Spatial Resolution Enhancement of Low-Resolution Image Sequences - A Comprehensive Review with Directions for Future Research

Article · December 1998

Source: CiteSeer

Robert L. Stevenson
University of Notre Dame

Spatial Resolution Enhancement of Low-Resolution Image Sequences
A Comprehensive Review with Directions for Future Research

Sean Borman, Robert Stevenson*

July 8, 1998

Abstract

Recent years have seen growing interest in the problem of super-resolution restoration of video sequences. Whereas in the traditional single image restoration problem only a single input image is available for processing, the task of reconstructing super-resolution images from multiple undersampled and degraded images can take advantage of the additional spatio-temporal data available in the image sequence. In particular, camera and scene motion lead to frames in the source video sequence containing similar, but not identical information. The additional information available in these frames make possible reconstruction of visually superior frames at higher resolution than that of the original data. In this paper we review the current state of the art and identify promising directions for future research.

*The authors are with the Laboratory for Image and Signal Analysis (LISA), University of Notre Dame, Notre Dame, IN 46556. E-mail: rls@nd.edu .

Executive Summary

This document presents a detailed review of existing techniques which address the problem of super-resolution video restoration. By "super-resolution," we refer to removal of blur caused by the imaging system (out of focus blur, motion blur, non-ideal sampling, etc.) as well as recovery of spatial frequency information beyond the diffraction limit of the optical system.

Super-resolution restoration from a still image is a well recognized example of an *ill-posed* inverse problem. Such problems may be approached using *regularization* methods which *constrain the feasible solution space* by employing *a-priori* knowledge. This may be achieved in two complimentary ways; (1) obtain additional novel observation data and (2) constrain the feasible solution space with *a-priori* assumptions on the form of the solution. Both techniques feature in modern super-resolution restoration methods which utilize (1) image *sequences* which provide additional spatio-temporal observation constraints (typically in the form of novel data arising from *sub-pixel motion*) and (2) various *a-priori* constraints on the super-resolution image (e.g. local smoothness, edge preservation, positivity, energy boundedness, etc.). The use of non-linear *a-priori* constraints provides the potential for *bandwidth extension* beyond the diffraction limit of the optical system.

Super-resolution techniques may be divided into two main classes; frequency domain and spatial domain. All frequency domain approaches are, to a greater or lesser extent, unable to accommodate general scene observation models including spatially varying degradations, non-global relative camera/scene motion, general *a-priori* constraints or general noise models. Spatial domain formulations can accommodate all these and provide enormous flexibility in the range of degradations and observation models which may be represented and are thus the methods of choice. Spatial domain observation models facilitate inclusion of additional data in the observation equation with the effect of reducing the feasible solution space.

It remains however to compute the solution to the ill-posed super-resolution inverse problem. Amongst the numerous solution techniques featuring in the literature, the Bayesian *Maximum A-Posteriori* (MAP) estimation method, and the method of *Projection Onto Convex Sets* (POCS) are most promising. MAP estimation provides a rigorous theoretical framework, several desirable mathematical properties and makes explicit use of *a-priori* information in the form of a prior probability density on the solution image. POCS defines the feasible solution space as the intersection of convex constraint sets and provides a convenient method for the inclusion of *a-priori* constraints. MAP estimation and POCS are complimentary, with initial work already presented on promising hybrid MAP/POCS solution methods.

We identify three critical factors affecting super-resolution restoration. Firstly, *reliable sub-pixel motion information is essential*. Poor motion estimates are more detrimental to restoration than a lack of motion information. Secondly, *observation models must accurately describe the imaging system and its degradations*. Thirdly, *restoration methods must provide the maximum potential for inclusion of* a-priori *information*. These observations, and the discussion of constraints on the solution space suggest the following approaches and actions:

Motion estimation: Existing sub-pixel motion estimation methods should be studied and extended with emphasis on providing robust performance and reliability measures on motion estimates. Regularized motion estimation methods, as well as simultaneous motion estimation and restoration methods should be considered. Feature based techniques which provide sparse motion maps but reliable "edge based" motion estimates provide useful high frequency data constraints.

Motion models: Motion models and estimation methods should be chosen according to a-priori knowledge of scene/camera motion. Model based motion estimators should be considered for multiple independent motion. As motion is estimated from *degraded* observation data, a study of the reliability of these estimates should be made. Motion constraints should be applied to the restored super-resolution image sequence. The last two points appear to have been overlooked in the literature.

Observation models: Observation models which accurately account for degradations occurring in the imaging system (and thus accurately describe the relationship between the super-resolution image and the observed data) constrain the image solution space. Little attention has been paid to modeling CCD image sensor geometry, spatio-temporal integration, noise and readout characteristics. Better modeling of the observation process promises improved restoration.

Restoration Algorithms: Hybrid MAP/POCS approaches combine mathematical rigor and uniqueness of solution with convenient description of *a-priori* constraints. Restoration algorithms should be capable of including reliability measures on individual motion estimates as well as accommodate model based motion estimators. Simultaneous multi-frame super-resolution restoration which has not been addressed in the literature, will provide a powerful opportunity for further solution space constraints.

Nomenclature

In this section we describe the conventions which will be adopted throughout this document.

These are general notational conventions followed in the document.

General Notational Conventions	
f	Scene
y	Observation
n	Noise
z	SR reconstruction

The following typographical conventions apply to all mathematical objects in this document.

Typographical Conventions		
f	Lowercase italic	Scalar
$\mathbf{f}$	Lowercase boldface	Vector
$\mathbf{F}$	Uppercase boldface	Stacked vector or matrix

Conventions used to describe the image formed at the image plane are listed below:

Illumination at Focal Plane (Desired Image f)	
Spatial coordinates	
(x_1, x_2)	Location in image plane coordinate system
Time invariant imagery (Single image)	
$f(x_1, x_2)$	Projection of optical system at image plane
Time varying imagery (Multiple images)	
t	Continuous time variable
$f(x_1, x_2, t)$	Time varying projection of optical system at image plane

Assuming no noise is present, the image sensor array will record the incident illumination to yield the following (ideal) measurements:

Noiseless Image (f)	
Detector array coordinates	
(x_1, x_2)	Location in image plane coordinate system
$[m_1, m_2]$	Pixel location in detector array
Time invariant imagery (Single image)	
T_a	Aperture time (reciprocal of shutter speed) at image acquisition
$f(x_1, x_2)$ $f[m_1, m_2]$	Continuous scene Image incident at detector array $m_1 \in \{1, 2, \ldots, M_1\}$, $m_2 \in \{1, 2, \ldots, M_2\}$
f	Lexicographic ordering of $f[m_1, m_2]$ (column vector)
f_i	i^{th} element of column vector **f** and, i^{th} pixel in the lexicographic ordering of image $f_r[m_1, m_2]$
Time varying imagery (Multiple images)	
R	Total number of acquired images
r	Acquired image number $r \in \{1, 2, \ldots, R\}$
T_{a_r}	Aperture time for r^{th} image acquisition
t	Continuous time variable
t_r	Time instant of the beginning of the r^{th} image acquisition
$f(x_1, x_2, t)$	Time varying projection of optical system at image plane
$f[m_1, m_2, r]$	r^{th} image recorded by detector array. In general, $m_1 \in \{1, 2, \ldots, M_{1_r}\}$, $m_2 \in \{1, 2, \ldots, M_{2_r}\}$
$\mathbf{f_r}$	Lexicographic ordering of $f[m_1, m_2, r]$ (column vector)
f_{r_i}	i^{th} element of column vector $\mathbf{f_r}$ and, i^{th} pixel in the lexicographic ordering of image $f_r[m_1, m_2, r]$
F	$\mathbf{F} = \left[\mathbf{f_1}^T \;\; \mathbf{f_2}^T \;\; \mathbf{f_3}^T \;\; \cdots \;\; \mathbf{f_R}^T\right]^T$ Column vector containing R lexicographically ordered images

In the presence of observation noise however, we are unable to acquire $\mathbf{f}$ but instead measure $\mathbf{y}$ as tabulated below:

Recorded Imagery (Noisy Measured Data $\mathbf{y}$)	
Detector array coordinates	
$[m_1, m_2]$	Pixel location in detector array
Time invariant imagery (Single image)	
T_a	Aperture time (reciprocal of shutter speed) at image acquisition
$y[m_1, m_2]$	Image recorded by detector array (typically discrete valued) $m_1 \in \{1, 2, \ldots, M_1\}$, $m_2 \in \{1, 2, \ldots, M_2\}$
$\mathbf{y}$	Lexicographic ordering of $y[m_1, m_2]$ (column vector)
y_i	i^{th} element of column vector $\mathbf{y}$ and, i^{th} pixel in the lexicographic ordering of image $y_r[m_1, m_2]$
Time varying imagery (Multiple images)	
R	Number of recorded images
r	Acquired image number $r \in \{1, 2, \ldots, R\}$
t	Continuous time variable
t_r	Time instant of the beginning of the r^{th} image acquisition
T_{a_r}	Aperture time for r^{th} image acquisition
$y[m_1, m_2, t_r]$	Image acquired by detector array at time $t = t_r$ $m_1 \in \{1, 2, \ldots, M_{1_r}\}$, $m_2 \in \{1, 2, \ldots, M_{2_r}\}$
$\mathbf{y_{t_r}}$	Lexicographic ordering of $y[m_1, m_2, t_r]$ (column vector)
$y_{t_{r\,i}}$	i^{th} element of column vector $\mathbf{y}_{t_r}$
$y[m_1, m_2, r]$	r^{th} image recorded by detector array (same as $y[m_1, m_2, t_r]$)
$\mathbf{y_r}$	Lexicographic ordering of $y[m_1, m_2, r]$ (column vector)
y_{r_i}	i^{th} element of column vector $\mathbf{y_r}$ and, i^{th} pixel in the lexicographic ordering of image $y_r[m_1, m_2, r]$
$\mathbf{Y}$	$\mathbf{Y} = \left[\mathbf{y_1}^T \;\; \mathbf{y_2}^T \;\; \mathbf{y_3}^T \;\; \cdots \;\; \mathbf{y_R}^T\right]^T$ Column vector containing R lexicographically ordered images

Where noise is included the observation model, the following notation will be utilized.

Noise Model ($\mathbf{n}$)	
Detector array coordinates	
$[m_1, m_2]$	Pixel location in detector array
Time invariant imagery (Single image)	
T_a	Aperture time (reciprocal of shutter speed) at image acquisition
$n[m_1, m_2]$	Noise component of measurement from pixel $[m_1, m_2]$ in detector array $m_1 \in \{1, 2, \ldots, M_1\}$, $m_2 \in \{1, 2, \ldots, M_2\}$
$\mathbf{n}$	Lexicographic ordering of $n[m_1, m_2]$ (column vector)
n_i	i^{th} element of column vector $\mathbf{n}$
Time varying imagery (Multiple images)	
R	Number of recorded images
r	Acquired image number $r \in \{1, 2, \ldots, R\}$
T_{a_r}	Aperture time for r^{th} image acquisition
t	Continuous time variable
t_r	Time instant of the beginning of the r^{th} image acquisition
$n[m_1, m_2, r]$	Noise component in r^{th} image acquisition by detector array
$\mathbf{n_r}$	Lexicographic ordering of $n[m_1, m_2, r]$ (column vector)
n_{r_i}	i^{th} element of column vector $\mathbf{n_r}$
$\mathbf{N}$	$\mathbf{N} = \left[\mathbf{n_1}^T \ \mathbf{n_2}^T \ \mathbf{n_3}^T \ \cdots \ \mathbf{n_R}^T\right]^T$ Stacked column vector containing noise component in the p image acquisitions

The notation for the computed super-resolution image is described below:

Super-resolution Images (Reconstructed Images z)	
Super-resolution image coordinates	
$[n_1, n_2]$	Pixel location in super-resolution image
Super-resolution still image (Single image)	
$z[n_1, n_2]$	SR still image $n_1 \in \{1, 2, \ldots, N_1\}$, $n_2 \in \{1, 2, \ldots, N_2\}$ Typically $N_1 > M_{1_r}$, $N_2 > M_{2_r}$, $1 \leq r \leq R$
$\mathbf{z}$	Lexicographic ordering of SR image $z[n_1, n_2]$
z_i	i^{th} element of **z** and, i^{th} pixel in the lexicographically ordered SR image **z**
$\hat{\mathbf{z}}^{(j)}$	j^{th} iterative approximation to solution for SR image **z**
Spatial resolution enhanced video (frame rate same as LR sequence)	
$z[n_1, n_2, r]$	SR image coincident with r^{th} LR frame
$\mathbf{z_r}$	Lexicographic ordering of SR image $z[n_1, n_2, r]$
z_{r_i}	i^{th} element of column vector $\mathbf{z_r}$
Spatio-temporal resolution enhanced video (frame rate may differ from LR video)	
s	Index of the reconstructed super-resolution image $1 \leq s \leq S$
t_s	s^{th} time instant of SR restoration. $t_s \in \mathbb{R}$, $t_j > t_i$ for $j > i$
$z[n_1, n_2, t_s]$	SR image computed for time $t = t_s$
$\mathbf{z}_{t_s}$	Lexicographic ordering of SR image $z[n_1, n_2, t_s]$
$z_{t_s i}$	i^{th} element of column vector $\mathbf{z}_{t_s}$
$z[n_1, n_2, s]$	s^{th} computed SR image (time $t = t_s$)
$\mathbf{z_r}$	Lexicographic ordering of SR image $z[n_1, n_2, s]$
z_{r_i}	i^{th} element of column vector $\mathbf{z_s}$
$\mathbf{Z}$	$\mathbf{Z} = \left[\mathbf{z_1}^T \;\; \mathbf{z_2}^T \;\; \mathbf{z_3}^T \;\; \cdots \;\; \mathbf{z_S}^T\right]^T$ Stacked column vector containing S super-resolution images

Glossary

DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
GMRF	Gaussian Markov Random Field
GGMRF	Generalized Gaussian Markov Random Field
HMRF	Huber Markov Random Field
HRVS	High Resolution Video Still
LMMSE	Linear Minimum Mean Square Error
LR	Low Resolution
LSI	Linear Shift Invariant
LSV	Linear Shift Varying
MAP	Maximum A-Posteriori
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MRF	Markov Random Field
MSE	Mean Square Error
POCS	Projection Onto Convex Sets
PSF	Point Spread function
RLS	Recursive Least Squares
RTLS	Recursive Total Least Squares
SVPSF	Space Variant Point Spread function
SR	Super-resolution
TLS	Total Least Squares

Contents

1 Introduction

This report considers the problem of spatial resolution enhancement of low-resolution video sequences. It reviews existing techniques, discussing the relative merits and demerits of these approaches, and identifies promising directions for future research.

In the section that follows, we provide some background information which will provide a overall perspective of existing and future techniques in spatial and temporal enhancement of video sequences.

1.1 A Hierarchy of Spatio-Temporal Video Enhancement Research

Spatial resolution enhancement of low-resolution video sequences has emerged from earlier work; first robust image interpolation for single frame resolution enhancement, followed by improved resolution still images from video. To complete the hierarchy, one may consider also the more general problem of spatial and temporal resolution enhancement of a low-resolution video sequence. The hierarchy of increasingly general techniques is illustrated in Figure 1 below.

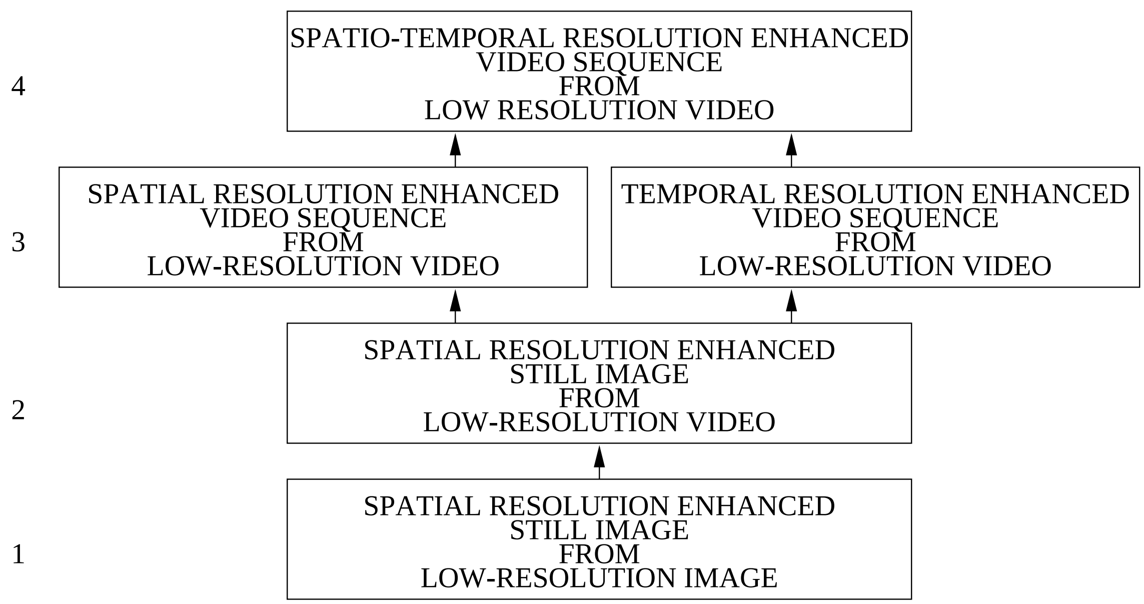

Figure 1: Hierarchy describing spatial and temporal resolution enhancement techniques.

The focus of this report is spatial resolution enhancement of low-resolution video which lies at the third level in this hierarchy. Though work directly addressing this problem will be reviewed, we shall concentrate considerable effort in reviewing work at level 2 (resolution enhanced stills from low-resolution video) since this is the basis for much of the existing work on video resolution enhancement. In fact, most existing techniques for spatial resolution enhancement are extensions of those used for reconstructing still images from a video sequence, so we are justified in our approach. Level 1 describes single image interpolation work from which techniques are sometimes utilized in reconstruction of still images from video.

In the material that follows, the term super-resolution (SR) may be considered synonymous with spatial resolution enhancement.

1.2 Formal Problem Definition

The super-resolution video enhancement problem may be formally described as follows:

> Let $f(x_1, x_2, t),\ x_1, x_2, t \in \mathbb{R}$ denote the time-varying scene in the image plane coordinate system. Given a sequence of R low-resolution sampled images $y[m_1, m_2, r]$ with $m_1 \in \{1, 2, \ldots, M_{1_r}\}$, $m_2 \in \{1, 2, \ldots, M_{2_r}\}$ and $r \in \{1, 2, \ldots, R\}$ acquired by imaging of the scene $f(x_1, x_2, t)$ at times $t_1 < t_2 < \cdots < t_r < \cdots < t_R$, our objective is to form S estimates $\hat{f}[n_1, n_2, s],\ 1 \leq s \leq S$ of $f(x_1, x_2, \tau_s)$ on the discrete, *super-resolution* sampling grid indexed by $[n_1, n_2]$ with $n_1 \in \{1, 2, \ldots, N_{1_s}\}$, $n_2 \in \{1, 2, \ldots, N_{2_s}\}$ at the arbitrary time instants $t_1 \leq \tau_1 < \tau_2 < \cdots < \tau_s < \cdots < \tau_S \leq t_R$. Typically we choose $N_{1_s} > M_{1_r},\ N_{2_s} > M_{2_r}, \forall\ r, s$ and / or $S > R$. Super-resolution refers to the reconstruction of images $\hat{f}[n_1, n_2, s]$ that are visually superior to the original low resolution observations. This often implies *bandwidth extrapolation* beyond the passband of the imaging system.

Super-resolution reconstruction methods typically make use of a set of low-resolution frames when computing each super-resolution frame. Is is the additional spatio-temporal information available in the *sequence* of low-resolution images enables reconstruction at resolutions higher than that of the original data.

1.3 Super-Resolution Enhancement as an Ill-Posed Inverse Problem

One of the fundamental ideas we shall repeatedly encounter is the fact that super-resolution reconstruction is an ill-posed inverse problem [1]. The problem is typically that a multiplicity of possible solutions exists given a set of observation images. The accepted approach to tackling such problems is to constrain the solution space according to *a-priori* knowledge on the form of the solution. This may include such constraints such as smoothness, positivity and so on. Inclusion of such constraints is critical to achieving high quality super-resolution reconstructions, so much emphasis will be placed on techniques which enable the inclusion of *a-priori* knowledge.

1.4 Super-Resolution Video Reconstruction

Many of the papers we shall discuss in this report do not directly consider the problem of super-resolution *video* restoration, but instead concentrate on restoration of a single super-resolution still image from a short, low resolution image sequence. All of these techniques may, however, be applied to video restoration by using a shifting window of processed frames as illustrated in Figure 2. For a given super-resolution frame, a “sliding window” determines the set of low resolution frames to be processed to produce the output. The window is moved forward to produce successive super-resolution frames in the output sequence.

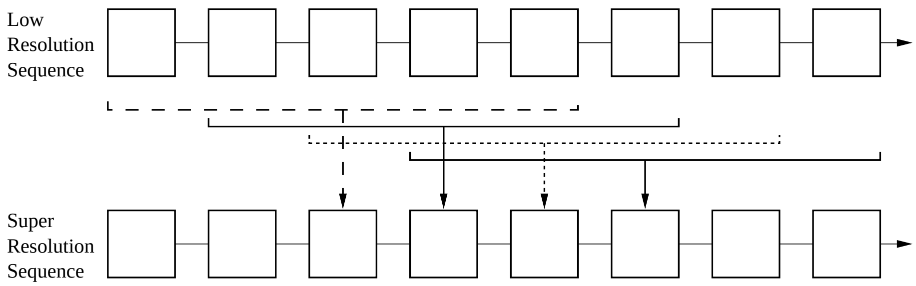

Figure 2: Super-resolution video enhancement from a low resolution image sequence.

1.5 Applications

Applications for the techniques of high resolution reconstruction from sequences appear to be growing rapidly. Super-resolution techniques have already been applied in the following areas:

- Satellite imaging
- Video Surveillance
- Video enhancement and restoration
- Video standards conversion
- Microscopy
- Digital Mosaicing
- Aperture displacement cameras [2, 3]
- Medical computed tomographic imaging

This list appears to be growing rapidly as super-resolution techniques become better known.

1.6 Overview of this Report

We will review the history and current state of the art by categorizing existing techniques into several divisions. At the highest level, super-resolution techniques can be divided in to *frequency domain* or *spatial domain* algorithms. Much of the earlier work concentrated on the frequency domain formulation, but as more general degradation models were considered, later research has tended to concentrate almost exclusively on spatial domain formulations.

We will discuss the frequency domain formulation of the SR problem in Section 2 which is followed by spatial domain techniques in Section 3. Each of these sections include short concluding summaries. The report is concluded with a final summary and an identification of promising directions for future work.

2 Frequency Domain Methods

2.1 Introduction

In this section we discuss a major class of super-resolution methods which utilize a frequency domain formulation of the super-resolution problem. The techniques discussed utilize the shifting property of the Fourier transform to model global translational scene motion, and take advantage of the sampling theory to enable effect restoration made possible by the availability of multiple observation images.

It is interesting to note that the methods we discuss here include the earliest investigation of the super-resolution problem, and although there are significant disadvantages in the frequency domain formulation, work has continued in this area until relatively recently when spatial domain techniques, with their increased flexibility, have become more prominent. This does not however mean to say that frequency domain techniques be ignored. Indeed, under the assumption of global translational motion, frequency domain methods are computationally highly attractive.

We begin our review of frequency domain methods with the seminal work of Tsai and Huang [4].

2.2 Reconstruction via Alias Removal

The earliest formulation, and proposed solution to the multi-frame super-resolution problem was undertaken by Tsai and Huang [4] in 1984, motivated by the need for improved resolution images from Landsat image data. Landsat acquires images of the same areas of the earth in the course of its orbits, thus producing a sequence of similar, but not identical images. Observed images are modeled as under-sampled versions of a unchanging scene undergoing global translational motion. Impulse sampling is assumed, but the sampling rate fails to meet the Nyquist criterion [5]. Neither the effects of blurring due to satellite motion during image acquisition nor observation noise are considered.

The authors propose a frequency domain formulation based on the shift and aliasing properties [6] of the continuous and discrete Fourier transforms for the reconstruction of a band-limited image from a set under-sampled, and therefore aliased, observation images. The shift and aliasing properties are used to formulate a system of equations which relate the aliased discrete Fourier transform (DFT) coefficients of the observed images to samples of the continuous Fourier transform (CFT) of the unknown original scene. The system of equations is solved for the frequency domain coefficients of the original scene, which is then recovered using the inverse DFT. Formulation of the system of equations requires knowledge of the translational motion between frames to sub-pixel accuracy. Solution of the equations requires that each observation contribute *independent* equations, which places restrictions on the inter-frame motion that contributes useful data.

With the continuous scene denoted by $f(x_1, x_2)$, global translations of $f(x_1, x_2)$ yield R shifted images, $f_r(x_1, x_2) = f(x_1 + \Delta_{x_{1_r}}, x_2 + \Delta_{x_{2_r}})$ where $r = 1, 2, \ldots, R$. The CFT of the scene is given by $\mathcal{F}(u_1, u_2)$ and that of the translations by $\mathcal{F}_r(u_1, u_2)$. The shifted images are impulse sampled to yield observed images $y_r[m_1, m_2] = f(m_1 T_{x_1} + \Delta_{x_{1_r}}, m_2 T_{x_2} + \Delta_{x_{2_r}})$ with $m_1 = 0, 1, \ldots, M_1 - 1$ and $m_2 = 0, 1, \ldots, M_2 - 1$. The R corresponding 2D DFT's are denoted $\mathcal{Y}_r[k_1, k_2]$. The CFT of the scene and the DFT's of the shifted and sampled images are related via aliasing [6],

$$\mathcal{Y}_r[k_1,k_2] = \frac{1}{T_{x_1}T_{x_2}} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \mathcal{F}_r\left(\frac{k_1}{M_1T_{x_1}} + p_1 f_{s_1}, \frac{k_2}{M_2T_{x_2}} + p_2 f_{s_2}\right) \tag{1}$$

where $f_{s_1}=1/T_{x_1}$ and $f_{s_2}=1/T_{x_2}$ are the sampling rates in the x_1 and x_2 dimensions respectively. We will also use the shifting property of the CFT,

$$\mathcal{F}_r(u_1,u_2) = e^{j2\pi(\Delta_{x_{1_r}}u_1+\Delta_{x_{2_r}}u_2)}\mathcal{F}(u_1,u_2) \tag{2}$$

If $f(x_1,x_2)$ is band-limited, there exist L_i such that $\mathcal{F}(u_1,u_2) \to 0$ for $|u_i| \geq L_i f_{s_i}$, $i=1,2$. Assuming $f(x_1,x_2)$ is band-limited, we may use the shifting property in (2) to rewrite the alias relationship of (1) in matrix form as,

$$\mathbf{Y} = \Phi\mathbf{F} \tag{3}$$

$\mathbf{Y}$ is a $R\times 1$ column vector with the r^{th} element being the DFT coefficients $\mathcal{Y}_r[k_1,k_2]$ of the observed image $y_r[m_1,m_2]$. Φ is a matrix which relates the DFT of the observation data to samples of the unknown CFT of $f(x_1,x_2)$ contained in the $4L_1L_2\times 1$ vector $\mathbf{F}$.

Super-resolution reconstruction therefore is reduced to finding the DFT's of the R observed images, determining Φ, solving the system of equations (3) for $\mathbf{F}$ and then using the inverse DFT to obtain the reconstructed image. It is shown that Φ may be factored so as to reduce the computational cost in solving for the CFT coefficients in $\mathbf{F}$.

Since the system matrix Φ requires knowledge of the translation parameters $\Delta_{x_1}, \Delta_{x_2}$, which are not typically known *a-priori*, these parameters must be estimated before reconstruction is possible. Super-resolution reconstruction is thus effected using a two step process: motion estimation to determine the translation parameters, followed by restoration of the improved resolution image.

The authors address the problem of registration using a novel approach which appears not to have gained significant recognition. Since the observed images $y_r[m_1,m_2]$ are under-sampled, there is some question as to the accuracy of standard techniques for motion estimation which typically utilize two (though sometimes more) frames when computing motion estimates. It is well recognized [7] that the accuracy of the motion estimates is arguably *the* limiting factor in super-resolution reconstruction performance, so any fruitful consideration of this problem promises significant returns. A simultaneous *multi-frame* image registration algorithm is proposed which is shown to deliver reliable registration parameters even under the conditions of severe undersampling, provided a sufficiently large number of observation frames are available. This idea does not appear to have been sufficiently explored in the super-resolution community and holds considerable promise as a means of improving current super-resolution methods. It should be pointed out however that the idea addresses only the relatively simple problem of global translation estimation.

In summary, the Tsai-Huang frequency domain method, though computationally attractive, has significant disadvantages. The assumption of ideal sampling is unrealistic. The possibility of an optical system point spread function, or even that of spatially integrating sensors is not addressed. Observation noise is not considered, which is a major shortcoming given that noise will have a detrimental effect on the solution of (3). Blurring due to finite aperture time is also not considered. The global translation model is, for many applications, inappropriate.

It is important to recognize that the global translation model is intimately bound with the proposed frequency domain approach. Indeed, global translation in the spatial domain is related, via the Fourier shifting property, to phase shifts in the spatial-frequency domain. This fortuitous relationship enabled the formulation presented, however it is not difficult to see that global translation is exceptional in that more general motion models will not lend themselves so cleanly to a frequency domain formulation. It is worth noting that *spatially varying* (non-global) motion models, which are of great interest in super-resolution reconstruction of scenes containing independent motion, could not be easily accommodated using the frequency domain formulation. This lack of flexibility is a major disadvantage of all frequency domain approaches similar to [4].

Two of the limitations of the Tsai-Huang method are addressed by Tekalp, Ozkan and Sezan in [8]. The authors propose a frequency domain approach which extends [4] by including the effects of a LSI PSF $h(x_1, x_2)$ as well as observation noise. Periodic sampling is still assumed, but is not necessarily ideal, since the PSF of the sensor can be integrated into the system PSF. As in [4], a translation only motion model is used, which enables the formulation of a Fourier domain system model which relates the noisy, degraded and under-sampled observed images to the original 2-D scene. By utilizing L^2 observed images, each L times undersampled in the x_1 and x_2 dimensions, a system of L^2 equations in L^2 unknowns is formulated for each sampled frequency domain point. These equations are solved in the least squared sense due to the presence of observation noise.

The authors assume the spatial domain observation model,

$$y_r[m_1, m_2] = \sum_{p_1=0}^{M_1-1} \sum_{p_2=0}^{M_2-1} \left[f(x_1 - \Delta_{x_{1_r}}, x_2 - \Delta_{x_{2_r}}) ** h(x_1, x_2) \right] \delta(x_1 - p_1 T_{x_1}, x_2 - p_2 T_{x_2}) + n_r(p_1 T_{x_1}, p_2 T_{x_2}) \tag{4}$$

Similar to [4], the equivalent frequency domain equations are derived for (4), which relates the Fourier coefficients of the undersampled observation images of the shifted and blurred original high-resolution image. After least-squares solution to the frequency domain equations, the estimate of the original image is reconstructed by inverse DFT of the sampled CFT data.

In addition to the extensions of [4], the authors propose a spatial domain projection onto convex sets (POCS) reconstruction method as well as a spatial domain interpolation-restoration approach. These are discussed in Sections 3.6.1 and 3.2 respectively.

In the text, “Digital Video Processing,” Tekalp [9] dedicates a chapter to the discussion of select techniques in super-resolution reconstruction. Tekalp provides an introduction to the super-resolution problem and discusses various observation models which are capable of including the effects of scene motion and sensor and optical system point spread functions. The frequency domain model of Tsai and Huang [4] , as well as the extensions proposed in [8] are presented in a tutorial exposition. The POCS method proposed in [8] (see Section 3.6.1) is also reviewed and examples are presented. No novel results are presented in this work.

Kaltenbacher and Hardie [10] utilize the frequency domain formulation proposed by Tsai and Huang [4] for restoration of aliased, under-sampled low resolution frames, but propose an alternative method for estimating the inter-frame global translational parameters required for restoration. Their approach is considerably cheaper computationally than that of [4].

Recorded low resolution images $y[m_1, m_2]$ are expanded in a first order Taylor series,

$$y[m_1, m_2] \approx y[m_{1_0}, m_{2_0}] + (m_1 - m_{1_0})\frac{\partial y[m_{1_0}, m_{2_0}]}{\partial m_1} + (m_2 - m_{2_0})\frac{\partial y[m_{1_0}, m_{2_0}]}{\partial m_2} \tag{5}$$

Given the assumption of global translational motion, the terms $\Delta_{m_1} = (m_1 - m_{1_0})$ and $\Delta_{m_2} = (m_2 - m_{2_0})$ represent global translation between a pair of observed low resolution frames. Δ_{m_1} and Δ_{m_2} are estimated by minimizing the square error,

$$\epsilon = \sum_{m_1}\sum_{m_2}\left[y[m_1, m_2] - y[m_{1_0}, m_{2_0}] - \Delta_{m_1}\frac{\partial y[m_{1_0}, m_{2_0}]}{\partial m_1} - \Delta_{m_2}\frac{\partial y[m_{1_0}, m_{2_0}]}{\partial m_2}\right]^2 \tag{6}$$

Minimizing ϵ is achieved straightforwardly by differentiation with respect to Δ_{m_1} and Δ_{m_2} and setting the result to zero, which, after manipulation yields the matrix equation,

$$M\Delta = V \tag{7}$$

with,

$$\begin{aligned}
M &= \begin{bmatrix} \sum_{m_1}\sum_{m_2}\left(\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_1}\right)^2 & \sum_{m_1}\sum_{m_2}\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_1}\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_2} \\ \sum_{m_1}\sum_{m_2}\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_1}\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_2} & \sum_{m_1}\sum_{m_2}\left(\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_2}\right)^2 \end{bmatrix} \\
\Delta &= \begin{bmatrix} \Delta_{m_1} \\ \Delta_{m_2} \end{bmatrix} \\
V &= \begin{bmatrix} \sum_{m_1}\sum_{m_2}\left(y[m_1, m_2] - y[m_{1_0}, m_{2_0}]\right)\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_1} \\ \sum_{m_1}\sum_{m_2}\left(y[m_1, m_2] - y[m_{1_0}, m_{2_0}]\right)\frac{\partial y[m_{1_0},m_{2_0}]}{\partial m_2} \end{bmatrix}
\end{aligned} \tag{8}$$

The spatial derivative terms are estimated by first smoothing the observed images and then applying Prewitt operators. The resulting equations are solved to yield sub-pixel estimates of the global translation parameters Δ_{m_1} and Δ_{m_2}. Various performance tests demonstrating the efficacy of the method are presented. The motion estimates are used as parameters in the Tsai-Huang super-resolution reconstruction algorithm. The authors suggest the use of a least squares solution to the linear system of equations relating the aliased low-resolution observation image spectra to the unaliased, super-resolution spectrum, when the number of observed images results in more equations than unknowns. This leads naturally to a Moore-Penrose pseudo-inverse solution.

In summary, this paper contributes a simple method for sub-pixel accuracy *global translational motion estimation*, which is more attractive computationally than the algorithm proposed in [4]. Super-resolution reconstruction theory is not significantly furthered beyond [4].

2.3 Recursive Least Squares Techniques

We consider now an approach based on a least squares formulation for the solution of (3) which is implemented in a recursive fashion to improve computational efficiency.

Kim, Bose and Valenzuela [11] utilize the frequency domain theoretical framework as well as the global translation observation model proposed by Tsai and Huang [4], however extend the formulation to consider observation noise as well as the effects of spatial blurring. An excellent review of the frequency domain reconstruction method of [4] precedes the authors' primary contribution - a recursive least-squares, and a weighted recursive least squares solution method for solving the linear system of equations (3) in the presence of observation noise. Sufficient conditions for non-singularity of the system matrix Φ in (3) are also derived, in terms of the relative global translation parameters of the undersampled scene observations.

The proposed recursive solution approach is computationally attractive, while the least squares formulation provides the advantage of a measure of robustness in the case of an under- or over-determined system (3).

Though the authors address limitations of [4], by including LSI blur (which removes the restriction of impulse sampling), LSV blur, and observation noise, the restriction to a global translation motion observation model remains. Issues concerning estimation of the motion parameters themselves are not discussed - the translation parameters are assumed to be known *a-priori*. It is noted that zeroes in the the blur spectral response will result in the problem become ill-posed due to non-uniqueness of solution due to the loss of all information at the response zeroes. This point is not addressed.

Kim and Su [12, 13] extend their earlier work [11] by addressing the issue of the ill-posedness of the inverse problem resulting from the presence of zeroes in the blur PSF. The formulation of the SR problem is still that of [4], but includes like [11], a LSI PSF in the frequency domain observation model. Since the presence of the blur PSF in general results in an ill-posed inverse problem [14], the authors propose replacing their earlier recursive least-squares solution for (3) with a Tikhonov regularization Tikhonov77 method. The system of equations (3), which are repeated below for convenience.

$$\mathbf{Y} = \Phi\mathbf{F} \tag{9}$$

In the presence of observation noise this system will typically be inconsistent, and Φ will be ill-conditioned due to zeroes in the blur PSF. A regularized solution may be found by finding $\mathbf{F}$ which minimizes the expression,

$$\|\Phi\mathbf{F} - \mathbf{Y}\|^2 + \gamma(\mathbf{F}) \tag{10}$$

where $\gamma(\cdot)$ is a *regularization functional*. In this work, $\gamma(\cdot) = \lambda\|\mathbf{F} - \mathbf{c}\|^2$, where $\mathbf{c}$ is an approximation to the (as yet unknown) solution.

Algebraic manipulations yield the minimum of the expression in (10) as,

$$\hat{\mathbf{F}} = \left(\Phi^T\Phi + \lambda\mathbf{I}\right)^{-1}\left(\Phi^T\mathbf{Y} + \lambda\mathbf{c}\right) \tag{11}$$

Further manipulations show that it is possible to solve (11) using an iterative method and avoid the necessity for the matrix inversion step. Since the solution is unknown, $\mathbf{c}$ is initially set to zero, and thereafter set to the result of the previous iteration.

Though this paper addresses the ill-posedness of the SR inverse problem, there are several criticisms which can be leveled at the the approach taken. Firstly, the stabilizing function (squared error) is unrealistic for images, tending to result in overly smoothed solutions. Secondly the use of an estimate of the unknown solution $\mathbf{c}$, leaves unanswered questions as to the stability of the proposed recursive solution method.

Su and Kim [15] address one of the fundamental limitations of the Tsai-Huang [4] formulation - the inability to accommodate non-global motion models - yet still utilize the frequency domain framework of [4]. The proposed method allows a degree of non-global translational motion by utilizing the Tsai-Huang approach on overlapping, *motion compensated sub-blocks* of the observation image sequence.

The procedure is as follows: Observation images are decomposed into overlapping blocks; translational motion is estimated for these blocks in the observation sequence frames using [16]; Tsai-Huang reconstruction is performed for each block independently; interpolation is used to correct for occlusions and exposures; the resulting blocks are assembled to produce the final image.

The extension of [4] to accommodate non-global motion is a significant extension to the frequency domain theory, however it should be noted that the fundamental limitation to a translational motion model have not been addressed. This is due to the fact that the reconstruction phase is still the frequency domain approach which assumes translational model. There is significant question as the choice of the block decomposition, as well as the performance of the algorithm in cases where there is rapid motion, or multiple independent motion within blocks. Though effective in the test cases the algorithm presented is ad-hoc in nature, especially with regards merging the results of reconstruction of enhanced sub-blocks.

2.4 Recursive Total Least Squares Methods

A extensions of the recursive least squares work discussed in Section 2.3 is that of recursive *total* least squares which is known to provide some degree of robustness to errors in the observation model, which are likely, in the case of super-resolution reconstruction, to result from errors in motion estimation. Total least squares theory is well developed, and the reader is referred to the references [17, 18] for additional information. The work we discuss here does not, however, extend the observation model beyond that of the previous section.

Bose, Kim and Valenzuela [19, 20] extend the ideas developed in [11, 12, 13] to include a degree of robustness to errors in both the matrix Φ and the observation vector $\mathbf{Y}$ in (3). The technique for achieving this is the method of *total least squares* (TLS), [17, 18], which is known to provide favorable robustness to errors in both observations as well as the system matrix in matrix equations of the form $Ax = b$. The motivation for the use of the TLS approach is to provide a degree of robustness to errors in Φ which result from errors in the translational motion estimates required in the specification of Φ. Since it is well understood that motion estimates need be as accurate as possible to SR reconstruction, the justification for the TLS approach is clear.

The observation model used is of the form,

$$\mathbf{Y} = [\Phi + \mathbf{E}]\,\mathbf{F} + \mathbf{N} \tag{12}$$

where $\mathbf{E}$ is the error in Φ due to errors in motion estimates. The TLS problem may be formulated as a minimization problem,

$$\begin{array}{ll} \text{Minimize} & \|\,[\,\mathbf{N} \,\vdots\, \mathbf{E}\,]\,\| \\ \text{Subject to} & \mathbf{Y} - \mathbf{N} = [\Phi + \mathbf{E}]\mathbf{F} \end{array} \tag{13}$$

The Frobenius norm is used. A *recursive* total least squares (RTLS) algorithm based on [21] (see also [22]) is used to solve (13) efficiently. Since the formulation of the reconstruction problem is in the frequency domain, utilizing complex DFT coefficients, (13) is transformed to an equivalent real-valued minimization. As with their earlier work [11], the proposed algorithm is well suited to massively parallel implementation, since DFT values of $\mathbf{F}$ may be computed in a decoupled manner.

Though attention is directed to the problem of uncertainties in the global translation parameter estimates, the proposed method does not address more fundamental issues such as the inherent limitations of the underlying frequency domain approach which cannot incorporate general scene or camera motion models. Though not explicitly demonstrated in the paper, the proposed method can incorporate a LSI blur degradation of the observed images as was discussed in the earlier papers [11, 12, 13]. Motion blur due to non-zero aperture time is not considered.

Bose, Kim and Zhou [23], and Kim [24] provide a theoretical analysis of the performance of their RTLS reconstruction algorithm based on existing work in TLS theory. They show that under certain assumptions, the TLS formulation reaches the Cramér-Rao bound.

2.5 Multichannel Sampling Theorem Based Techniques

Ur and Gross propose a super-resolution reconstruction method based on the generalized sampling theorem of Papoulis [25] and a variant thereof, the multichannel sampling theorem, by Brown [26]. Although the implementation of the reconstruction is achieved in the spatial domain, the technique is fundamentally a frequency domain technique relying on the shift property of the Fourier transform to model the translation of the source imagery.

Consider a function $f(x)$ which is band-limited to $-\sigma < \omega < \sigma$ and is passed through R "mutually independent" linear channels, the outputs of which are sampled at the rate $2\sigma/R$ (under-sampled at $1/R$ of the Nyquist rate) to produce R discrete signals $y_r(mT)$, $T = R/2\sigma$, $m \in \mathbb{Z}$. By the multichannel sampling theorem, $f(x)$ may be reconstructed from $y_r(mT)$ by passing the sub-sampled signals through R linear filters, summing the outputs and performing band-limited interpolation. Specifically,

$$\hat{f}(x) = \sum_{r=1}^{R} \hat{f}_r(x) = \sum_{r=1}^{R} \sum_{m=-\infty}^{\infty} y_r(mT) \cdot h_r(x - mT) \tag{14}$$

where $\hat{f}(x)$ is a sampled version of $f(x)$ meeting the Nyquist criterion and which may be interpolated to recover $f(x)$ exactly. The R filters with impulse response $y_r(m)$ may be found according to equations given in [25] and [26]. A frequency domain derivation is used to obtain the above result. In particular, in the formulation by Brown [26] of the multichannel sampling theorem, matrices containing samples from the frequency domain representation of the R sampled signals are used to determine the set of filters $h_r(\cdot)$ which are applied in the time domain to the R degraded and sub-sampled signals to yield a sampled version of $f(t)$. This is similar to the frequency domain formulation of Tsai and Huang, in which systems of equations in the frequency domain are be solved to yield the frequency domain representation of $f(t)$. The multichannel sampling theorem yields a time domain method (multichannel linear filtering) which effects the equivalent reconstruction as is achieved in the Tsai-Huang method which operates by solving a linear system of frequency domain equations.

Ur and Gross consider the linear degradation channels to include the effects of a blur PSF as well as global translation which may be modeled as a delay. Observing that the operations of blurring and translation are commutative and assuming a single blur common to all the channels, it is show that the super-resolution problem may be separated into two distinct processes: ”merging” the under-sampled signals into a single σ-band-limited function, followed by deblurring of the merged signal. Since the deblurring operation is independent of the merging process, it is possible to derive a closed form solution for computing the merged signal from the R degraded and under-sampled channel outputs. Any standard image restoration algorithm may then be applied to the merged signal to correct for the blur degradation.

In terms of application, the Ur and Gross method assumes that the global translation parameters are known *a-priori*. No attention is directed to the motion estimation problem.

As mentioned above, the Ur-Gross approach is, in effect, a spatial domain analog of the the Tsai-Huang frequency domain formulation, the only significant difference being the inclusion of a single PSF common to all the observations. Observation noise and motion blur is not considered. Since the Ur-Gross proposal is effectively a spatial domain implementation equivalent to the frequency domain methods discussed in Section 2.2, it suffers from the same limitations in range of feasible motion models. The authors are thus limited to global motion models only, and like Tsai and Huang, considered only the simplest case of global translational motion.

2.6 Summary

Super-resolution reconstruction via the frequency domain approach discussed in the previous sections has significant advantages:

- Simplicity

  The principles behind the frequency domain approaches discussed are readily understandable in terms of basic Fourier theory. Though additional complexity exists in the implementation, the fundamental principles are clear.

- Computational complexity

  Many of the techniques discussed in this section are computationally attractive. In the Tsai-Huang [4] formulation and its derivative forms, the system of frequency domain equations are decoupled - that is a single frequency sample point of the super-resolution image may be computed independently of any other. This fact makes this formulation highly amenable to massively parallel implementation.

- Intuitive super-resolution mechanism

  The techniques discussed are all based on the de-aliasing techniques of [4]. This formulation makes explicit the underlying mechanism of resolution improvement - restoration of frequency components beyond the Nyquist limit of the individual observation image samples. Later we shall encounter methods where the mechanism of super-resolution restoration are not as abundantly obvious.

There are, however, significant disadvantages which must be addressed if more general super-resolution video restoration problems are to be tackled:

- Global translation motion model

  It is important to note that all the frequency domain methods we have discussed utilize a global translational motion model. Even [15], which describes a method with local motion by a block decomposition of the low-resolution observation images, falls into this category since only a global translation model for the decomposed blocks is possible. The global translation model we observe is no coincidence. This is a fundamental limitation of the frequency domain approach as we show in the next point.

- Inflexibility regarding motion models

  The frequency domain methods discussed require the existence of a transformation which is the Fourier domain equivalent of the spatial domain motion model. This requirement imposes severe limitations on the range of feasible motion models due to the inability to formulate Fourier domain transformations which are equivalent to spatially varying motion models typically required in sophisticated SR applications.

- Degradation models

  For reasons similar the previous point, it is also difficult to include spatially varying degradation models in the frequency domain reconstruction formulation.

- Inclusion of spatial domain *a-priori* knowledge for regularization

  Since super-resolution reconstruction is ill-posed, regularization is often required in order to achieve acceptable quality solution imagery. Though certain of the methods we have discussed include regularizing functionals, these are, as a result of the frequency domain formulation, necessarily spatially invariant. Often the most useful *a-priori* knowledge used to constrain the solution space to effect regularization is via spatial domain constraints, which are highly convenient and intuitive. Frequency domain methods do not lend themselves well to the inclusion of such constraints.

In the limited case of global translation motion, there is significant benefit in frequency domain approaches to super-resolution restoration, however since we are interested in more general classes of motion, as well as degradations, it is clear that the frequency domain approach is insufficient. In the sections that follow, we shall review the literature of spatial domain super-resolution methods, some of which completely address the disadvantages noted for frequency domain approaches.

3 Spatial Domain Methods

3.1 Introduction

Having examined a number of frequency domain approaches to super-resolution reconstruction, we turn our attention to the second major division in super-resolution literature - spatial domain methods. The papers we discuss in this section formulate the super-resolution reconstruction in the *spatial domain*. The major advantages provided by this approach include:

- General observation models, which may include:
  - Arbitrary motion models (global or non-global)
  - Motion blurring due to non-zero aperture time
  - Optical system degradations (spatially varying or invariant)
  - Effects of non-ideal sampling (spatially varying or invariant)
  - Ability to model complex degradations (such as compression blocking artifacts)
- Powerful methods for inclusion of *a-priori* constraints
  - Spatial domain image models such as Markov Random Fields
  - Set based constraints (POCS formulation)
  - Nonlinear models capable of bandwidth extrapolation.

3.2 Interpolation of Non-Uniformly Spaced Samples

In this section we consider a simple approach to constructing super-resolution images from an image sequence based on spatial domain interpolation. The low-resolution observation image sequence are registered, resulting in a composite image composed of samples on a non-uniformly spaced sampling grid. These non-uniformly spaced sample points are interpolated and resampled on the high-resolution sampling grid. Though this approach may initially appear attractive, it is, however, overly simplistic as it does not take into consideration the fact that samples of the low resolution images do not result from ideal sampling but are, in fact, spatial averages. The result is that the reconstructed image does not contain the full range of frequency content that can be reconstructed given the available low-resolution observation data.

Keren, Peleg and Brada [27] describe a spatial domain approach to image registration using a global translation and rotation model, as well as a two stage approach to super-resolution reconstruction. Here we discuss the motion estimation techniques and the first stage of the super-resolution reconstruction method which is a simple interpolation technique. The second stage, which falls into the category of the simulate and correct type methods will be discussed in Section 3.4.

Consider the continuous image pair $f_a(x_1, x_2)$ and $f_b(x_1, x_2)$ related by global translation and rotation as

$$f_b(x_1, x_2) = f_a(x_1 \cos\theta - x_2 \sin\theta + \Delta_{x_1}, x_2 \cos\theta - x_1 \sin\theta + \Delta_{x_2}) \tag{15}$$

where $\Delta_{x_1}, \Delta_{x_2}, \theta$ are the translation parameters in the x_1 dimension, translation in the x_2 dimension and rotation, respectively. By expanding the trigonometric functions in a second order Taylor series and then expanding f_a as a Taylor series about (x_1, x_2), an approximate error $\epsilon(\Delta_{x_1}, \Delta_{x_2}, \theta)$, between f_b and f_a may be determined as,

$$\epsilon(\Delta_{x_1}, \Delta_{x_2}, \theta) = \sum \left[f(x_1, x_2) + \left(\Delta_{x_1} - x_2\theta - x_1\frac{\theta^2}{2} \right) \frac{\partial f}{\partial x_1} + \left(\Delta_{x_2} + x_1\theta - x_2\frac{\theta^2}{2} \right) \frac{\partial f}{\partial x_2} \right]^2 \tag{16}$$

The summation is over overlapping portions of f_b and f_a. The minimum of $\epsilon(\Delta_{x_1}, \Delta_{x_2}, \theta)$ is found by differentiation and equating to zero, yielding a set of three equations in the three unknowns. This system is solved using an iterative method, utilizing a Gaussian pyramid to expedite processing. The result is the registration parameters $\Delta_{x_1}, \Delta_{x_2}, \theta$. It is reported that the approach yields parameters with sub-pixel accuracy provided θ is small, due to the small angle approximation to the trigonometric functions describing the image rotation.

Given R observed frames the registration procedure described above is used to extract Δ_{x_1},Δ_{x_2} and θ for each of $R - 1$ images relative to a chosen reference image. The images are registered and a high resolution reconstruction grid is imposed on the "stack" of observed images. Each pixel in the high resolution image is chosen as a outlier trimmed mean of the value of the set of observation image pixels the centers of which fall within the area of the high resolution pixel under consideration. Post-processing of the interpolated image may be effected using a de-blurring filter, under the assumption of a Gaussian blur.

The resulting image, which is a composite of the R observed low resolution images does not significantly extend the frequency domain content of the reconstructed image beyond that of any individual low-resolution image since the reconstruction is predominantly the result of linear combination of the data which is known not to extend frequency content of the signal. Though the paper does contribute early ideas concerning image registration, the image interpolation procedure is overly simplistic in its approach to the super-resolution problem. No attempt is made to de-alias the observation data - a feature of all frequency domain methods in Section 2 and though perhaps not as explicitly, by many of the spatial domain methods we shall discuss. In effect, a only a fraction of the available *information* is utilized in the reconstruction. The proposed method does not provide a systematic framework for achieving super-resolution reconstruction. There is little indication as to the nature of the image improvement and no analysis of the method is proffered.

Another interpolation based approach is proposed in an early paper of Aizawa, Komatsu and Saito [28]. They examine the problem of acquiring high-resolution imagery from stereo cameras. Though ostensibly distinct from the problem of super-resolution reconstruction from an image sequence, it is in fact an analog in that an image pair captured using a single camera with relative camera/scene motion produces, like a stereo pair, similar, but not identical views of the scene. Of course the range of possible motion is more general in the case of temporally distinct images, however under the observation model assumptions used here, this is not a significant complication.

By considering the spatial frequency response of a regularly spaced two dimensional array of sampling apertures (CCD pixels) the authors show that there is a limit (Nyquist's theorem) to the extent of frequency components which may be reconstructed from the sampled image. By considering the possibility of sampling at spatial positions *between* the array pixels, it is demonstrated that the effective frequency response of the combined (double image) system is increased.

In practice this is implemented by utilizing two (and in their later papers, more) cameras which are positioned so as to provide images which are shifted by sub-pixel distances relative to the first (reference) image. The image from the second camera is registered relative to the reference image sampling lattice to sub-pixel accuracy using gradient based or block matching techniques. Due to effects such as perspective projection, differing camera alignment, etc., the samples from the second camera do not fall regularly between the reference sampling grid, but are non-uniformly spaced. The result is a set of regular samples (from the reference image) amongst which are interspersed samples from the second image. In order to reconstruct a high resolution frame, the authors are thus faced with the problem of interpolation and resampling to obtain the desired high resolution image. This is effected using a method for reconstruction from non-uniformly spaced samples [29]. We demonstrate the idea in the one dimensional case. Consider the function $f(x)$ sampled at times x_m which are non-uniformly spaced. If a one-to-one mapping $\gamma(\cdot)$ exists such that $mT = \gamma(x_m)$ and if $f(\gamma^{-1}(x))$ is band-limited to $\omega_0 = \pi/T$ then $f(x)$ may be reconstructed as:

$$f(x) = \sum_{m=-\infty}^{\infty} f(x_m) \frac{\sin[\omega_0(\gamma(x) - mT)]}{\omega_0(\gamma(x) - mT)} \tag{17}$$

This is immediately recognized as band-limited sinc interpolation with a nonlinear mapping used to correct for the non-uniform spacing of the samples. The case for two dimensions entails minor additional complications regarding the choice of $\gamma(\cdot, \cdot)$, but the principle remains the same.

In later papers by the authors, the ideas discussed here are extended to the case of multiple cameras (closer to our multiple image reconstruction problem), and an iterative simulate and correct approach based on an approach proposed by Landweber [30] is used to estimate the uniformly spaced samples of the super-resolution reconstruction. We shall discuss these papers later in this section.

Tekalp, Ozkan and Sezan [8] propose, amongst other techniques discussed in Sections 2.2 and 3.6.1, a two step procedure where the upsampling of the low resolution images and the restoration to correct for a sensor PSF are performed sequentially. The low resolution frames are registered and combined to form a non-uniformly sampled high resolution image which is then interpolated and resampled on a uniform grid to produce the reconstructed high resolution frame. They suggest the use of the thin-plate spline method of Franke [31], the iterative method of Sauer and Allebach [32] or the POCS based method of Yeh and Stark [33] to effect the interpolation and resampling.

Since the interpolated image still suffers from sensor blur, as well as errors resulting from interpolation process itself, a restoration step is applied. The authors suggest the use of any of the commonly known deconvolution methods that take into account the presence of noise.

Aizawa, Komatsu, Saito and Igarashi extend [28] by introducing an iterative super-resolution reconstruction procedure based on their merging and interpolation work. In [34, 35, 36], the problem of acquiring high-resolution imagery from two or more cameras is considered and is extended [37, 38] to include cameras which differ in the dimensions of the CCD array sampling apertures. The objective is the acquisition of images beyond the physical resolution limitations of the individual sensor arrays. As discussed earlier, this problem is related to that of super-resolution image reconstruction problem from a single camera image sequence.

As in [28], the R low resolution observation images are merged, using motion compensation, to form a non-uniformly sampled high resolution image $\tilde{\mathbf{z}}$. Desired, however, is the uniformly sampled image $\mathbf{z}$. Various interpolation based techniques are considered, [29, 39, 40, 41], however a method based on the Landweber algorithm [30] is used to iteratively estimate the uniformly sampled high resolution image from the non-uniformly sampled image produced by the registration process. The desired, uniformly sampled image $\mathbf{z}$ and the merged image $\tilde{\mathbf{z}}$ are related by,

$$\tilde{\mathbf{z}} = \mathbf{A}\mathbf{z} \tag{18}$$

where $\mathbf{A}$ represents the non-uniform sampling process. Inversion of (18) is infeasible due to the dimensionality, and likely singularity of $\mathbf{A}$. Instead, an iterative procedure which corrects for the non-uniform sampling process is used.

An initial guess $\mathbf{z}^{(0)}$ of the high resolution image (with uniformly spaced samples) is made, and is then iteratively updated by the recurrence formula,

$$\mathbf{z}^{(j+1)} = \mathbf{z}^{(j)} + \alpha \cdot \mathbf{A}^{*} \circ \left(\tilde{\mathbf{z}} - \mathbf{A} \circ \mathbf{z}^{(j)}\right) \tag{19}$$

where $\mathbf{A}$ is the non-uniform sampling process, $\mathbf{A}^{*}$ is the adjoint operator of $\mathbf{A}$, α is a control parameter, and $\tilde{\mathbf{z}}$ is the non-uniformly spaced, merged image. The parameter α is chosen so as to ensure that (19) is convergent.

In Section 3.4 we will encounter an identical Landweber type iteration, however it is important to note that the operation represented by $\mathbf{A}$ then models the *imaging process* itself, not the mapping between a pair of super-resolution images as is the case here.

The authors later propose a single camera temporal integration imaging procedure which utilizes deformable, quadrilateral based, motion estimation [42, 43]. This work is, in effect, an image region registration and interpolation procedure with limited ability for super-resolution.

The techniques we have discussed in this section have the advantage of simplicity but do not address several important issues. The observation models used are generally unrealistic as they do not properly account for the effects of optical blurring, motion blurring or noise. Though [8] did consider an observation PSF, there is much question as to the optimality of separated merging and restoration phases of super-resolution reconstruction. None of the methods discussed include *a-priori* information in an attempt to resolve the non-uniqueness of solution typical in the super-resolution problem.

3.3 Algebraic Filtered Backprojection

An early algebraic tomographic filtered backprojection approach to super-resolution reconstruction is that of Frieden and Aumann, [44]. The authors do not consider the problem of super-resolution image reconstruction from an image sequence, but the related problem of super-resolution image reconstruction from multiple 1-D scans of a stationary scene by a linear imaging array. Noting that the PSF in the 1-D scan system represents a line integral and that of the multiple image super-resolution problem represents an area integral, it is clear that the problems differ only in the form of the imaging system PSF. In [44], the linear imaging array detectors are assumed to be larger than the limiting resolution of the optical system. The imaging geometry provides overlapping scans of a given scene area, enabling reconstruction at a resolution higher than the limiting spatial sampling rate of the sensor array.

The imaging process is described by the discrete model,

$$\mathbf{y} = \mathbf{H}\mathbf{z} \tag{20}$$

where $\mathbf{y}$ is a vector of the observed data from the successive image scans, $\mathbf{H}$ is the system matrix which describes the imaging geometry, and $\mathbf{z}$ is a vector containing the unknown image to be estimated. Observation noise and/or a lack of data, as well as the considerable size of the matrix $\mathbf{H}$ preclude direct inversion to estimate $\mathbf{z}$. Instead the authors propose that an estimate be formed using the *backprojection* of the observation data as follow:

$$\hat{\mathbf{z}} = \mathbf{H}^T\mathbf{y} \tag{21}$$

It is important to note that the backprojection matrix $\mathbf{H}^T$ is *not* equivalent to the inverse $\mathbf{H}^{-1}$, indeed we have, using (20),

$$\hat{\mathbf{z}} = \mathbf{H}^T\mathbf{H}\mathbf{z} \tag{22}$$

If and only if $\mathbf{H}^T\mathbf{H} = \mathbf{I}$ will the backprojection be equivalent to the inverse. In the given application, the matrix $\mathbf{H}^T\mathbf{H}$ is Toeplitz, and thus associated with a well defined point spread function, allowing fast Fourier domain inverse filtering to restore the backprojected observation data. This is indeed the technique known as *filtered backprojection* in the tomography literature, though in the given application, the point spread function differs from that typically found in the case of computed tomographic reconstruction.

It is instructive to notice that the application of an inverse filter for $\mathbf{H}^T\mathbf{H}$ yields an estimate,

$$\hat{\mathbf{z}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{y} \tag{23}$$

which is immediately recognized as the minimum mean square error (MMSE) estimate for $\mathbf{z}$.

Frieden and Aumann make no allowances for the presence of observation noise. This has serious consequences since inverse filtering is well known to be highly noise sensitive due to the increasing amplitude response of the inverse filter with increasing frequency. The proposed reconstruction method does not fall into the class of regularized methods and as such cannot be expected to yield acceptable performance in the presence of noise or missing data. Despite these serious drawbacks, this paper has the distinction of being an early attempt to apply tomographic image reconstruction techniques to the super-resolution reconstruction problem.

3.4 Simulate and Correct Methods

We now consider a large class of super-resolution reconstruction methods which have in common a simulate and correct approach to reconstruction. Given an estimate of the super-resolution reconstruction and a model of the imaging process, the imaging process is simulated using the super-resolution estimate as the input to produce a set of simulated low resolution observation images. These images are compared with the actual observations, an error is computed and used to correct the estimate of the super-resolution image. This simulate/correct process is iterated until some stopping condition is met - typically the minimization of some error criterion between the simulated and observed images.

For a given observation model there is considerable freedom in the choice of the mechanism for back-propagating the error to determine the correction term for the super-resolution estimate, as well as the error function to be minimized. In this section we will examine several approaches. Despite these differences, it is possible to cast the reconstruction procedure into a single theoretical framework, with the differences in approaches evident in the choice of the projection and backprojection operators as well as the error functional to be minimized.

Consider the observation model relating the R low resolution observation images described by the stacked vector $\mathbf{Y}$ of lexicographically ordered pixels of each image, to the (unknown) discretized scene $\mathbf{F}$ via the operator $\mathbf{H}$ which typically performs motion compensation and subsampling;

$$\mathbf{Y} = \mathbf{H}\mathbf{f} \tag{24}$$

Given an estimate $\hat{\mathbf{z}}$ of the super-resolution reconstruction of the scene $\mathbf{f}$, it is possible to *simulate* the low resolution observation images by applying (24) to the super-resolution estimate $\hat{\mathbf{z}}$ as,

$$\hat{\mathbf{Y}} = \mathbf{H}\hat{\mathbf{z}} \tag{25}$$

Iterative, simulate and correct procedures discussed here update the estimate of the super-resolution reconstruction by *backprojecting* the error between the simulated low resolution images $\hat{\mathbf{Y}}^{(j)}$ and the observed low resolution images $\mathbf{Y}$ via the backprojection operator $\mathbf{H}^{BP}$ which models the apportioning of "blame" for errors in the projection, to pixels in the super-resolution estimate $\hat{\mathbf{z}}^{(j)}$. Typically $\mathbf{H}^{BP}$ is designed to *approximate* the inverse of the operator $\mathbf{H}$.

$$\begin{aligned} \hat{\mathbf{z}}^{(j+1)} &= \hat{\mathbf{z}}^{(j)} + \mathbf{H}^{BP}\left(\mathbf{Y} - \hat{\mathbf{Y}}^{(j)}\right) \\ &= \hat{\mathbf{z}}^{(j)} + \mathbf{H}^{BP}\left(\mathbf{Y} - \mathbf{H}\hat{\mathbf{z}}^{(j)}\right) \end{aligned} \tag{26}$$

It is interesting to note that the simulate/correct approach bears close similarity to many well known problems. In computed tomographic reconstruction, the operator $\mathbf{H}$ is typically a projection matrix which describes the interaction of the x-ray with the object to be reconstructed, and $\mathbf{H}^{BP}$ is a back-projection operator. In the solution of integral equations, (26) is seen to be of the form of the Landweber iteration [30]. The choice of the projection and backprojection operators in (26) varies widely in the papers we discuss.

Peleg, Keren and Schweitzer [45] and Keren, Peleg and Brada [27] present the initial ideas which would, in their later papers, be developed into the tomographic backprojection type algorithms for super-resolution reconstruction. In [45], super-resolution reconstruction from a set of globally translated images of an unchanging 2-D scene is considered, while in [27], a more general global translation and rotation model is used. The method for estimating the motion parameters for this model was discussed in Section 3.2. In [45, 27], the effects of a spatially invariant PSF is included.

An iterative procedure for finding a super-resolution reconstruction is proposed. Given an estimate of the super-resolution image, a simulated imaging process yields a set of low resolution images which are compared with the observed low resolution images. The super-resolution estimate is then modified so as to reduce the error between the observed and simulated images. The process is terminated when the error has reached some predetermined level, or after a given number of iterations.

The R observed images are denoted $y_r[m_1, m_2]$ with $r = 1, 2, \ldots, R$. A super-resolution image, coincident with one of the R observed images, is to be reconstructed from the low-resolution sequence. The j^{th} estimate of the super-resolution image is given by $\hat{z}^{(j)}[n_1, n_2]$. Assuming the imaging process which resulted in the observed images is known, it is possible to simulated the imaging process with the estimate $\hat{z}^{(j)}[n_1, n_2]$ as an approximation to the original scene, yielding R low resolution simulated images $\hat{y}_r^{(j)}[m_1, m_2]$. The error between the j^{th} iteration simulated images and the original observed images is defined as,

$$\epsilon^j = \sum_{r=1}^{R} \sum_{m_1, m_2} \left| \hat{y}_r^{(j)}[m_1, m_2] - y_r[m_1, m_2] \right| \tag{27}$$

Since it is desired that the simulated images be identical to the observed images, the error ϵ^j must be minimized. Peleg *et. al.* propose the following modification of the current estimate $\hat{z}^{(j)}[n_1, n_2]$. For some pixel $[n_1, n_2]$ if $\hat{z}^{(j)}[n_1, n_2] = l$, then consider modifications $\hat{z}^{(j)}[n_1, n_2] = l - 1$ and $\hat{z}^{(j)}[n_1, n_2] = l+1$. Compute for these three pixel values $\{l-1, l, l+1\}$ the corresponding simulated low resolution images and the corresponding errors. Choose the pixel value for which the error in (27) is a minimum. Repeat the process for each pixel in the image over several iterations.

Notice that since this very simple optimization attempts to monotonically decrease the error on a pixel by pixel basis it is highly likely to converge to only a local minimum. This problem is addressed by utilizing a simulated annealing [46] approach for optimization. Though it has been shown in [47] that simulated annealing is provably convergent to the global minimum given a suitable *annealing schedule*, the required schedule is too slow to be of practical utility. Selecting an appropriate annealing schedule to deliver useful performance in acceptable time may be difficult, and since the process is run for a finite time, there remains a finite probability of attaining a local minimum even with this algorithm. Convergence of the proposed optimization is thus slow.

The issue of the uniqueness of the solution is not addressed. There are likely to be many images which, under the imaging process, yield a set of simulated images which match the observation data. The proposed framework provides no manner in which to favor certain solutions over others. Furthermore, the particular solution image reached in the space of possible solutions will be dependent on the initial estimate for the super-resolution image, as well as the order in which pixels are updated in a given super-resolution image estimate. The effects of noise are inadequately considered.

The simulated imaging process is equivalent to a projection operation in the computed tomography framework. Back-projection of the error is not, however obvious. The modification of the solution estimate typically makes more sophisticated use of the error information available. In this case, the error is simply a scalar and no attempt is made to utilize the per-pixel errors. Modification of the estimate is nonetheless made on an error minimizing basis.

Irani and Peleg [48, 49] extend the earlier work [45, 27] by improving the means of backprojecting the error between the simulated low resolution images and the observed data. They modify the error functional (27) to the form,

$$\epsilon^j = \sqrt{\sum_{r=1}^{R} \sum_{m_1,m_2} \left(y_r[m_1,m_2] - \hat{y}_r^{(j)}[m_1,m_2] \right)^2} \tag{28}$$

They use consider a simulated imaging process which includes a point spread function to model degradation in the imaging system. The primary modification to [45, 27] is the use of a backprojection kernel in the update of the estimate $\hat{z}^{(j)}[n_1,n_2]$. The update procedure is intuitively understood as follows: The value of each high-resolution pixel $\hat{z}^{(j)}[n_1,n_2]$ is updated according to *all* low resolution pixels which are dependent on its value (via the simulated imaging process). The contribution of a low resolution pixel $\hat{y}_r^{(j)}[m_1,m_2]$ is the error $y_r[m_1,m_2] - \hat{y}_r^{(j)}[m_1,m_2]$ multiplied by a factor which measures the relative contribution of the the high resolution pixel $\hat{z}^{(j)}[n_1,n_2]$ to the low resolution pixel $y_r[m_1,m_2]$. The result is that errors in low resolution pixels which are strongly influenced by a particular high resolution pixel contribute more strongly to to corrective term applied to the estimate $\hat{z}^{(j)}[n_1,n_2]$. The correction may be written in the form,

$$\hat{z}^{(j+1)}[n_1,n_2] = \hat{z}^{(j)}[n_1,n_2] + \sum_{r=1}^{R} \sum_{\zeta} \left(\hat{y}_r[m_1,m_2] - \hat{y}_r^{(j)}[m_1,m_2] \right) h^{BP}[n_1 - m_1, n_2 - m_2] \tag{29}$$

where ζ is the set of low resolution pixels $[m_1,m_2]$ in the j^{th} low resolution image dependent on the high resolution pixel $\hat{z}^{(j)}[n_1,n_2]$. $h^{BP}[\cdot,\cdot]$ is a *backprojection kernel* which scales the correction to the high resolution estimate depending on the influence of a given high resolution pixel on the low resolution pixel for which the error is computed. The authors observe that h^{BP} may be chosen so as to influence the solution to which the iteration converges; thus the backprojection kernel may be used to incorporate additional constraints on the solution such as smoothness. It is important to note, however, that these constraints are all linear in form. We shall encounter non-linear constraints in later papers. A theorem providing a sufficient condition on h^{BP} for convergence in the special case of de-blurring, is proved.

Irani and Peleg extend their earlier work [48, 49] in [50] where they propose a very general procedure for super-resolution reconstruction of scenes which contain arbitrary independent motion. The primary contribution of [50] is not in the super-resolution reconstruction algorithm per sé, which is unchanged from [48, 49], but in the incorporation of a multiple motions tracking algorithm which allows super-resolution reconstruction for partially occluded objects, transparent objects or some object of interest.

The details of the motion detection and estimation algorithms are discussed in in detail in [51], but due to the relevance of the proposed techniques, we shall summarize the principles. The proposed approach uses an iterative procedure to identify and track multiple, independent motions occurring in the scene. A single motion is first computed and an object corresponding to this motion is identified and tracked. This is called the “dominant” motion, and object, respectively. Once the dominant object is identified and tracked it is excluded from the region of analysis, and the process is repeated to identify other objects and their motions. The procedure is terminated when no further objects are identified. Object motion may be modeled using any 2-D parametric motion model. The procedure is also extended to include a “temporal integration” approach which builds an internal representation for each moving object over the course of several frames.

Shah and Zakhor [52, 53] utilize a Landweber [30] reconstruction method similar to that of Irani and Peleg. Their primary contribution, however, is not in the area of reconstruction techniques, but is a novel approach to resolving problems caused by unreliable motion estimates in the super-resolution restoration. A per-pixel block-based motion estimation method in which a *set* of possible motion vectors are stored for each pixel in each low resolution frame is suggested.

The authors illustrate how standard block matching techniques may select incorrect *motion* estimates even though they maximize some correlation measure. Indeed motion estimation is an ill-posed problem. Often multiple candidate motions are equally “optimal” with respect to the block correlation measure. Since the effectiveness of the super-resolution reconstruction scheme is critically dependent on the accuracy of the estimated motion vectors [7], erroneous motion estimates are highly undesirable, resulting in degraded performance in the form of objectionable motion artifacts. The situation can arise where the true motion is rejected in favor of an incorrect estimate which has a higher block correlation measure. By maintaining a *set* of “near optimal” matches, recourse to other feasible motion estimates is possible if the initial motion estimate for a particular pixel is found to be inconsistent with those for surrounding pixels.

The simulate and correct type methods we have discussed provide a useful strategy for solving the super-resolution problem by providing a mechanism for constraining the super-resolution reconstruction to conform to the observation data. A complication concerns the uniqueness of the simulate/correct type solution given the fact that the super-resolution problem is ill-posed. Many solutions that satisfy the constraints given by the observed low resolution frames exist. Though the algorithms discussed here are convergent, they are therefore not necessarily convergent to a unique solution. Indeed the solution found may be a function of the initial estimate for the Landweber iteration, or may be dependent on the order of pixel updates. This is usually undesirable.

A powerful method for resolving this non-uniqueness of solution is the use of *a-priori* knowledge which constrains the solution space. This is not easily achieved within the simulate/correct framework. Though Irani and Peleg citeIrani90,Irani91 suggest that choice of the backprojection kernel may impose smoothness constraints, their approach is neither flexible or general.

We therefore turn our attention to probabilistic methods (Section 3.5) and the POCS formulation (Section 3.6.1) where *a-priori* constraints may be directly included. It is interesting to note that though the Bayesian and POCS methods do not employ the Landweber type iteration which we encountered here, they do ensure that the super-resolution reconstruction is constrained to match the observed data. In this sense these methods are equivalent to those discussed in this section. They differ though in that the reconstruction need *also* match the *a-priori* constraints.

3.5 Probabilistic Methods

In this section we discuss one of the major research tracks in super-resolution reconstruction. Since super-resolution is an ill-posed inverse problem, techniques which are capable of including *a-priori* constraints are well suited to this application. In recent years, Bayesian methods, which inherently include *a-priori* constraints in the form of prior probability density functions, have enjoyed increasing popularity and are now central to the solution of ill-posed inverse problems in a wide range of applications. The Bayesian approach is synonymous with *Maximum A-Posteriori* estimation which is the starting point for our discussion of probabilistic techniques. We shall also briefly address other methods.

3.5.1 MAP Reconstruction Methods

Given the general observation model,

$$\mathbf{Y} = \mathbf{H}\mathbf{f} + \mathbf{N} \tag{30}$$

The *Maximum A-Posteriori* (MAP) approach to estimating $\mathbf{f}$ seeks the estimate $\hat{\mathbf{f}}_{\mathrm{MAP}}$ for which the *a-posteriori* probability, $\Pr\{\mathbf{f}|\mathbf{Y}\}$ is a maximum. Formally, we seek $\hat{\mathbf{f}}_{\mathrm{MAP}}$ as,

$$\hat{\mathbf{f}}_{\mathrm{MAP}} = \arg\max_{\mathbf{f}} \left[\Pr\{\mathbf{f}|\mathbf{Y}\}\right] \tag{31}$$

Applying Bayes' rule yields,

$$\hat{\mathbf{f}}_{\mathrm{MAP}} = \arg\max_{\mathbf{f}} \left[\frac{\Pr\{\mathbf{Y}|\mathbf{f}\}\Pr\{\mathbf{f}\}}{\Pr\{\mathbf{Y}\}}\right] \tag{32}$$

and since the maximum $\hat{\mathbf{f}}_{\mathrm{MAP}}$ is independent of $\mathbf{Y}$ we have,

$$\hat{\mathbf{f}}_{\mathrm{MAP}} = \arg\max_{\mathbf{f}} \left[\Pr\{\mathbf{Y}|\mathbf{f}\}\Pr\{\mathbf{f}\}\right] \tag{33}$$

Since the logarithm is a monotonic increasing function, this is equivalent to finding,

$$\hat{\mathbf{f}}_{\mathrm{MAP}} = \arg\max_{\mathbf{f}} \left[\log\Pr\{\mathbf{Y}|\mathbf{f}\} + \log\Pr\{\mathbf{f}\}\right] \tag{34}$$

where $\log\Pr\{\mathbf{Y}|\mathbf{f}\}$ is the *log-likelihood function* and $\log\Pr\{\mathbf{f}\}$ is the log of the *a-priori density* of $\mathbf{f}$. Since $\mathbf{Y} = \mathbf{H}\mathbf{f} + \mathbf{N}$, it is easy to see that the likelihood function is determined by the probability density of the noise $f_{\mathbf{N}}(\cdot)$ as,

$$\Pr\{\mathbf{Y}|\mathbf{f}\} = f_{\mathbf{N}}(\mathbf{Y} - \mathbf{H}\mathbf{f}) \tag{35}$$

Typically since the noise is assumed to be Gaussian, then the use of the natural logarithm in the above derivation removes the exponential term from the density $f_{\mathbf{N}}(\cdot)$. Additionally, it is common to utilize a Markov Random Field prior which has a Gibbs probability density of the form,

$$\Pr\{\mathbf{f}\} = \frac{1}{\mathcal{Z}}\exp\left\{-\frac{1}{\beta}\mathcal{U}(\mathbf{f})\right\} \tag{36}$$

where $\mathcal{Z}$ is a normalizing constant, β is the so called "temperature" parameter and $\mathcal{U}(\mathbf{f})$ is the "energy" of $\mathbf{f}$. The use of the logarithm in the formulation for the MAP solution thus greatly simplifies manipulations in these cases.

If the noise is assumed to be Gaussian and the prior is chosen to to be a convex function of $\mathbf{f}$, then it is easily seen [54] that the optimization of (31) is convex, so that the solution $\hat{\mathbf{f}}_{\mathrm{MAP}}$ is assured to exist and is unique. This is a very significant advantage of the MAP formulation.

We begin our discussion of MAP reconstruction methods with Schultz and Stevenson. The relatively detailed development will provide a clear theoretical framework which in essence describes many of the existing MAP reconstruction methods, allowing other work to be mentioned in terms of differences from this work.

Schultz and Stevenson extend their earlier work [55, 56] on Bayesian (MAP) image interpolation for improved definition using a Huber Markov Random Field (HMRF) to the problem of super-resolution image reconstruction in [54, 57, 58, 7]. They propose a motion compensated subsampling matrix based observation model which accounts for both scene and camera motion which occurs between images acquisitions. Blurring due to motion during a non-zero aperture time is not considered in the their early work. Here we summarize the observation model, which remains essentially unchanged in their later papers.

Assume that p (odd) low-resolution frames, $y[m_1, m_2, k]$ with $k \in \{c-\frac{p-1}{2}, \ldots, c, \ldots, c+\frac{p-1}{2}\}$ and $m_1 \in \{1, 2, \ldots, M_1\}$, $m_2 \in \{1, 2, \ldots, M_2\}$ are observed. The objective is to reconstruct a super-resolution image $f[n_1, n_2, c]$ with $n_1 \in \{1, 2, \ldots, qM_1\}$, $n_2 \in \{1, 2, \ldots, qM_2\}$ and $q \in \mathbb{N}$, coincident with $y[m_1, m_2, c]$, the center frame in the observed image sequence. A subsampling model, which models the spatial integration of sensors in the detector array, is proposed for the c^{th} observed frame:

$$y[m_1, m_2, c] = \sum_{n_1=qm_1-q+1}^{qm_1} \sum_{n_2=qm_2-q+1}^{qm_2} f[n_1, n_2, c] \tag{37}$$

This relationship for the c^{th} (center) frame may be more succinctly written using lexicographic ordering of the LR and SR images as,

$$\mathbf{y_c} = \mathbf{H_c f_c} \tag{38}$$

where $\mathbf{H_c} \in \mathbb{R}^{M_1 M_2 \times q^2 M_1 M_2}$ is the subsampling matrix relating the SR image $\mathbf{f_c}$ with the observed frame $\mathbf{y_c}$. The remaining observed images $\mathbf{y_k}$ are related to $\mathbf{f_c}$ via *motion-compensated* subsampling matrices which compensate for the effects of motion occuring between frames as,

$$\mathbf{y_k} = \mathbf{H_k f_c} + \mathbf{u_k},\ \text{for } k \in \{c - \frac{p-1}{2}, \ldots, c-1, c+1, \ldots, c + \frac{p-1}{2}\} \tag{39}$$

$\mathbf{H_k} \in \mathbb{R}^{M_1 M_2 \times q^2 M_1 M_2}$ is the motion-compensated subsampling matrix which relates the k^{th} LR observation to the SR image $\mathbf{f_c}$ which is temporally coincident with the center frame in the LR sequence. The vector $\mathbf{u_k}$ contains pixels which are unobservable from $\mathbf{f_c}$, but present in $\mathbf{f_k}$. The elements of $\mathbf{u_k}$ are not known since $\mathbf{f_c}$ is unknown. Notice that rows of $\mathbf{H_k}$ which contain useful information are those for which elements of $\mathbf{y_k}$ are observed *entirely* from motion compensated elements of $\mathbf{f_c}$. To improve robustness, rows for which this is not true are removed, yielding a reduced set of equations,

$$\mathbf{y'_k} = \mathbf{H'_k f_c} \tag{40}$$

In practice $\mathbf{H'_k}$ is unknown and must be estimated from the observed LR frames $\mathbf{y_k}$ and $\mathbf{y_c}$. This results in,

$$\mathbf{y'_k} = \mathbf{\hat{H}'_k f_c} + \mathbf{n_k} \tag{41}$$

where $\mathbf{n_k}$ contain the errors resulting from the use of the estimate $\mathbf{\hat{H}'_k}$. The elements of $\mathbf{n_k}$ are assumed to be independent identically distributed Gaussian random variables.
With p observed frames we have the system of equations:

$$\begin{array}{llllll}
\mathbf{y}'_{\mathbf{c}-\left(\frac{\mathbf{p}-1}{2}\right)} & = & \mathbf{\hat{H}}'_{\mathbf{c}-\left(\frac{\mathbf{p}-1}{2}\right)} & \mathbf{f_c} & + & \mathbf{n}_{\mathbf{c}-\left(\frac{\mathbf{p}-1}{2}\right)} \\
\vdots & & \vdots & \vdots & & \vdots \\
\mathbf{y}'_{\mathbf{c}-1} & = & \mathbf{\hat{H}}'_{\mathbf{c}-1} & \mathbf{f_c} & + & \mathbf{n}_{\mathbf{c}-1} \\
\mathbf{y_c} & = & \mathbf{H_c} & \mathbf{f_c} & + & \mathbf{0} \\
\mathbf{y}'_{\mathbf{c}+1} & = & \mathbf{\hat{H}}'_{\mathbf{c}+1} & \mathbf{f_c} & + & \mathbf{n}_{\mathbf{c}+1} \\
\vdots & & \vdots & \vdots & & \vdots \\
\mathbf{y}'_{\mathbf{c}+\left(\frac{\mathbf{p}-1}{2}\right)} & = & \mathbf{\hat{H}}'_{\mathbf{c}+\left(\frac{\mathbf{p}-1}{2}\right)} & \mathbf{f_c} & + & \mathbf{n}_{\mathbf{c}+\left(\frac{\mathbf{p}-1}{2}\right)}
\end{array} \tag{42}$$

Which may be written as a stacked set of equations,

$$\mathbf{Y} = \mathbf{H f_c} + \mathbf{N} \tag{43}$$

The SR image $\mathbf{f_c}$ is estimated using the MAP criterion as,

$$\mathbf{\hat{f}_c} = \arg\max_{\mathbf{f_c}} \left[\log \Pr\left\{\mathbf{f_c} | \{\mathbf{y_k}\}\right\}\right] \tag{44}$$

which after applying Bayes' rule may be written,

$$\mathbf{\hat{f}_c} = \arg\max_{\mathbf{f_c}} \left[\log \Pr\left\{\mathbf{f_c}\right\} + \log \Pr\left\{\{\mathbf{y_k}\} | \mathbf{f_c}\right\}\right] \tag{45}$$

Schultz and Stevenson use the Huber Markov random field (HMRF) for the prior term $\log \Pr\{\mathbf{f_c}\}$, which is a discontinuity preserving image model, which allows edge reconstruction while imposing smoothness constraints on reconstruction. It is assumed that motion estimation errors between frames are independent thus, the likelihood term may be written in the form $\Pr\{\{\mathbf{y_k}\}|\mathbf{f_c}\} = \prod_{\mathrm{k}} \Pr\{\mathbf{y_k}|\mathbf{f_c}\}$. Taking into account that $\mathbf{H_c}$ is known exactly, finding $\mathbf{\hat{f}_c}$ requires the solution of the *constrained* optimization,

$$\begin{array}{ll}
\text{Find} & \mathbf{\hat{f}_c} = \arg\max_{\mathbf{f_c} \in \mathcal{F}} \left[\log \Pr\left\{\mathbf{f_c}\right\} + \log \Pr\left\{\{\mathbf{y_k}, \mathrm{k} \neq \mathrm{c}\} | \mathbf{f_c}\right\}\right] \\
\text{Subject to} & \mathcal{F} = \left\{\mathbf{f} : \mathbf{H_c f} = \mathbf{y_c}\right\}
\end{array} \tag{46}$$

This constrained optimization is solved using a gradient projection technique, the details of which may be found in [7].

The observation model assumes local translation motion which is computed using block matching for motion estimation. A hierarchical block motion estimation techniques is used for both computation efficiency as well as to ensure consistency in the motion estimates.

Schultz and Stevenson have also successfully applied their super-resolution reconstruction technique to the problem of scan conversion - where interlaced scan images are combined to produce motion compensated, de-intelaced, super-resolution images [59].

Schultz, Stevenson and Meng, noting the importance of accurate motion information for super-resolution reconstruction have examined techniques for sub-pixel motion estimation.

In [60, 61] the MAP super-resolution reconstruction procedure of [7] is reviewed, and various motion estimation methods are studied. In particular, the block motion estimation procedure of [7], an eight parameter projective model proposed based on [62], as well as the Horn Schunck [63] optical flow method are discussed and utilized in various super-resolution reconstruction experiments. Methods for detecting and eliminating inaccurate motion vector estimates are also presented. The result is improved super-resolution reconstructions since objectionable artifacts resulting from incorrect motion vectors are minimized.

Noting that motion estimation is itself and ill-posed problem, Schultz and Stevenson [64] propose the use of a regularized motion estimation procedure which favors smoothness of the motion fields while still allowing discontinuities. This is achieved using a Bayesian motion estimation techniques and Huber-Markov random field *a-priori* smoothness constraints with Horn Schunck optical flow estimation. Super-resolution image reconstructions from an image sequence with motion estimated using the Bayesian motion estimator are compared with those computed using the block motion estimation technique of [7] and is found to produce visually superior results.

It is interesting to note that in [60, 64, 61] motion estimates are computed from only image *pairs*. It is reasonable to anticipate improved reliability in the motion estimates computed using multiple images. Also of note is the fact that the motion estimation techniques are not object based. No attempt is made, as was the case with [50] to construct internal representations for multiple independent motions within the scene. These points should be addressed.

Hardie, Barnard and Armstrong [65] present a MAP super-resolution reconstruction procedure which is essentially identical in form to that of Schultz and Stevenson [7]. They consider the cases of global as well as non-global motion estimation. For the former, they use the translation and rotation model of [49] and for the latter, they draw on optical flow techniques [66] with foreground/background segmentation. The super-resolution formulation is equivalent to that of Schultz and Stevenson under the assumption of Gaussian noise and a Gauss-Markov random field prior model. This work is extended in [67] to consider the problem of simultaneous estimation of both the super-resolution image $\hat{\mathbf{f}}$ and the motion parameters $\hat{\mathbf{s}}$. A MAP formulation similar to that of [65] is used, but included in the optimization are the unknown motion parameters as,

$$\hat{\mathbf{f}}, \hat{\mathbf{s}} = \arg\max_{\mathbf{f},\mathbf{s}} \left[\Pr\left\{ \mathbf{f}, \mathbf{s} | \mathbf{Y} \right\} \right] \tag{47}$$

Since estimation of $\hat{\mathbf{f}}$ requires the motion estimates $\hat{\mathbf{s}}$, a procedure is suggested in which $\hat{\mathbf{f}}$ and $\hat{\mathbf{s}}$ are estimated alternately, that is, $\hat{\mathbf{s}}$ is fixed, and $\hat{\mathbf{f}}$ estimated, and then $\hat{\mathbf{f}}$ fixed and $\hat{\mathbf{s}}$ estimated. Though this procedure can be slow to converge, it does have significant advantages. Motion is no longer estimated from the low resolution images directly as is the case in most super-resolution algorithms, instead, the motion parameters are chosen as those which minimize the error between simulated low resolution observation images (as computed using the observation model and the current super-resolution image estimate), and the observed low resolution frames. The priors used for the super-resolution image and the motion are Gaussian. The optimization in (47) is achieved using iterated coordinate descent. The authors observe that it is not possible to use the faster conjugate gradient method, since the cost functional is non-constant from iteration to iteration due to the joint estimation of both the motion parameters and the super-resolution reconstruction.

A more detailed exposition of a observation model, which includes global rotation, translation and a sensor and optical system PSF may be found in [68] where image registration is performed using the method of Irani and Peleg [49] and Lucas and Kanade [69]. In this paper, which estimates motion in a separate step from the super-resolution reconstruction phase, gradient descent and conjugate gradient descent methods are used to maximize the posterior probability. The details of these methods are presented.

Lorette, Shekarforoush and Zerubia [70] also present a MAP formulation of the super-resolution reconstruction problem. Elsewhere [71] they consider the problem of 3-D super-resolution reconstruction, but [70] study three non-linear penalty functions for their Markov-random field prior model which allow edge reconstruction while ensuring smoothness. It should be noted that Huber MRF of Schultz and Stevenson performs exactly this function.

Cheeseman, Kanefsky, Kraft and Stutz [72, 73, 74] have been applying MAP super-resolution reconstruction to Viking, Voyager and more recently Mars Pathfinder imagery. Their formulation assumes Gaussian noise and utilized a prior terms which leads to a linear system of equations which are solved using Gauss-Jacobi methods.

Hunt, Sementilli, Mariappan, Marcellin, Nadar, Yuen and Sheppard have pursued work on a Poisson MAP super-resolution formulation and have provided a theoretical analysis of the performance of this super-resolution technique in [75, 76, 77, 78, 79, 80]. This work concentrates on restoration of astronomical images and as such make fundamental assumptions concerning bounded support of the reconstructed objects which are not valid in the more general problem super-resolution video restoration. Furthermore, the assumption of Poisson photon statistics in the observation are not relevant since imagery captured for video super-resolution do not, in general, have low photon counts. A substantial portion of this work concentrates on super-resolution restoration techniques which are well known within the astronomical community but are of questionable applicability to our problem. This body of research is noted here primarily for completeness, as well as a source for theoretical ideas.

3.5.2 Maximum Likelihood and Expectation Maximization based Algorithms

Tom and Katsaggelos [81, 82] examine the super-resolution reconstruction problem as composed of three separate steps - registration of the low resolution images, restoration of these images followed by an interpolation step which yields the super-resolution reconstruction. In [81], a simultaneous registration and restoration approach is proposed in an effort to improve super-resolution reconstruction. The work is based on ideas on multichannel image identification and restoration in developed in [83]. In [84] these ideas are extended to perform combined registration, restoration and interpolation.

These papers are essentially based on a maximum likelihood (ML) formulation of the registration, restoration and interpolation phases of super-resolution reconstruction. A fundamental problem with this approach is the fact that ML estimation is poorly suited to the solution of ill-posed inverse problems due high noise sensitivity when the reconstruction problem is under-specified. Throughout this section, the importance of regularized solutions using *a-priori* constraints has been emphasized as a necessary condition for yielding high-quality reconstructions. These papers do not sufficiently address this issue despite the useful ideas contributed in terms of simultaneous restoration and registration.

Elad and Feuer [85, 86] have also considered (along with other techniques discussed later) a ML estimation formulation, but the same criticisms apply to this work. In fact they modify the ML estimate to include a regularization functional in [85], which in effect is equivalent to a MAP formulation with a Gaussian prior or the Tikhonov-Arsenin [87] type regularization. It is also suggested the the simulate and correct methods yield the ML estimate of the super-resolution reconstruction.

3.5.3 Summary

The probabilistic reconstruction techniques we have considered here; the MAP formulation utilizing Markov random field prior models in particular; have become one of the most popular approaches for super-resolution reconstruction. There are several reasons for this:

- The MAP framework allows direct incorporation of *a-priori* constraints on the solution - essential for finding high quality solutions to the ill-posed super-resolution inverse problem.
- Markov random field priors have become almost standard as they provide a highly convenient, intuitive and realistic method for modeling typical image characteristics using only a local neighbor interaction model. With judicious choice of the form of the penalization of the activity measure in the MRF model, priors may be defined which provide the smoothness constraints necessary to ensure acceptable reconstruction under conditions of limited or poor data, but also allow effective reconstruction of edges.
- Optimization using the MAP framework and convex cost functionals in MRF priors leads to a globally convex optimization problem, ensuring not only the existence and uniqueness of the solution, but also enabling the use of well understood and efficient descent type algorithms.
- Simultaneous motion estimation and restoration is possible within the MAP framework.

The MAP formulation is thus one of the most promising and flexible approaches to super-resolution image reconstruction. It's only competitor in terms of convenience and flexibility are the POCS based methods which we discuss next.

3.6 Set Theoretic Methods

3.6.1 Projection Onto Convex Sets

One of the prominent approaches to super-resolution reconstruction is based on the method of projection onto convex sets. In this formulation, constraint sets are defined which limit the feasible solution space of the super-resolution reconstruction. Constraints are defined as convex sets in the vector space $\mathbb{R}^{N_1 \times N_2}$ which represents the space containing all possible super-resolution reconstructions. Sets that represent desirable characteristics of the solution are defined, such as positivity, bounded energy, fidelity to data, smoothness and so on. The solution space of the super-resolution reconstruction problem is thus the intersection of the convex constraint sets. Projection onto convex sets (POCS) refers to an iterative procedure which, given any point in the vector space, locates a point which satisfies all the convex constraint sets.

Given k convex constraint sets in $\mathbb{R}^{N_1 \times N_2}$ such that the intersection of the sets is non-empty, POCS projects a point in the vector space onto each constraint set, repeating until a point is reached which is in the intersection of the k sets. It can be shown that provided the constraint sets are convex that this iteration converges. A detailed theoretical discussion of the POCS method may be found in [88]. Discussions on the use of POCS techniques in image restoration are presented in [89, 90].

POCS has attracted much attention in recent years in a multitude of image reconstruction and restoration applications. Three reasons for this stand out:

- Simplicity

  POCS is very intuitive and generally simple to implement. The only potential source of difficulty is the determination of the projection operators.

- Flexible Spatial Domain Observation Model

  Because the POCS method is typically formulate in the spatial domain, very general motion and observation models may be used. The complexity of the motion and observation model has little impact of the POCS solution procedure.

- Powerful inclusion of *a priori* information

  Perhaps the most useful aspect of the POCS formulation is the ease with which *a-priori* information may be included. It is generally simple to define convex constraint sets which incorporate desired solution characteristics. These sets may impose restrictions such as positivity or bounded energy which are difficult to represent in terms of cost functionals.

We begin our examination of POCS based techniques with a detailed exposition of an earlier paper which demonstrates many of the fundamental ideas. This will enable a more concise explanation of later work.

Stark and Oskoui [91] propose an early POCS based solution to super-resolution image reconstruction problems. In particular, the paper addresses the scanning linear array problem originally discussed by Frieden and Aumann (see Section 3.3) as well as the problem of restoring a super-resolution image from multiple plane array images. We begin our discussion of the POCS reconstruction technique with a brief outline of the proposed imaging model. We develop the mathematics for the case of a time invariant image. It will be clear how the method is extended to the case of time varying imagery with multiple image acquisitions.

The optical system projects an image $f(x,y)$ onto the image sensor array which is assumed to be a regular array of M_1 by M_2 sensor elements. The output y_i of the i^{th} detector in the sensor array with spatial response characteristic $\sigma_i(x,y)$ is given by,

$$y_i = \int\int_{-\infty}^{\infty} f(x,y)\sigma_i(x,y) \quad , \quad 1 \le i \le M_1 M_2 \tag{48}$$

This integration over the continuous spatial variables (x,y) may be discretized on the super-resolution reconstruction grid $[n_1, n_2]$ yielding,

$$y_i = \sum_{n_1}\sum_{n_2} f[n_1,n_2]\sigma_i[n_1,n_2] \quad , \quad 1 \le n_1 \le N_1 \ , \ 1 \le n_2 \le N_2 \tag{49}$$

Where $N_1 > M_1$, $N_2 > M_2$ are the dimensions of the super-resolution reconstruction array. The i^{th} detector spatial response characteristic $\sigma_i(x,y)$ is discretized to yield $\sigma_i[n_1,n_2]$ which is the fractional area of the super-resolution pixel $[n_1,n_2]$ contained within the response region of the i^{th} low-resolution detector. This assumes a uniform, unity response of the detector over its response region. In particular,

$$\sigma_i[n_1,n_2] = \begin{cases} 0 & \text{if SR pixel } [n_1,n_2] \text{ is completely outside of } i^{\text{th}} \text{ detector response region} \\ 1 & \text{if SR pixel } [n_1,n_2] \text{ is completely within } i^{\text{th}} \text{ detector response region} \\ r_i & (0 < r_i < 1) \text{ if SR pixel } [n_1,n_2] \text{ is partially within } i^{\text{th}} \text{ detector response region} \end{cases} \tag{50}$$

Lexicographic ordering of $f[n_1,n_2]$ and $\sigma_i[n_1,n_2]$ yields column vectors **f** and $\sigma_{\mathbf{i}}^T$ allowing us to instead write,

$$y_i = \sum_j f_j \sigma_{i_j} \quad , \quad 1 \le j \le N_1 N_2 \tag{51}$$

Or more succinctly,

$$y_i = \boldsymbol{\sigma}_{\mathbf{i}}^T \mathbf{f} \tag{52}$$

Now consider the sets C_i as,

$$C_i = \left\{ \mathbf{f} : \boldsymbol{\sigma}_{\mathbf{i}}^T \mathbf{f} = y_i \right\} \quad , \quad 1 \le i \le M_1 M_2 \tag{53}$$

The set C_i in (53) is the set of all discretized super-resolution images for which the response of the i^{th} sensor in the detector array is y_i, the observed value. The set C_i is defined for each pixel observation i , $1 \le i \le M_1 M_2$, and thus there are a total of $M_1 M_2$ such sets for a single image observation. These sets place a constraint on the possible values which may be assumed by the solution **f** – in particular (53) ensures that any $\mathbf{f} \in C_i$, $1 \le i \le M_1 M_2$ is constrained to be consistent with the measured datum y_i. Noticing that the constraints on **f** are linear, it is in principle possible to obtain a sufficient number of equations, so that a solution for **f** may be found by matrix inversion.

For reconstruction of a super-resolution image with $N_1 N_2$ pixels by matrix inversion, $N_1 N_2$ *independent* observation equations are required. Notice immediately that since $N_1 > M_1$, $N_2 > M_2$, data from more than one image acquisition is required to obtain a sufficient number of equations to enable solution by matrix inversion. Indeed the requirement to obtain $N_1 N_2$ *independent* equations implies that novel information must be present data obtained in the acquisition of the additional images. This novel information is present in small differences in the acquired images usually resulting from scene or camera motion. In practice however, direct matrix is infeasible.

In reality, additional image observations contribute further sets of the form of (53). These additional constraints augment the under-determined system of equations derived for a single image so that the system approaches the fully-determined case.

The collection of sets C_i are often called data consistency constraints and in this case can be shown to be closed and convex. This allows the definition of a projection operator $\mathcal{P}_i$ as follows:

$$\mathcal{P}_i \mathbf{f} = \begin{cases} \mathbf{f} & \text{if } \boldsymbol{\sigma}_\mathbf{i}^T \mathbf{f} = y_i \\ \mathbf{f} + \frac{y_i - \boldsymbol{\sigma}_\mathbf{i}^T \mathbf{f}}{\boldsymbol{\sigma}_\mathbf{i}^T \boldsymbol{\sigma}_\mathbf{i}} \boldsymbol{\sigma}_\mathbf{i} & \text{otherwise} \end{cases} \tag{54}$$

$\mathcal{P}_i \mathbf{f}$ is the projection of the point $\mathbf{f}$ onto the set C_i. In the POCS formulation, an initial guess $\mathbf{f}^{(\mathbf{0})}$ for the super-resolution image is projected onto each the constraint sets C_i to ensure consistency with each measured datum y_i and where K is the number of pixel measurements.

$$\mathbf{f}^{(\mathbf{n}+\mathbf{1})} = \mathcal{P}_1 \mathcal{P}_2 \mathcal{P}_3 \cdots \mathcal{P}_K \mathbf{f}^{(\mathbf{n})} \tag{55}$$

This set of projections is applied repeatedly to yield an updated estimate of the super-resolution image. Closedness and convexity of the constraint sets ensure convergence of the iteration to some point satisfying all the constraint sets [88]. It is imperative to note however, that this point is in general non-unique, but is a point on the surface of the convex polytope defined by the intersection of the constraint sets. The solution is in general dependent on the initial guess.

It is interesting to note that application of the POCS method using only the data consistency constraints in (53) is identical to the Algebraic Reconstruction Technique found in the computed tomography literature [92].

In addition to the data consistence constraints, additional constraints which represent *a-priori* knowledge of the form of the solution may be included. In particular the range of values in the solution image may be constrained to the set C_A,

$$C_A = \{\mathbf{f} : \alpha \leq f_i \leq \beta \ , \ \alpha < \beta\} \tag{56}$$

Another possibility is to ensure that the solution be contained in a set of solutions with bounded energy as:

$$C_E = \{\mathbf{f} : \|\mathbf{f}\| \leq E\} \tag{57}$$

Or if the solution is known to be similar to some reference image $\mathbf{f_R}$ then the solution may be constrained to the sets,

$$C_R = \{\mathbf{f} : \|\mathbf{f} - \mathbf{f_R}\| \leq \epsilon_R\} \tag{58}$$

Bounded support of the solution may also be imposed by defining a set of points A for which the solution is required to be zero, such as

$$C_S = \{\mathbf{f} : f_i = 0 \text{ for } i \in A\} \tag{59}$$

Additional constraints on the solution may be defined in a similar manner. The iteration in (55) is augmented to include the additional constraint sets. The inclusion of prior knowledge in this fashion constrains the solution space thus enabling robust performance in the presence of noise, inconsistent data or missing data.

This paper is significant as it is the first application of the POCS method to the problem of super-resolution reconstruction problem. There are, however, several drawbacks to the approach taken in the paper, many of which are addressed in later work. For now, we note that the proposed observation model does not incorporate noise. The observation model for the k^{th} single observed image is of the form,

$$\begin{aligned} \mathbf{y_k} &= \mathbf{H_k f} \\ \mathbf{H_k} &= \left[\boldsymbol{\sigma}^T_{\mathbf{1_k}} \ \boldsymbol{\sigma}^T_{\mathbf{2_k}} \ \cdots \ \boldsymbol{\sigma}^T_{\mathbf{M_{1_k} M_{2_k}}}\right]^T \end{aligned} \tag{60}$$

which may be written in the case of p image observations as,

$$\begin{aligned} \mathbf{Y} &= \mathbf{Hf} \\ \mathbf{Y} &= \left[\mathbf{y}^T_{\mathbf{1}} \ \mathbf{y}^T_{\mathbf{2}} \ \cdots \ \mathbf{y}^T_{\mathbf{k}} \ \cdots \ \mathbf{y}^T_{\mathbf{p}}\right]^T \\ \mathbf{H} &= \left[\mathbf{H}^T_{\mathbf{1}} \ \mathbf{H}^T_{\mathbf{2}} \ \cdots \ \mathbf{H}^T_{\mathbf{k}} \ \cdots \ \mathbf{H}^T_{\mathbf{p}}\right]^T \end{aligned} \tag{61}$$

Tekalp, Ozkan and Sezan [8] propose three different super-resolution reconstruction approaches. Here we will concentrate on the POCS based approach which extends the earlier work by Stark and Oskoui [91]. The other proposed approaches, which are a frequency domain formulation and a interpolation/restoration method are discussed in Section 2.2 and 3.2 respectively. The primary contribution of this paper is an observation model which includes noise. In particular, the authors propose a single image observation model of the form,

$$\mathbf{y} = \mathbf{Hf} + \mathbf{n} \tag{62}$$

where the system matrix $\mathbf{H}$ includes the effects the sensor PSF. Only global translational motion is modeled. In order to account for the observation noise, the data consistency constraint for the i^{th} pixel observation (as proposed by Stark and Oskoui in (53)) is modified to the form,

$$C_i = \{\mathbf{f} : |r_i| < \delta_0\} \tag{63}$$

where r_i is the i^{th} element of the residual,

$$\mathbf{r} = \mathbf{y} - \mathbf{Hf} \tag{64}$$

Noticing that

$$\mathbf{y} - \mathbf{Hf} = \mathbf{n} \tag{65}$$

It is clear that the residual $\mathbf{r}$ is nothing more than the observation noise $\mathbf{n}$ and so δ_0 represent a confidence in the observation and may be set according to the noise statistics. The authors proceed to define the required projection operator for the set defined in (63) and show the results of experiments where a super-resolution image is reconstructed from several globally translated, undersampled frames. Though this paper has addressed the problem of observation noise, the motion model used assumes only global translation between image acquisitions and does not consider the effects of motion blur. This is not, however, a result of any inherent limitation with the POCS based as we see in [93].

Patti, Sezan and Tekalp address the shortcomings of [8] in [93]. This paper utilizes the same general observation model $\mathbf{y} = \mathbf{Hf} + \mathbf{n}$ and a POCS formulation for the solution of these equations. The primary development in this paper is the motion model included in the system matrix $\mathbf{H}$ which departs significantly from previous work. Most significantly, the matrix $\mathbf{H}$ incorporates a space-varying PSF (SVPSF) which models the degradations caused by sampling with a low-resolution sensor, as well as the effects of relative motion between the sensor and the scene (which includes blurring due to a non-zero aperture time). The SVPSF results from three modeling stages; first the blur caused by the relative motion between the scene and the sensor during the aperture time; next the effects of the physical dimensions and response characteristics of the low resolution sensors; and finally the effects of sampling by the sensor array. The combined model defines the system matrix $\mathbf{H}$.

The motion model proposed, which appears to be inspired by Sawchuk [94] allows for the inclusion of a very general class of scene motion. In particular, the authors show how it is in theory possible to include general scene motion, provided occlusions are not present. A theoretical derivation illustrate how the SVPSF may be found for this general case. This model is illustrated using the example of temporally piecewise constant motion velocities, with the case of global translational motion fully developed and used in experiments. Reconstruction of the super-resolution image is effected using a POCS framework which follows their earlier work [8].

In summary, this paper provides a POCS based framework for super-resolution reconstruction which accounts for aliasing, sensor blur, motion blur and additive noise. It is interesting to note that spatial blurring due to motion occurring during the aperture time had not been addressed prior to this work. Additionally this work relaxed the constraint that blur and noise affecting each acquired image be LSI.

In addition to their work addressing the super-resolution reconstruction problem, Patti, Sezan and Tekalp have also published in the related but not identical problem of standards conversion, [95], in which the objective is robust spatial and temporal resampling of a video signal in one spatio-temporal sampling lattice to some other lattice. This work extends [8, 93] to include an arbitrary input spatio-temporal sampling lattice. The form of the discrete observation model is unchanged from the SVPSF model proposed in [93]. A POCS formulation identical to that found in [93] is utilized for reconstruction. In [96] reconstruction of super-resolution stills from interlaced video sequences is addressed. This paper contains several extensions in terms of motion detection and analysis, utilizing both a global affine dominant motion model, estimated using gradient based techniques as well as segmented local motion regions.

Full details of the work of Patti, Sezan and Tekalp may be found in [97] which consolidates and elaborates on the earlier papers [8, 93, 95]. The observation model incorporates a spatially varying point spread function which takes into account, camera and scene motion, motion blurring resulting from non-zero aperture time, the physical dimensions of LR sensor elements and blurring caused by imaging optics. Sensor noise is considered. A new development is the inclusion of sampling of the continuous scene on an arbitrary space-time lattices. A detailed derivation of the motion modeling is provided as well as the techniques used for motion estimation. Reconstruction is via the POCS method discussed previously. The results of several experiments are presented, demonstrating the capabilities of the proposed method.

It is interesting to note that the simulate and correct methods discussed in Section 3.4 are a special case of the POCS procedure where the data consistency constraint (53) requires equality rather that allowing an error tolerance.

The extension of observation models to include arbitrary local motion compounds the problem of reconstruction artifacts caused by unreliable motion estimates. Estimating local motion is generally less reliable than the case for global motion estimates. This is due simply to a lack of data. Global motion may often be reliably estimated since a large number of image pixels (usually every pixel in the image) are utilized in the computation. In the case of local motion, however, an estimate must be computed using a small neighborhood of pixels and is therefore less reliable. Since it is necessary to obtain sub-pixel accurate motion estimates for super-resolution reconstruction, it is imperative that motion estimates used be accurate. When the reliability of the motion estimates is questionable, it is important that they do not adversely affect the accuracy of the super-resolution reconstruction. In [98], Eren, Sezan and Tekalp extend [97] to consider means of improving the robustness of the super-resolution reconstruction methods to errors in motion estimates. To do this, two new concepts are introduced, a *validity map* and a *segmentation map*. The validity map disables projections based on observations with inaccurate motion estimates, while the segmentation map enables object-based tracking and processing where more accurate object-based motion models may be used to improve the quality of the reconstruction. The segmentation map also allows reconstruction of specific objects in manner similar to [50].

3.6.2 Bounding Ellipsoid Method

A variant of the POCS based formulation using an ellipsoid to bound the constraint sets has been investigated by Tom and Katsaggelos [99, 100] and briefly mentioned by Elad and Feuer [85, 86]. Given a set of ellipsoidal constraint sets, a bounding ellipsoid is computed [101]. The centroid of this ellipsoid is taken as the super-resolution estimate. Since direct computation of this point is infeasible, an iterative procedure is used. It is interesting to note that this approach takes a form closely related to regularized methods. The observation model used in this work is similar to that proposed by Schultz and Stevenson [7].

3.6.3 Summary

We comment here primarily on the POCS based methods. At the outset of the discussion concerning POCS methods, we noted the following advantages to the POCS formulation:

- Simplicity
- Flexible Spatial Domain Observation Model
- Powerful inclusion of *a priori* information

Our discussion clearly demonstrated these points. POCS is not, however, without disadvantages:

- Non-Uniqueness of Solution

  The solution space for the POCS method is defined as the intersection of the convex constraint sets. Unless the intersection is a point set, the solution is therefore *non-unique*. This is a serious drawback of the POCS based formulation. In the MAP formulation, convex priors were favored in order to ensure the existence and uniqueness of the solution. In POCS, this is typically impossible.

- Dependence of Solution on Initial Guess

  Worse yet is the fact that the solution determined using the POCS framework is dependent on the (arbitrary) initial estimate, as well as the order of application of the projections. POCS repeatedly projects the current estimate of the super-resolution image onto the constraint sets. The solution is the first point in the vector space of possible solutions which satisfies all the constraint sets. When not using relaxed projections, the solution lies on the surface of the volume of intersection of the constraint sets.

- Computation Cost

  POCS based methods require considerable computation and a large number of iterations to achieve convergence. An analysis of POCS methods shows that without the use of relaxed projections or similar modifications, POCS can be very slow to converge as the solution is approached. For this reason POCS is commonly not run to convergence but to a point where the estimate is visually acceptable.

It may be noted that the bounding ellipsoid method does have a unique solution within the intersection of the convex constraint sets. The problem is that this solution is completely arbitrary. There is little reason why the centroid should be favored amongst possible solutions within the set intersection. In contrast, in the MAP formulation, the solution is unique and maximizes that *a-posteriori* density.

3.7 Hybrid ML/MAP/POCS Methods

Some work has been undertaken on combined ML/MAP/POCS based approaches to super-resolution reconstruction. In particular the desirable characteristics of MAP estimation and those of the very flexible POCS method could be combined in a hybrid optimization.

Elad and Feuer [85, 86], after reviewing several existing super-resolution techniques, propose a hybrid ML/POCS based method which uses the statistical ML formulation to pose super-resolution as a statistical estimation problem, while utilizing projections based constraints to effect regularization.

It is interesting to note, however, that this approach is evident in the earlier work of Schultz and Stevenson [7] where a constrained MAP super-resolution reconstruction algorithm, in which an alternating projections based constraint and maximization of the *a-posteriori* density iteration is used. Schultz and Stevenson did not, however, explicitly identify this approach as novel.

The idea is to minimize the a-posteriori density or likelihood function (in the MAP and ML estimation frameworks respectively) while ensuring that the solution remains within constraint sets specified to reduce the feasible solution space. In [7] the constraint set ensured that downsampling of the super-resolution image exactly matched the reference frame of the low resolution image sequence.

The hybrid ML/MAP/POCS optimization approach is highly promising as it combines the most favorable characteristics of statistical methods (optimal estimation theoretic solution, mathematical rigor and direct inclusion of *a-priori* constraints) and POCS based approaches (powerful mechanism for inclusion of linear and nonlinear, set theoretic *a-priori* constraints).

The observation model used by Elad and Feuer is closely based on that of Schultz and Stevenson [7].

3.8 Optimal and Adaptive Filtering Methods

Several researchers have proposed inverse filtering approaches to super-resolution reconstruction. Here we briefly review efforts in this area. These techniques are considered primarily for completeness, as several are sub-optimal in terms of inclusion of *a-priori* constraints.

Jacquemod, Odet and Goute [102] proposed a simple deconvolution restoration approach that assumes sub-pixel translational motion. A deconvolution filter suitable for restoration of merged observation images is determined. This approach is poorly suited to the incorporation of more general observation models and is limited in terms of inclusion of *a-priori* constraints.

Erdem, Sezan and Ozkan [103] have proposed a LMMSE filtering approach, the motion compensated multiframe Wiener filter, for restoration of image sequences degraded by LSI spatial blur and additive noise. A global translation model is assumed, but motion blurring is not incorporated. Though the motion and degradation models are limited, and non-linear *a-priori* constraints are difficult to incorporate in their formulation, this paper is noteworthy in that *simultaneous* multiframe restoration is undertaken. There are several reasons why this approach is attractive, as we shall discuss in Section 5. A similar framework is discussed by Srinivas and Srinath [104].

Techniques based on adaptive filtering, especially the Kalman filter, have also seen application in super-resolution reconstruction. Patti and Sezan [105] and Elad and Feuer [106, 107] are examples. In [105] a motion compensated model Kalman filter capable of super-resolution reconstruction under spatially varying blurs is proposed. Though their Kalman filtering formulation is computationally efficient, it is, in effect, still a linear minimum mean square error estimator. Nonlinear image modeling constraints which provide bandwidth extrapolation cannot be easily incorporated into this framework. Similar comments apply to the work of Elad and Feuer [106, 107] which is reduced to a LMMSE estimate of the super-resolution image. Kim [24] dedicates a chapter of his doctoral dissertation to a similar Kalman filtering based approach to super-resolution reconstruction.

3.9 Tikhonov-Arsenin Regularized Methods

Due the the ill-posedness of the super-resolution problem, Hong, Kang and Katsaggelos [108, 109] have proposed the use of Tikhonov-Arsenin [87] regularization for super-resolution reconstruction. This is a deterministic regularization approach utilizing regularization functionals to impose smoothness constraints on the space of feasible solutions. The regularization functional used is, in effect, equivalent to a Gaussian Markov random field (GMRF) prior in the Bayesian (MAP) framework. GMRF priors are well known to produce overly smoothed reconstructions [110]. Indeed the poor performance of the GMRF prior is the reason why priors such as the Huber MRF and Generalized Gaussian MRF, which impose smoothness while still enabling edge reconstruction, were introduced. This work is thus a limited, special case of the more general Bayesian framework such as that of Schultz and Stevenson [7]. The observation model used in these papers is similar to the motion compensated subsampling matrix of [7].

3.10 Summary of Spatial Domain Methods

Super-resolution reconstruction via the spatial domain approach addresses many of the shortcomings of frequency domain approaches:

- Motion models

  Spatial domain methods, using the linear observation model of the form $\mathbf{Y} = \mathbf{H}\mathbf{f}$ are capable of including an almost unlimited range of motion models. Since the matrix $\mathbf{H}$ models observation image pixels as a linear combination of *any* combination of pixels in the super-resolution image $\mathbf{f}$, there is enormous flexibility in the formulation of the motion model. There is no limit, as was the case with frequency domain approaches to super-resolution reconstruction to global models. It is just as simple to include a local motion model as a global model using the spatial domain formulation.

- Degradation models

  The system matrix $\mathbf{H}$ also allows almost trivially simple inclusion of linear degradations such as motion blurring resulting from a non-zero aperture time (modeled as spatial integration over the motion trajectory), spatially varying or invariant blurs, missing pixels and so on. It is extremely cumbersome, if not impossible to include such degradations using the frequency domain super-resolution reconstruction framework.

- Inclusion of spatial domain *a-priori* knowledge for regularization

  As we have discussed at length, inclusion of *a-priori* information is necessary for the solution of ill-posed inverse problems such as super-resolution reconstruction. Markov random fields as well as the spatial domain POCS formulation provide almost trivially simple, yet very powerful methods to incorporate *a-priori* constraints into the reconstruction process. Working with spatial domain constraints is highly intuitive and direct. Furthermore, the ability with MRF models to provide smoothness along with edge preservation (impossible in the frequency domain) is highly desirable.

- Powerful mechanism for Bandwidth Extrapolation

  The combination of data from multiple images, as well as the use of realistic *a-priori* constraints on the reconstructed image endow spatial domain methods with a powerful mechanism for image bandwidth extrapolation. MRF models which utilize non-linear penalty functions are especially useful as they are capable of introducing novel frequency information. It is even possible for spatial domain methods to extrapolate frequency information beyond the diffraction limitations of the optical system.

- Theoretical Framework

  Probabilistic methods, especially the MAP estimation method, provide a solid mathematical framework within which further theoretical developments can be made. Theory applying to optimality of estimators is directly applicable, as well as bounds on errors and so on. This is a rich framework which is not available in the frequency domain approach.

Spatial domain methods do however come at some cost:

- Simplicity

  Unlike the frequency domain approach where reconstruction was a relatively simple process, the optimizations involved in spatial domain methods are more complex than their frequency domain counterparts.

- Computational complexity

  The increased flexibility of spatial domain methods tend to come at the cost of much increased computational requirements. This is especially true of methods which utilize non-convex priors for which one must resort to slow simulated annealing or graduated non-convexity approaches.

- Intuitive super-resolution mechanism

  In the frequency domain formulation, the mechanism for super-resolution reconstruction was abundantly clear - dealiasing of shifted, undersampled frames. For spatial domain techniques the mechanism for resolution enhancement is not as obvious.

It should be clear from this discussion that for super-resolution reconstruction of scenes involving anything more than global translational motion, that spatial domain techniques are the preferred approach.

We have discussed the advantages and disadvantages common to spatial domain techniques in general, but have not yet directed our attention to a closer scrutiny as to which of the techniques hold the most promise for the future. We shall do this in Section 4.

4 Summary and Comparison of Techniques

We begin by presenting a tabular form comparison of the two main classes of super-resolution reconstruction approaches - frequency and spatial domain. The table is divided into two sections. The upper portion deals with the formulation of the observation, motion and degradation models, while the lower portion makes generalizations concerning the solution approaches. It is important to realize that these are generalizations, and as a result, exceptions exist within the wide range of techniques we have discussed. We choose the most optimistic approaches within both the frequency and the spatial domain methods.

	Frequency Domain	**Spatial Domain**
Observation model	Frequency domain	Spatial domain
Motion models	Global translation	Almost unlimited
Degradation model	Limited	LSI or LSV
Noise model	Limited	Very flexible, even spatially varying
SR Mechanism	Dealiasing	Dealiasing & BW extrapolation using *a-priori* constraints
Simplicity	Very simple	Generally complex
Computational Cost	Low	High
A-priori constraints	Limited	Almost unlimited
Regularization	Limited	Excellent
Extensibility	Poor	Excellent
Performance	Good for specific applications	Good

From the table it is evident that apart from the computational and technical complexity of spatial domain methods, these methods are superior to the frequency domain super-resolution formulation. Within the class of spatial domain super-resolution reconstruction methods, two major techniques stand out as most promising; the Bayesian (MAP) approach and the set theoretic POCS methods. These are compared in the table below:

	Bayesian (MAP)	**POCS**
Theoretical Framework	Rich	Limited
A-priori constraints	Prior PDF (typically convex) Easy to incorporate No “hard” constraints	Convex Sets Easy to incorporate Very powerful yet simple
SR solution	Unique MAP estimate	Non-unique Volume of intersection of sets
Optimization	Iterative, standard methods Good convergence	Iterative Often slow convergence
Computational Cost	High	High
Complications	Optimization difficult for non-convex priors	Projection operators can be difficult to define

In terms of motion estimation, several important observations may be drawn from the papers reviewed:

- Sub-pixel motion

  Sub-pixel relative motion between observation frames contributes novel information which provides constraints for the solution of the inverse problem.

- Accuracy of motion estimates

  It is essential that motion estimates be accurate to sub-pixel dimensions. This is obvious in light of the previous point.

- Density of motion estimates

  Motion estimates need not necessarily be dense in order to effect super-resolution reconstruction. It is more important that motion estimates be accurate than dense. Intuitively, motion estimates computed in near constant valued regions are unlikely to be accurate, but are also of limited utility to the super-resolution algorithm since little resolution enhancement is possible in such regions. The most fruitful motion estimates may be found where they can be most reliably estimated - in regions of high spatial variance.

- Regularized motion estimation

  Motion estimation is itself an ill-posed problem. Given realizations from an image sequence, the motion estimates are unlikely to be unique. Regularization can be fruitfully applied to ensure consistent motion maps, especially with non-parametric motion models.

- General motion models

  We encountered several motion models including global translation, rotation, affine and projective, as well as various non-parametric models. Model (object) based motion estimation featured in later papers, allowing super-resolution of objects subject to partial occlusion, transparency, under motion, and so on.

5 Directions for Future Research

In this section we identify three research areas in which developments promise improvements in the state of the art of super-resolution video reconstruction - motion estimation, degradation models and restoration algorithms. For each of these we list specific avenues of investigation and provide some justification as why performance gains can be expected.

- Motion Estimation Techniques

  Super-resolution enhancement of arbitrary scenes containing global as well as multiple independent object motion is the ultimate goal of this research effort. Achieving this is critically dependent on the development of robust, model based, sub-pixel accuracy motion estimation and segmentation techniques. This is presently an open research problem. However the following directions of research are expected to yield returns:

  - Review and analyze sub-pixel motion estimation techniques.
  - Investigate errors in motion estimates computed from noisy, undersampled frames.
  - Examine multiframe motion estimation techniques to improve accuracy and reliability.
  - In the case of non-parametric motion, utilize constrained motion estimation to ensure consistency in motion maps.
  - Utilize regularized motion estimation methods to achieve this.
  - Investigate the use of non-dense motion maps.
  - Consider mechanisms for achieving motion estimation with reliability measures which should assist in reconstruction.
  - Develop general scene, and multiple independent motion identification and tracking. This is expected to require:
    - global and local motion models.
    - iterative motion estimation, identification and segmentation.
    - independent model based motion for independently moving objects.
    - motion prediction.
  - Examine simultaneous motion estimation and super-resolution reconstruction approaches.

- Degradation Models

  Observation models which accurately account for degradations occurring in the imaging system (and thus accurately describe the relationship between the super-resolution image and the observed data) promise improved super-resolution reconstructions. For this reason we anticipate that work in the following areas will lead to improvements in reconstruction quality.

  - Color imagery

    Few of the existing super-resolution techniques address the problem of super-resolution restoration of color video. This problem differs from the single band case as there exists significant correlation between color bands which should be incorporated in the observation model. Since applying super-resolution techniques to each band independently is sub-optimal, color super-resolution restoration should be investigated.

  - Compressed image sequences

    The requirement for digital storage of video sequences has led to the emergence of several lossy image compression schemes. Source data in these formats are degraded via color subsampling and quantization effects. Super-resolution restoration of such sequences promises greatest returns if these degradations can be effectively modeled.

  - Low quality video camera sequences

    Consumer video cameras typically yield low quality images. Since restoration of sequences captured using such equipment is a likely application scenario, an effort should be made to understand the degradations typical of these devices. Degradations inherent in the magnetic media recording and playback process should be modeled, as well as the effects of the CCD array (see next point).

  - CCD modeling

    CCD arrays are currently the most common image plane array used in video cameras. It is interesting to note, however, that little work has been undertaken which attempts to model the degradations occurring in these devices. This is especially relevant in consumer electronic devices where low price necessitates performance compromises. In particular effort should be focused in modeling of:

    * sensor geometry.
    * spatio-temporal integration characteristics.
    * noise and readout effects.

- Restoration Algorithms
  - Hybrid MAP/POCS restoration

    In the previous section we determined that spatial domain methods, in particular the Bayesian and set theoretic POCS super-resolution restoration methods were most promising. We noted that each of these approaches provided significant strengths. Though the techniques are quite different it is possible to combine the MAP and POCS formulations to yield a hybrid super-resolution reconstruction technique similar to those discussed in Section 3.7, thereby gaining the benefits of both methods. The hybrid method is MAP based but with constraint projections inserted into the iterative maximization of the *a-posteriori* density in a generalization of the constrained MAP optimization of [7].
  - Simultaneous motion estimation / restoration

    Several papers have considered simultaneously estimating motion and other observation model parameters within the reconstruction iteration. This approach can be expected to yield improved reconstructions since motion estimation and reconstruction are interrelated. Separate motion estimation and restoration, as is commonly done, is sub-optimal as a result of this interdependence.
  - Simultaneous multiframe super-resolution restoration

    In Section 1.4 we discussed the manner in which super-resolution sequence reconstruction was effected using an independent frame by frame restoration approach. Techniques that simultaneously restore multiple frames of the super-resolution sequence can be expected to achieve higher performance since additional spatio-temporal constraints on the super-resolution image ensemble may be included. Though this is a well known technique in restoration theory, it has not seen any application in super-resolution reconstruction.

6 Conclusions

This report examined techniques used for spatial resolution enhancement of low-resolution video sequences. We have presented a comprehensive literature review within a taxonomy of existing techniques. We have discussed the advantages and disadvantages of these methods and have proposed directions for future work which we expect will provide improvements in the state of the art, in terms of both reconstruction quality as well as applicability to "real world" scenes captured using typical consumer quality hardware.

References

[1] J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, CT, 1923.

[2] R. Lenz and U. Lenz, "Calibration of a color CCD camera with 3000x2300 picture elements," in *Close-range photogrammetry meets machine vision*, Zurich, Switzerland, Sept. 1990, vol. 1395 of *Proceedings of the SPIE*, pp. 104–111.

[3] L. Ping and W. Yu-Mei, "Image Resolution Improvement by Swing CCD Array Imager in Two Dimensions and its Signal Processing," in *Visual Communications and Image Processing*, Chicago, IL, Sept. 1994, vol. 2308 of *Proceedings of the SPIE*, pp. 915–920.

[4] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, R. Y. Tsai and T. S. Huang, Eds., vol. 1, pp. 317–339. JAI Press Inc., 1984.

[5] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[6] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*, Addison-Wesley, 1987.

[7] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions of Image Processing*, vol. 5, no. 6, pp. 996–1011, June 1996.

[8] A. M. Tekalp, M. K. Ozkan, and M. I. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," in *Proceedings of the IEEE International Conference on Acousics, Speech and Signal Processing*, San Francisco, CA, 1992, vol. III, pp. 169–172.

[9] A. M. Tekalp, *Digital Video Processing*, chapter 17, Prentice Hall, Upper Saddle River, NJ, 1995.

[10] E. Kaltenbacher and R. C. Hardie, "High-resolution infrared image reconstruction using multiple low resolution aliased frames," in *Proceedings of the IEEE National Aerospace Electronics Conference (NAECON)*, Dayton, OH, May 1996, vol. 2, pp. 702–709.

[11] S. P. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 6, pp. 1013–1027, 1990.

[12] S. P. Kim and W. Su, "Recursive high-resolution reconstruction of blurred multiframe images," in *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, Toronto, Canada, May 1991, vol. 4, pp. 2977–2980.

[13] S. P. Kim and W.-Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Transactions on Image Processing*, vol. 2, pp. 534–539, Oct. 1993.

[14] R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Kluwer Academic Publishers, 1991.

[15] W.-Y. Su and S. P. Kim, "High-resolution restoration of dynamic image sequences," *International Journal of Imaging Systems and Technology*, vol. 5, no. 4, pp. 330–339, 1994.

[16] S. P. Kim and W.-Y. Su, "Subpixel accuracy image registration by spectrum cancellation," in *Proceedings of the IEEE International Conference of Acoustics, Speech and Signal Processing*, Minneapolis, MN, Apr. 1993, vol. V, pp. 153–156.

[17] G. H. Golub and C. F. Van Loan, "An Analysis of the Total Least Squares Problem," *SIAM Journal on Numerical Analysis*, vol. 17, pp. 883–893, 1980.

[18] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 2nd edition, 1989.

[19] N. K. Bose, H. C. Kim, and H. M. Valenzuela, "Recursive Total Least Squares Algorithm for Image Reconstruction from Noisy, Undersampled Multiframes," *Multidimensional Systems and Signal Processing*, vol. 4, no. 3, pp. 253–268, July 1993.

[20] N. K. Bose, H. C. Kim, and H. M. Valenzuela, "Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframes," in *Proceedings of the IEEE conference on Acoustics, Speech and Signal Processing*, Minneapolis, MN, Apr. 1993, vol. 5, pp. 269–272.

[21] C. E. Davila, "Recursive Total Least Squares Algorithms for Adaptive Filtering," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Toronto, Canada, 1991, pp. 1853–1856.

[22] C. E. Davila, "Efficient Recursive Total Least Squares Algorithms for FIR Adaptive Filtering," *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 268–280, Feb. 1994.

[23] N. K. Bose, H. C. Kim, and B. Zhou, "Performance analysis of the TLS algorithm for image reconstruction from a sequence of undersampled noisy and blurred frames," in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, Nov. 1994, vol. III, pp. 571–575.

[24] H. C. Kim, *High Resolution Image Reconstruction from Undersampled Multifreames*, Ph.D. thesis, Pennsylvania State University, May 1994.

[25] A. Papoulis, "Generalized sampling expansion," *IEEE Transactions on Circuits and Systems*, vol. 24, no. 11, pp. 652–654, Nov. 1977.

[26] J. L. Brown, "Multichannel sampling of low-pass signals," *IEEE Transactions on Circuits and Systems*, vol. 28, no. 2, pp. 101–106, ??? 1981.

[27] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using subpixel displacements," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1988, pp. 742–746.

[28] K. Aizawa, T. Komatsu, and T. Saito, "Acquisition of very high resolution images using stereo cameras," in *Visual Communications and Image Processing*, 1991, vol. 1605 of *Proceedings of the SPIE*, pp. 318–328.

[29] J. J. Clark, M. R. Palmer, and P. D. Laurence, "A transformation method for the reconstruction of functions from non-uniformly spaced samples," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 33, no. 4, pp. 1151–1165, Oct. 1985.

[30] L. Landweber, "An iterative formula for Fredholm integral equations of the first kind," *American Journal of Mathematics*, vol. 73, pp. 615–624, 1951.

[31] R. Franke, "Smooth Interpolation of Scattered Data by Local Thin Plate Splines," *Computers and Mathematics with Applications*, vol. 8, no. 4, pp. 273–281, 1982.

[32] K. D. Sauer and J. P. Allebach, "Iterative Reconstruction of Band-Limited Images from Nonuniformly Spaced Samples," *IEEE Transactions on Circuits and Systems*, vol. 34, no. 12, pp. 1497–1506, Dec. 1987.

[33] S. Yeh and H. Stark, "Iterative and One-step Reconstruction from Nonuniform Samples by Convex Projections," *Journal of the Optical Society of America A*, vol. 7, pp. 491–499, 1990.

[34] K. Aizawa, T. Komatsu, and T. Saito, "A scheme for acquiring very high resolution images using multiple cameras," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, Mar. 1992, vol. 3, pp. 289–292.

[35] T. Komatsu, K. Aizawa, and T. Saito, "Very high resolution image acquisition via reconstruction using multiple lower-resolution cameras various in their pixel apertures," in *Proceedings of the International Workshop on HDTV*, 1992, pp. 581–588.

[36] T. Komatsu, K. Aizawa, T. Igarashi, and T. Saito, "Signal-processing based method for acquiring very high resolution images with multiple cameras and its theoretical analysis," *IEE Proceedings I*, vol. 140, pp. 19–25, Feb. 1993.

[37] T. Komatsu, T.Igarashi, K. Aizawa, and T. Saito, "Very high resolution imaging scheme with multiple different aperture cameras," *Signal Processing Image Communication*, vol. 5, pp. 511–526, Dec. 1993.

[38] T. Saito, T. Komatsu, and K. Aizawa, "An image processing algorithm for a super high definition imaging scheme with multiple different-aperture cameras," in *Proceedings of the IEEE International Conference of Acoustics, Speech and Signal Processing*, Adelaide, Australia, Apr. 1994, vol. 5, pp. 45–48.

[39] S. P. Kim and N. K. Bose, "Reconstruct of 2-D bandlimited discrete signals from nonuniform samples," *Proceeding of the IEE*, vol. 137, pp. 197–204, 1990.

[40] R. W. Franke, "Scattered data interpolation: tests of some methods," *Mathematics of Computation*, vol. 38, no. 157, pp. 181–200, Jan. 1982.

[41] M. J. Carlotto and V. T. Tom, "Interpolation of two-dimensional surfaces using gerchberg algorithm," *SPIE*, vol. 359, pp. 226–232, 1982.

[42] Y. Nakazawa, T. Saito, T. Sekimori, and K. Aizawa, "Two approaches for image-processing based high resolution image acquisition," in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, 1994, vol. III, pp. 147–151.

[43] Y. Nakazawa, T. Komatsu, and T. Saito, "High resolution image acquisition based on temporal integration with hierarchical estimation of image warping," in *Proceedings of IEEE International Conference on Image Processing*, Washington, DC, 1995, vol. III, pp. 244–247.

[44] B. R. Frieden and H. G. Aumann, "Image reconstruction from multiple 1-D scans using filtered localized projection," *Applied Optics*, vol. 26, no. 17, pp. 3615–3621, Sept. 1987.

[45] S. Peleg, D. Keren, and L. Schweitzer, "Improving image resolution by using subpixel motion," *Pattern Recognition Letters*, vol. 5, no. 3, pp. 223–226, Mar. 1987.

[46] C. D. Gellat Jr. S. Kirkpatrick and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 13 May 1983.

[47] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Patern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, June 1984.

[48] M. Irani and S. Peleg, "Super Resolution From Image Sequences," in *Proceedings of the 10th International Conference on Pattern Recognition*, Atlantic City, NJ, June 1990, vol. 2, pp. 115–120.

[49] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 3, pp. 231–239, May 1991.

[50] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion and transparency," *Journal of Visual Communications and Image Representation*, vol. 4, pp. 324–335, Dec. 1993.

[51] M. Irani, B. Rousso, and S. Peleg, "Computing occluding and transparent motions," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 5–16, Feb. 1994.

[52] N. R. Shah and A. Zakhor, "Multiframe spatial resolution enhancement of color video," in *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, Sept. 1996, vol. I, pp. 985–988.

[53] N. R. Shah and A. Zakhor, "Resolution Enhancement of Color Video Sequences," *Submitted to IEEE Transactions on Image Processing*.

[54] R. R. Schultz and R. L. Stevenson, "Improved Definition Video Frame Enhancement," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Detroit, MI, May 1995, vol. 4, pp. 2169–2172.

[55] R. R. Schultz and R. L. Stevenson, “Improved definition image expansion,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, Mar. 1992, vol. 3, pp. 173–176.

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

[57] R. R. Schultz and R. L. Stevenson, “Extraction of High-Resolution Frames from Video Sequences,” in *Proceedings of the IEEE Workshop on Nonlinear Image Processing*, Neos Manos, Greece, June 1995.

[58] R. R. Schultz and R. L. Stevenson, “Video resolution enhancement,” in *Image and Video Processing III*, San Jose, CA, Feb. 1995, Proceedings of the SPIE, pp. 23–34.

[59] R. R. Schultz and R. L. Stevenson, “Motion-compensated scan conversion of interlaced video sequences,” in *Image and Video Processing IV*, San Jose, CA, 1996, Proceedings of the SPIE, pp. 107–118.

[60] R. R. Schultz, L. Meng, and R. L. Stevenson, “Subpixel motion estimation for multiframe resolution enhancement,” in *Visual Communications and Image Processing*, J. Biemond and E. J. Delp, Eds., San Jose, CA, Feb. 1997, vol. 3024 of *Proceedings of the SPIE*, pp. 1317–1328.

[61] R. R. Schultz, L. Meng, and R. L. Stevenson, “Subpixel motion estimation for super-resolution image sequence enhancement,” *Journal of Visual Communication and Image Representation, special issue on High-Fidelity Media Processing*, vol. 9, no. 1, pp. 38–50, Mar. 1998.

[62] S. Mann and R. W. Pickard, “Virtual bellows: constructing high quality stills from video,” in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, 1994, pp. 363–367.

[63] B. K. Horn and B. Schunk, “Determining optical flow,” *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[64] R. R. Schultz and R. L. Stevenson, “Bayesian estimation of subpixel-resolution motion fields and high-resolution video stills,” in *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, Oct. 1997, vol. III, pp. 62–65.

[65] R. C. Hardie, T. R. Tuinstra, J. Bognar, K. J. Barnard, and E. Armstrong, “High resolution image reconstruction from digital video with global and non-global scene motion,” in *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, Oct. 1997, vol. I, pp. 153–156.

[66] A. M. Tekalp, *Digital Video Processing*, Prentice Hall, Upper Saddle River, NJ, 1995.

[67] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, “Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images,” *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.

[68] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, "High Resolution Image Reconstruction From a Sequence of Rotated and Translated Frames and its Application to an Infrared Imaging System," *Optical Engineering*, vol. 73, no. 1, pp. 247–260, Jan. 1998.

[69] B.D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 33, no. 4, pp. 1151–1165, Oct. 1985.

[70] A. Lorette, H. Shekarforoush, and J. Zerubia, "Super-resolution with adaptive regularization," in *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, Oct. 1997, vol. I, pp. 169–172.

[71] H. Shekarforoush, M. Berthod, and J. Zerubia, "3D super-resolution using generalized sampling expansion," in *Proceedings of the IEEE International Conference on Image Processing*, Washington, DC, Oct. 1995, vol. II, pp. 300–303.

[72] P. Cheeseman, B. Kanefsky, J. Stutz, and R. Kraft, "Subpixel resolution from multiple images," *Lunar and Planetary Science*, vol. XXV, pp. 241–242, Mar. 1994.

[73] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," Tech. Rep. FIA-94-12, NASA Ames Research Center, Moffet Field, CA 1994, 1994.

[74] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," in *Maximum Entropy and Bayesian Methods*, G. R. Heidbreder, Ed., pp. 293–308. Kluwer, Santa Barbara, CA, 1996.

[75] B. R. Hunt, "Imagery Super-Resolution: Emerging Prospects," in *Applications of Digital Image Processing XIV*, San Diego, CA, July 1991, vol. 1567 of *Proceedings of the SPIE*, pp. 600–608.

[76] B. R. Hunt, "Quantifying the Super-Resolution Capabilities of the CLEAN Image Processing Algorithm," in *Applications of Digital Image Processing XV*, San Diego, CA, July 1992, Proceedings of the SPIE, pp. 300–306.

[77] P. J. Sementilli, M. S. Nadar, and B. R. Hunt, "Poisson MAP super-resolution estimator with smoothness constraint," in *Neural and stochastic methods in image and signal processing II*, San Diego, CA, July 1993, vol. 2032 of *Proceedings of the SPIE*, pp. 2–13.

[78] P. J. Sementilli, M. S. Nadar, and B. R. Hunt, "Empirical Evaluation of a Bound on Image Superresolution Performance," in *Image reconstruction and restoration*, San Diego, CA, July 1994, vol. 2302 of *Proceedings of the SPIE*, pp. 178–187.

[79] P. Yuen and B. R. Hunt, "Constraining information for improvements to MAP image super-resolution," in *Image resonstruction and restoration*, San Diego, CA, July 1994, vol. 2302 of *Proceedings of the SPIE*, pp. 144–155.

[80] B. R. Hunt, "Super-Resolution of Images: Algorithms, Principles, Performance," *International Journal of Imaging Systems and Technology*, vol. 6, no. 4, pp. 297–308, 1995.

[81] B. C. Tom, A. K. Katsaggelos, and N. P. Galatsanos, "Reconstruction of a high resolution image from registration and restoration of low resolution images," in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, 1994, vol. III, pp. 553–557.

[82] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high resolution image from multiple degraded mis-registered low resolution images," in *Visual Communications and Image Processing*, Chicago, IL, Sept. 1994, vol. 2308 of *Proceedings of the SPIE*, pp. 971–981.

[83] B. C. Tom and A. K. Katsaggelos, "Multi-Channel Image Identification and Restoration Using the Expectation Maximization Algorithm," in *Applications of Digital Image Processing XVII*, San Diego, CA, July 1994, vol. 2298 of *Proceedings of the SPIE*, pp. 316–331.

[84] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration and interpolation of low-resolution images," in *Proceedings of the IEEE International Conference on Image Processing*, Washington, DC, 1995, vol. 2, pp. 539–542.

[85] M. Elad and A. Feuer, "Super-Resolution Reconstruction of an Image," in *Proceedings of the 19th IEEE Conference in Israel*, Jerusalem, Israel, Nov. 1996, pp. 391–394.

[86] M. Elad and A. Feuer, "Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1646–1658, Dec. 1997.

[87] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, V. H. Winston & Sons, Washington, D.C., 1977.

[88] D. C. Youla and H. Webb, "Image Restoration by the Method od Convex Projections: Part 1–theory," *IEEE Transactions on Medical Imaging*, vol. MI-1, no. 2, pp. 81–94, Oct. 1982.

[89] P. Oskoui-Fard and H. Stark, "Tomographic image reconstruction using the theory of convex projections," *IEEE Transactions on Medical Imaging*, vol. 7, no. 1, pp. 45–58, Mar. 1988.

[90] M. I. Sezan and H. Stark, "Tomographic Image Reconstruction from Incomplete View Data by Convex Projections and Direct Fourier Inversion," *IEEE Transactions on Medical Imaging*, vol. MI-3, no. 2, pp. 91–98, June 1984.

[91] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *Journal of the Optical Society of America A*, vol. 6, no. 11, pp. 1715–1726, 1989.

[92] G. T. Herman, *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*, Academic Press, New York, 1980.

[93] A. J. Patti, M. I. Sezan, and A. M. Tekalp, "High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur," in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, 1994, vol. I, pp. 343–347.

[94] A. Sawchuk, "Space-variant image motion degradation and restoration," *Proceedings of the IEEE*, vol. 60, no. 7, pp. 854–861, July 1972.

[95] A. J. Patti, M. I. Sezan, and A. M. Tekalp, "High resolution standards conversion of low resolution video," in *Proceedings of the IEEE International Conference of Acoustics, Speech and Signal Processing*, Detroit, MI, 1995, vol. 4, pp. 2197–2200.

[96] A. J. Patti, M. Sezan, and A. M. Tekalp, "Robust methods for high quality stills from interlaced video in the presence of dominant motion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 328–342, Apr. 1997.

[97] A. J. Patti, M. I. Sezan, and A. M. Tekalp, "Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1064–1076, Aug. 1997.

[98] P. E. Eren, M. I. Sezan, and A.M. Tekalp, "Robust, object-based high-resolution image reconstruction from low-resolution video," *IEEE Transactions on Image Processing*, vol. 6, no. 10, pp. 1446–1451, Oct. 1997.

[99] B. C. Tom and A. K. Katsaggelos, "An Iterative Algorithm for Improving the Resolution of Video Sequences," in *Visual Communications and Image Processing*, Orlando, FL, Mar. 1996, vol. 2727 of *Proceedings of the SPIE*, pp. 1430–1438.

[100] B. C. Tom and A. K. Katsaggelos, "Resolution enhancement of video sequences using motion compensation," in *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, Sept. 1996, vol. I, pp. 713–716.

[101] A. K. Katsaggelos, "A Multiple Input Image Restoration Approach," *Journal of Visual Communication and Image Representation*, vol. 1, no. 1, pp. 93–103, Sept. 1990.

[102] G. Jacquemod, C. Odet, and R. Goutte, "Image resolution enhancement using subpixel camera displacement," *Signal Processing*, vol. 26, no. 1, pp. 139–146, 1992.

[103] A. T. Erdem, M. I. Sezan, and M. K. Ozkan, "Motion-compensated multiframe wiener restoration of blurred and noisy image sequences," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, Mar. 1992, vol. 3, pp. 293–296.

[104] C. Srinivas and M. D. Srinath, "A stochastic model-based approach for simultaneous restoration of multiple misregistered images," in *Visual Communications and Image Processing*, Lausanne, Switzerland, Oct. 1990, vol. 1360 of *Proceedings of the SPIE*, pp. 1416–1427.

[105] A. J. Patti, A. M. Tekalp, and M. I. Sezan, "A New Motion Compensated Reduced Order Model Kalman Filter for Space-Varying Restoration of Progressive and Interlaced Video," *IEEE Transactions on Image Processing*, vol. 7, no. 4, pp. 543–554, Apr. 1998.

[106] M. Elad and A. Feuer, "Super-Resolution Restoration of Continuous Image Sequence Using the LMS Algorithm," in *Proceedings of the 18th IEEE Conference in Israel*, Tel-Aviv, Israel, Mar. 1995, vol. ???, pp. ???–???

[107] M. Elad and A. Feuer, "Super-Resolution Restoration of Continuous Image Sequence - Adaptive Filtering Approach," *Submitted to IEEE Transactions on Image Processing in December 1995.*

[108] M.-C. Hong, M. G. Kang, and A. K. Katsaggelos, "A regularized multichannel restoration approach for globally optimal high resolution video sequence," in *Visual Communications and Image Processing*, J. Biemond and E. J. Delp, Eds., San Jose, CA, Feb. 1997, vol. 3024 of *Proceedings of the SPIE*, pp. 1306–1316.

[109] M.-C. Hong, M. G. Kang, and A. K. Katsaggelos, "An iterative weighted regularized algorithm for improving the resolution of video sequences," in *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, Oct. 1997, vol. II, pp. 474–477.

[110] C. Bouman and K. Sauer, "A generalized gaussian image model for edge preserving map estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 3, pp. 296–310, July 1993.