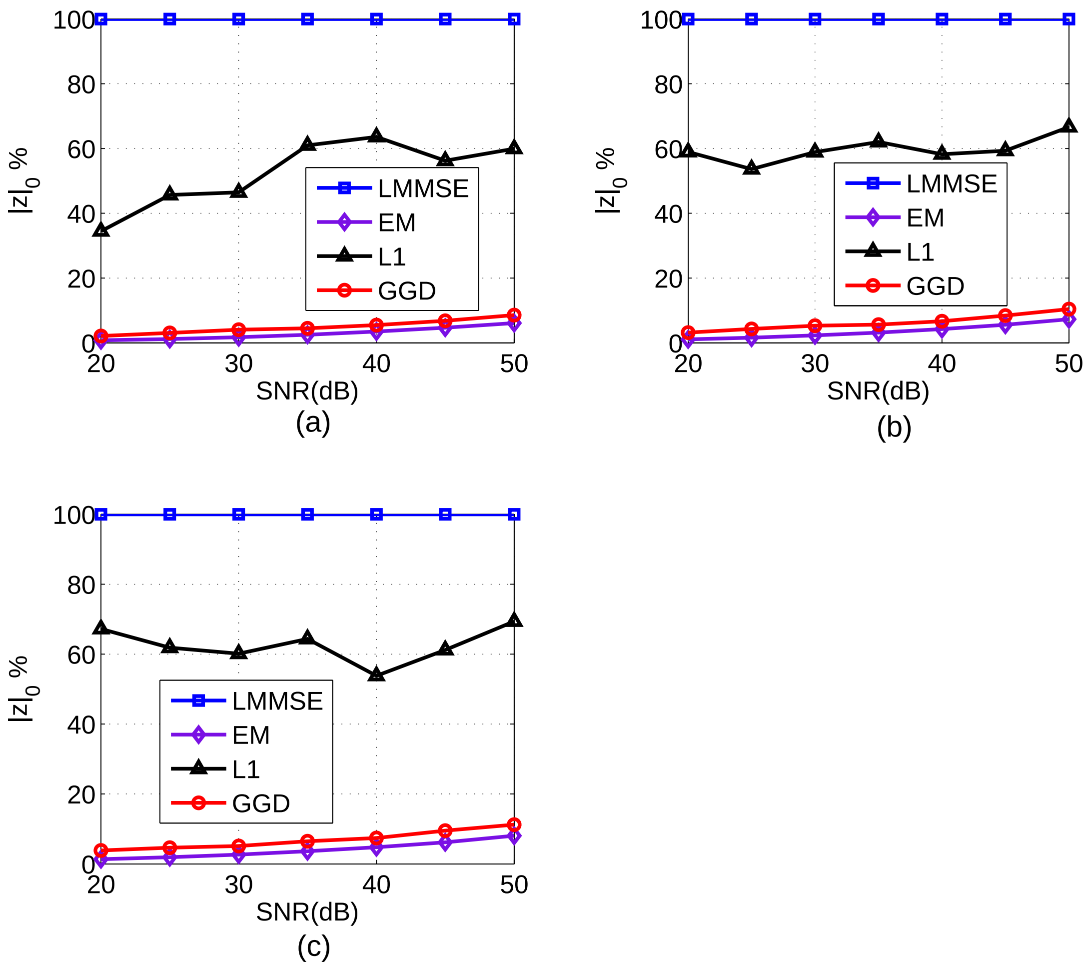


Figure 5.15: The plot of L_0 norm (%) of reconstruction versus SNR for the case $D = 4$ pixel down-sampling, $w = 1.0$ degree of optical blur, and (a) $K = 4$, (b) $K = 8$, and (c) $K = 12$ low-resolution images.

mental setup as shown in Fig. 5.16. A plasma monitor is used to display objects at a distance of 1.5 meters from the cameras. The cameras are deployed such that their geometries with respect to a reference camera can be modeled using affine transformations. Unlike the imaging systems considered in the previous section, accurate knowledge of all camera parameters is not available. The point spread functions (PSF) of these cameras also need to be estimated to characterize the degradation introduced by the optical blur. We place point source objects at multiple locations in the object space and capture the images from each camera. We estimate the PSF from these image measurements assuming that the response of each camera is shift-invariant. The PSFs thus estimated for two cameras are shown in Fig. 5.17. Coarse values of affine camera geometry parameters (i.e., scale, rotation, and shifts in two image coordinates) are estimated by aligning the image with respect to the reference image. Image registration is performed by choosing a set of corresponding points in the reference image and the image to be aligned [85]. These parameters are further fine-tuned by minimizing mean square error between the reference image and registered image.

We demonstrate the sparsity-constrained regularization method first by reconstructing an object that is sparse in the space-domain. Fig. 5.18(a) shows a measured LR image that is downsampled by $D = 2$ pixels. The images in Fig. 5.18(b) and (c) are reconstructed from $K = 4$ such LR images using least square (minimum L2 norm) estimation and the ℓ_1 penalty method respectively. Note that because the object is sparse in the space-domain, the operator $Q = HI^T = H$ in equations 5.5-5.7 where I is the identity matrix. Fig. 5.18(d) shows a 4×4 pixel downsampled image. Fig. 5.18(e) and (f) show object reconstructions from $K = 4$ such LR images using least squares estimation and the ℓ_1 penalty method respectively. ℓ_1 penalty

Figure 5.16: A snapshot of the experimental setup. The cameras are arranged in an affine geometry and the objects are displayed on the plasma monitor.

.0661	.0835	.0396
.1714	.2305	.1292
.0838	.1223	.0737

.0452	.0828	.0549
.1641	.2877	.1817
.0504	.0842	.0491

Figure 5.17: Estimated point spread functions of two cameras.

Figure 5.18: Sample reconstructions of a sparse object in space domain. (a) A captured image that is degraded by $D = 2$ pixel downsampling, (b) the least square (minimum L_2 norm) estimate, and (c) the space-domain ℓ_1 penalty reconstruction from $K = 4$ frames. (d) A captured image that is degraded by $D = 4$ pixel downsampling, (e) the least square estimate, and (f) the space-domain ℓ_1 penalty reconstruction from $K = 4$ frames.

produces reconstructions with finer details than does least squares estimation. Next we apply our algorithms to reconstruct an object that is sparse in the wavelet domain. We display the object shown in Fig. 5.7(a) on the plasma monitor. The first and the second row in Fig. 5.19 display captured LR images and reconstructed images using four different algorithms for the case of $D = 2$ and $D = 4$ pixel downsampling respectively. The sampled captured images in Fig. 5.19(a) and (f) are interpolated to original size using a pixel replication method. The images in the second, third, fourth and fifth columns are reconstructed from $K = 4$ LR images using the LMMSE, EM, ℓ_1 penalty, and GGD penalty method respectively. The images reconstructed using these algorithms display more object details than the observed image. The reconstructed image quality for the case of $D = 2$ is superior to that of $D = 4$. All three wavelet-based methods produce object estimates having similar visual quality.

5.5 Conclusions

In this chapter we described a novel multiframe image restoration method to obtain a high-resolution object estimate from multiple low-resolution images that are warped, blurred, and corrupted by measurement noise. We incorporated the prior knowledge that an object may be sparse in some transform domain (e.g. wavelet) to improve the restoration. We formulated the multiframe image restoration problem as a maximum a-posteriori estimation. We obtained an estimate by solving a series of linearized regularized least squares problems. The object sparsity prior is incorporated using the ℓ_1 norm as the regularization operator (ℓ_1 penalty). We also modeled objects in the wavelet domain using generalized Gaussian density functions, and the model parameters are estimated from a set of training objects. The regularization operator

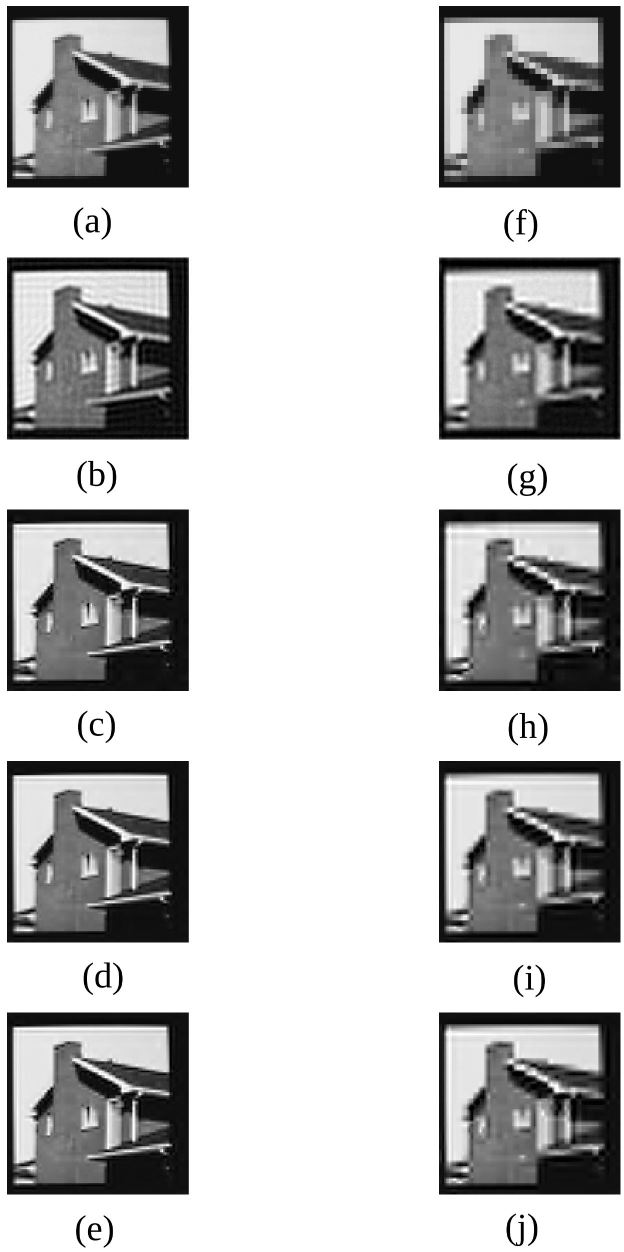


Figure 5.19: Sample reconstructions of a sparse object in wavelet domain. (a) A captured image that is degraded by $D = 2$ pixel downsampling, and images reconstructed using $K = 4$ such frames and the (b) LMMSE, (c) EM, (d) ℓ_1 penalty, and (e) GGD penalty methods. (f) A captured image that is degraded by $D = 4$ pixel downsampling, and images reconstructed using $K = 4$ such frames and the (g) LMMSE, (h) EM, (i) ℓ_1 penalty, and (j) GGD penalty methods.

is derived from these parameters (GGD penalty). We compared the performance of our algorithms with that of an EM algorithm for ℓ_1 norm minimization, as well as that of the LMMSE estimator. We considered imaging systems having different pixel sizes and different numbers of cameras deployed according to affine geometries. We quantified the reconstruction fidelity using both RMSE and visually weighted RMSE metrics. In the case of $D = 2$ pixel down-sampling the ℓ_1 penalty and GGD penalty methods produce reconstructions having smaller RMSE than LMMSE and EM algorithms for SNR values greater than 18dB and $K = 1 - 4$ LR images. VRMSE is generally the smallest for the GGD penalty method for these cases. In the case of $D = 4$ the ℓ_1 penalty method offers reconstructions with the smallest RMSE for a wide range of SNR and K. VRMSE values are the smallest for the GGD penalty method for SNR values less than 30dB. Using $K = 8$ LR images that are down-sampled by 4×4 pixels at $SNR = 50dB$ the reconstruction errors of object estimates obtained from ℓ_1 and GGD penalty algorithm are 5.5% smaller than the EM method and 14.3% smaller than LMMSE method. For both cases of $D = 2$ and $D = 4$, the EM method and the GGD penalty method offer the most sparse reconstructions. We also applied these algorithms to produce an object estimate from images gathered in a laboratory experiment. We demonstrated successful application of sparsity-constrained restoration of space-domain and the wavelet-domain sparse objects.

REFERENCES

[1] S. Chaudhuri Ed. *Super-Resolution Imaging.* Norwell, MA: Kluwer, 2001.

[2] R. Tsai and T. Huang. Multiframe image restoration and registration. *Advances in Computer Vision and Image Processing*, 1:317–339, 1984.

[3] H. Ur and D. Gross. Improved resolution from subpixel shifted pictures. *CVGIP: Graphical Models and Image Processing*, 54(2):181–186, March 1992.

[4] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, 53(3):231–239, May 1991.

[5] S. Kim, N. Bose, and H. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(6):1013–1027, October 1990.

[6] S. P. Kim and W. Y. Su. Recursive high-resolution reconstruction of blurred multiframe images. *IEEE Transactions on Image Processing*, 2(4):534–539, October 1993.

[7] B. Tom, A. Katsaggelos, and N. Galatsanos. Reconstruction of a high resolution image from registration and restoration of low resolution images. In *Proceedings International Conference on Image Processing*, volume 3, pages 553–557, November 1994.

[8] B. Tom and A. Katsaggelos. Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images. In *Proceedings International Conference on Image Processing*, volume 2, pages 539–542, 1995.

[9] M. Elad and A. Feuer. Restoration of a single superresolution image from several blurred,noisy, and undersampled measured images. *IEEE Transactions on Image Processing*, 6(12):1646–1658, December 1997.

[10] N. Nguyen, P. Milanfar, and G. Golub. A computationally efficient superresolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, 10(4):573–583, April 2001.

[11] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, October 2004.

[12] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli. How to take advantage of aliasing in bandlimited signals. *IEEE conf. on Accoustics Speech and Signal Processing*, 3:948–951, 2004.

[13] D. Robinson and P. Milanfar. Statistical performance analysis of super-resolution. *IEEE Transactions on Image Processing*, 15(6):1413–1428, June 2004.

[14] D. Robinson and P. Milanfar. Fundamental performance limits in image registration. *IEEE Transactions on Image Processing*, 13:1185–1199, 2004.

[15] M. Elad and A. Feuer. Super-resolution reconstruction of image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9):817–834, September 1999.

[16] H. Shekarforoush and R. Chellappa. Data-driven multichannel superresolution with application to video sequence. *JOSA-A*, 16(3):481–492, March 1999.

[17] R. Schultz and R. Stevenson. Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6):996–1011, June 1996.

[18] D. Chen and R. Schultz. Extraction of high-resolution video stills from mpeg imagesequences. In *Proceedings International Conference on Image Processing*, pages 465–469, 1998.

[19] J. Mateos, A. Katsaggelos, and R. Molina. Simultaneous motion estimation and resolution enhancement of compressed low resolution video. In *Proceedings International Conference on Image Processing*, volume 2, pages 653–656, 2000.

[20] H. Andrews and B. Hunt. *Digital Image Restoration*. Prentic-Hall Inc, New Jersey, 1977.

[21] H. Barrett and K. Myers. *Foundations of Image Science*. Wiley Series in Pure and Applied Optics, 2004.

[22] S. Osher, M. Burger, D. Goldfarb, J.Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, 2005.

[23] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena*, 60(3):259–268, November 1992.

[24] C. Vogel and M. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Transactions on Image Processing*, 7(6):813–824, June 1998.

[25] R. Hardie, K. Bernard, and E. Armstrong. Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE transactions on Image Processing*, 6:1621–1633, 1997.

[26] R. Hardie, K. Barnard, J. Bognar, E. Armstrong, and E. Watson. High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system. *Opt. Eng.*, 37:247–260, 1998.

[27] D. Rajan and S. Chaudhari. An mrf-based approach to generation of super-resolution images from blurred observation. *Journal of Mathematical Imaging and Vision*, 16:5–15, 2002.

[28] M. Sezan and H. Stark. Image restoration by convex projections in the presence of noise. *Applied Optics*, 22:2781–2789, 1983.

[29] M. Belge, M. Kilmer, and E. Miller. Wavelet domain image restoration with adaptive edge-preserving regularization. *IEEE Trans. Image Processing*, 9(4):597–608, 2000.

[30] N. Nguyen and P. Milanfar. A wavelet-based interpolation-restoration method for superresolution. *Circuits, Systems, and Signal Processing*, 16:321–338, 2000.

[31] N. Bose and S. Lertrattanapanich. Advances in wavelet superresolution. In *Proceedings of International Conference on Sampling Theory and Application*, pages 5–12, May 2001.

[32] D. Capel and A. Zisserman. Super-resolution from multiple views using learnt image models. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 627–634, 2001.

[33] Simon Baker and Takeo Kanade. Limits on super-resolution and how to break them. *IEEE Trans. Pattern Analysis And Machine Intelligence*, 24:1167–1183, 2002.

[34] A. Fruchter and R. Hook. Drizzle: A method for the linear reconstruction of undersampled images. *Publications of the Astronomical Society of the Pacific*, 114:144–152, 2002.

[35] R. Wagner, R. Nowak, and R. Baraniuk. Distributed image compression for sensor networks using correspondence analysis and super-resolution. In *Proceedings International Conference on Image Processing*, volume 1, pages 14–17, 2001.

[36] N. Omrane and P. Palmer. Superresolution for translated satellite images using the walsh functions. In *Proceedings of SPIE Image and Signal Processing for Remote Sensing IX*, volume 5238, pages 62–71, 2001.

[37] H. Greenspan, G. Oz, N. Kiryati, and S. Peled. Super-resolution in mri. In *Proceedings IEEE International Symposium on Biomedical Imaging*, volume 5238, pages 943–946, 2002.

[38] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka. Thin observation module by bound optics (tombo): concept and experimental verification. *Applied Optics*, 40:1806–1813, 2001.

[39] M. Christensen, V. Bhakta, D. Rajan, T. Mirani, S. Douglas, S. Wood, and M. Haney. Adaptive flat multiresolution multiplexed computational imaging architecture utilizing micromirror arrays to steer subimager fields of view. *Applied Optics*, 45:2884–2892, 2006.

[40] K. Aizawa, T. Komatsu, and T. Saito. A scheme for acquiring very high resolution images using multiple cameras. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 3:23–26, 1992.

[41] A. Segall, R. Molina, and A. Katsaggelos. High-resolution images from low-resolution compressed video. *IEEE Signal Processing Magazine*, 20:37–48, 2003.

[42] J. Lim. *Two-Dimensional Signal and Image Processing*. Prentice Hall Ptr., New Jersey, 1990.

[43] D. Gesbert, P. Duhamel, and S. Mayrargue. On-line blind multichannel equalization based on mutually referenced filters. *IEEE Trans. Signal Processing*, 45:2307–2317, 1997. , Vol. 45, pp.

[44] G. Giannakis and C. Tepedelenlioglu. Direct blind equalizers of multiple fir channels: A deterministic approach. *IEEE Trans. Signal Processing*, 47:62–74, 1999.

[45] J. Goodman. *Statistical Optics*. Wiley, New York, 1985.

[46] P. Milanfar and A. Shakouri. A statistical analysis of diffraction-limited imaging. *IEEE ICIP-2002*, 1:22–25, 2002.

[47] W. Press, S. Teukolsky, W. Vettering, and B. Flannery. *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, 2002.

[48] M. Chiang and T. Boult. Efficient image warping and super-resolution. In *Proceedings of the Third IEEE Workshop on Applications of Computer Vision*, pages 56–61, 1996.

[49] M. Chiang and T. Boult. Local blur estimation and super-resolution. In *Proceeding of Computer Vision and Pattern Recognition IEEE Computer Society Conference*, pages 821–826, 1997.

[50] R. Molina, J. Mateos, A. Katsaggelos, and M. Vega. Bayesian multichannel image restoration using compound gauss-markov random fields. *IEEE Trans. on Image Processing*, 12:1642–1654, 2003.

[51] C. Miller, B. Hunt, M. Marcellin, and M. Neifeld. Image restoration with the viterbi algorithm. *J. Opt. Soc. Am. A*, 17:265–275, 2000.

[52] M. Neifeld, R. Xuan, and M. Marcellin. Communication theoretic image restoration. *Applied Optics*, 39:269–276, 2000.

[53] M. Neifeld and Y. Wu. Parallell image restoration with a two-dimensional likelihood-based algorithm. *Applied Optics*, 41:4812–4824, 2002.

[54] M. Turk and A. Pentland. Face recognition using eigenfaces. In *IEEE Proceedings on Computer Vision and Pattern Recognition*, pages 586–591, 1991.

[55] X. Chen, K. Chugg, and M. Neifeld. Near-optimal parallel distributed data detection for page-oriented optical memories. *IEEE Journal on Selected Topics in Quantum Electronics*, 4:866–879, 1998.

[56] K. Chugg, A. Anastasopoulos, and X. Chen. *Iterative Detection: Adaptivity, Complexity Reduction, and Applications.* Kluwer Academic Publishers, 2001.

[57] C. Berrou, A. Glavieux, and P. Thitimajshima. Near shannon limit error-correcting coding and decoding: Turbo-codes(1). In *Proc. IEEE Int. Conf. Communications*, volume 64, pages 1064–1070, 1993.

[58] R. McEliece, D. MacKay, and J. Cheng. Turbo decoding as an instance of pearl's as an instance of pearl's "belief propagation" algorithm. *IEEE Journal on Selected Areas in Communications*, 16:140–152, 1998.

[59] F. Kschischang, B. Frey, and H. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47:498–519, 2001.

[60] P. Shankar, W. Hassenplaugh, R. Morrison, R. Stack, , and M. Neifeld. Multi-aperture imaging. *Applied Optics*, 45:2871–2883, 2006.

[61] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Information Theory*, 52:489–509, 2006.

[62] E. Candes and J. Romberg. signal recovery from random projections. In *Proceedings of SPIE, Computational Imaging III*, volume 5678, pages 76–86, 2005.

[63] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57:1413–1457, 2004.

[64] E. Cands, M. Wakin, and S. Boyd. Enhancing sparsity by reweighted l1 minimization. Technical report, Technical Report, California Institute of Technology, 2007.

[65] M. Duarte, M. Wakin, and R. Baraniuk. Fast reconstruction of piecewise smooth signals from random projections. In *Online Proceedings of the Workshop on Signal Processing with Adaptative Sparse Structured Representations.* http://spars05.irisa.fr/ACTES/TS5-3.pdf, 2005.

[66] D. Donoho. De-noising by soft-thresholding. *IEEE Trans. Information Theory*, 41:613–627, 1995.

[67] M. Figueiredo and R. Nowak. An em algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 12:906–916, 2003.

[68] D. Wipf and B. Rao. Sparse bayesian learning for basis selection. *IEEE Trans. Signal Processing*, 52:2153–2164, 2004.

[69] C. Paige and M. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Mathematical Software*, 8:43–71, 1982.

[70] C. Paige and M. Saunders. Lsqr: Sparse linear equations and least squares problems. *ACM Trans. Mathematical Software*, 8:195–209, 1982.

[71] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[72] E. Candes, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[73] M. Figueiredo, R. Nowak, and S. Wright. Gradient projections for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 1:586–597, 2007.

[74] M. Kilmer and D. O'Leary. Choosing regularization parameters in iterative methods for ill-posed problems. *SIAM J. Matrix Anal. Appl.*, 22(4):1204–1221, 2001.

[75] P. Hansen. *Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion.* SIAM, 1998.

[76] M. Jiang, L. Xia, G. Shou, and M. Tang. Combination of the lsqr method and a genetic algorithm for solving the electrocardiography inverse problem. *Physics in Medicine and Biology*, 52:1277–1294, 2007.

[77] N. Valdivia and E. Williams. Krylov subspace iterative methods for boundary element method based near-field acoustic holography. *The Journal of the Acoustical Society of America*, 117:711–724, 2005.

[78] B. Rao and K. Kreutz-Delgado. An affine scaling methodology for best basis selection. *IEEE Trans. Signal Processing*, 47(1):187–200, 1999. , Vol. 47, 1,.

[79] P. Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4):561–580, 1992. , Vol. 34, 4,.

[80] S. Mallat. A compact multiresolution representation: the wavelet model. In *Proc. IEEE Workshop Comp. Vision*, pages 2–7, FL, 1987.

[81] B. Jeffs and M. Gunsay. Restoration of blurred star field images by maximally sparse optimization. *IEEE Trans. Image Processing*, 2(2):202–211, 1993.

[82] G. Harikumar and Y. Bresler. A new algorithm for computing sparse solutions to linear inverse problems,. In *Proc. ICASSP*, volume 3, pages 1331–1334, 1996.

[83] J. Mannos and D. Sakrison. The effects of visual fidelity criterion on the encoding of image. *IEEE Trans. Information Theory*, 20(3):525–536, 1974.

[84] The USC-SIPI image database. Signal and image processing institute at the university of southern california, available at http://sipi.usc.edu/database. Technical report, University of Southern California, 2002.

[85] A. Goshtasby. Image registration by local approximation methods. *Image and Vision Computing*, 6:255–261, 1988.