A MULTI-FRAME SUPER-RESOLUTION ALGORITHM

USING POCS AND WAVELET

CHIU-CHIH CHEN

A THESIS

IN

THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF APPLIED SCIENCE

CONCORDIA UNIVERSITY

MONTRÉAL, QUÉBEC, CANADA

SEPTEMBER 2010

ABSTRACT

A Multi-Frame Super-Resolution Algorithm Using POCS and Wavelet

Chiu-Chih Chen

Super-Resolution (SR) is a generic term, referring to a series of digital image processing techniques in which a high resolution (HR) image is reconstructed from a set of low resolution (LR) video frames or images. In other words, a HR image is obtained by integrating several LR frames captured from the same scene within a very short period of time. Constructing a SR image is a process that may require a lot of computational resources. To solve this problem, the SR reconstruction process involves 3 steps, namely image registration, degrading function estimation and image restoration.

In this thesis, the fundamental process steps in SR image reconstruction algorithms are first introduced. Several known SR image reconstruction approaches are then discussed in detail. These SR reconstruction methods include: (1) traditional interpolation, (2) the frequency domain approach, (3) the inverse back-projection (IBP), (4) the conventional projections onto convex sets (POCS) and (5) regularized inverse optimization.

Based on the analysis of some of the existing methods, a Wavelet-based POCS SR image reconstruction method is proposed. The new method is an extension of the conventional POCS method, that performs some convex projection operations in the Wavelet domain. The stochastic Wavelet coefficient refinement technique is used to adjust the Wavelet

sub-image coefficients of the estimated HR image according to the stochastic F-distribution in order to eliminate the noisy or wrongly estimated pixels. The proposed SR method enhances the resulting quality of the reconstructed HR image, while retaining the simplicity of the conventional POCS method as well as increasing the convergence speed of POCS iterations. Simulation results show that the proposed Wavelet-based POCS iterative algorithm has led to some distinct features and performance improvement as compared to some of the SR approaches reviewed in this thesis.

Acknowledgments

I am extremely grateful for the support and encouragement from my MASc. My supervisor, Dr. Wei-Ping Zhu. Without his guidance, expertise, and dedication to my work, this thesis would not have been possible. He provided me with valuable advice, clear and stimulating instruction, and countless hours of his time so that I could see the completion of my thesis and MASc. program.

I would like to express my gratitude to Dr. S. Samadi and Dr. A. Amer. I had the pleasure to study their courses and learned many knowledge from them, which gave me a lot of useful induction towards understanding and able to complete this work.

I am grateful to the Department of Electrical and Computer Engineering staff members. I especially want to thank Miss. Diane Moffat. She was always available to answer my questions and assist me in every way possible.

I wish to thank my student colleagues Guanqun Chen and Majid Behbahani. With whom, over these years, I had the pleasure of doing research with.

I am forever indebted to my parents and in-laws for their emotional and financial support, and for the sacrifices they made so that I could reach my goal. And to my wife, Mina, thank you for your patience, encouragement and love for these past few years.

Contents

List of Figures

xi

xii

List of Tables

List of Symbols

n_1, n_2 : pixel indices of LR image (positive integers).

m_1, m_2 : pixel indices of HR image (positive integers).

μ : image size enlargement factor.

$\mathbf{y}_i$: acquired LR frames ($n_1 \times n_2$) .

i : LR frame counter (positive integers).

n : iteration counter(positive integers).

p : number of available frames

$\mathbf{D}$: image down sampling operator .

$\mathbf{B}$: image blurring or degrading operator .

$\mathbf{M}$: image motion or translation operator .

η : additive noise.

$\mathbf{H}$: complete degrading operator in matrix.

$\mathbf{x}$: ideal HR image ($m_1 \times m_2$).

$\widetilde{\mathbf{x}}$: estimated HR image.

$\phi(t)$: Wavelet scaling function.

$\psi^h(t,s)$: LH 2-D Wavelet function.

$\psi^v(t,s)$: HL 2-D Wavelet function.

$\psi^d(t,s)$: HH 2-D Wavelet function.

$a_{k,l}$: coefficient of LL band Wavelet approximation .

$b^h{}_{k,l}$: wavelet LH band Wavelet coefficient.

$b^v{}_{k,l}$: wavelet HL band Wavelet coefficient.

$b^d{}_{k,l}$: wavelet HH band Wavelet coefficient.

$W_{i,j}$: general term of Wavelet coefficient.

$S()$: a function generates appropriate scaling factor for Wavelet constraint projection.

C_D : data consistency constraint Sets.

C_W : Wavelet constraint sets.

C_A : amplitude constraint sets.

P_D : projection operator of data consistency constraint.

P_W : projection operator of Wavelet constraint.

P_A : projection operator of amplitude constraint.

$\sigma^2{}_{i,j}$: variance of Wavelet coefficient at (i,j).

$G(x)$: CDF of F-distribution.

α : Wavelet coefficient attenuation factor.

ε : convergence criterion

List of Acronyms

CCD : charge-coupled device

CDF : cumulative distribution function

CFT : continuous Fourier transform

CLS : constrained least squares

DFT : discreet Fourier transform

GCV : generalized cross-validation

HR : high resolution

IBP : inverse back-projection

LR : low resolution

LSI : linear spatial invariant

MAP : maximum a posteriori

ML : maximum likelihood

MRA : multi-resolution analysis

MRF : Markov random fields

PCG : preconditioning for conjugate gradient

POCS : projections onto convex sets

PSNR : peak signal to noise ratio

PSF : point spread function

SR : super-resolution

SSD : sum of squared differences

3

Chapter 1

Introduction

1.1 General

Super-resolution (SR) refers to a set of image enhancement algorithms that reconstruct high resolution (HR), fine detailed, high quality images from a set of degraded, low resolution (LR), coarse quality images. In most electronic visual applications, clear HR images are often required. An HR image is one that contains high density pixels within a fixed area and is able to reveal more details for other visual applications. Thus, the degree of image detail is important for many other image processes. For instance, in medical image applications such as tomography images, it is very essential to have high resolution in order to help doctors make correct decision for patients. Also, for object recognition, detection and identification applications, the image needs to have clear sharp contour to distinguish objects within an image or a video frame. When the visual applications are for military or astronomical purposes, the requirement for image enhancement algorithms is even more crucial. In many situations, there are some reasons that higher resolution images can not be achieved, such as limitations of camera sensor or the application budget constraints. However, those problems can be resolved by applying super-resolution algorithms, which are inexpensive and can be used even with antiquated equipment.

1

Super-resolution algorithms make higher spatial resolution by fusing useful information from the temporal domain data. In other words, super-resolution algorithms are the processes that convert temporal bandwidth into spatial resolution. Consecutive frames from the same scene provide the required elements to achieve spatial resolution enhancement. However, it is not guaranteed that all sequences of frames taken from the same scene will provide the information required to rebuild the high resolution scene. If the camera is fixed on a platform and there is not even the slightest object motion, the new information obtained from this kind of situation is useless. Therefore, at least slight motion is needed in order to obtain non-redundant information among consecutive frames. In other words, each LR frame must provide a different 'view' of the scene. There are two ways to obtain non-redundant information from the same scene; these are the multi-channel and multi-frame methods. In the case of multi-channel method for example, we can have multiple cameras take images by aiming all cameras at the same direction from different positions. The non-redundant information is then acquired as multi-channel method. The multi-frame method obtains useful sub-pixels from the same scene by creating frames from the relative motion of the camera at different time intervals. Examples of this latter method can be seen in the scanning of some security surveillance cameras as they sweep a still scene or in satellite cameras taking picture from the earth landscape while orbiting the earth. These two cases are examples of multi-frame data captured from a single camera.

Super-resolution is not a straightforward signal process. It combines several image and video signal processing techniques, such as image registration, degrading function estimation, image interpolation and image restoration. Each of these techniques is highly complex and the basics are briefly discussed in this thesis. Obtaining a SR image is a computationally intensive process. In order to manipulate hundreds of thousands of pixel data efficiently, each step employs outstanding algorithm among the various methods. Thus, in the future, with optimal high speed SR algorithm, an LR camera may be capturing a

real scene remotely and viewers could receive its HR version almost simultaneously. In practice, only few assumptions can be made prior to SR reconstruction process. Thus, the total amount of process data must be limited to relevant information. To this end, instead of being slowed by the challenges of SR algorithm such as insufficient information of sensor noise, motion modeling and camera characteristics, an SR algorithm should rely solely on the observed frame data and effectively utilize computational resources.

This thesis proposes a rapid and accurate super-resolution algorithm by imposing wavelet domain constrain using a wavelet sub-image coefficients refinement technique into the conventional POCS SR method. Before going into details of the proposed method, fundamental steps such as image registration, degrading function estimation, image restoration and image de-blurring are firstly introduced. Moreover, some well-known SR reconstruction approaches such as traditional interpolation, frequency domain approach, inverse back-projection (IBP), conventional projections onto convex sets (POCS) are examined for comparison.

1.2   The Super-Resolution Problem

Before formulating the super-resolution problem, some assumptions are made to allow the SR reconstruction algorithms to focus on the most important calculations and reduce the computational complexity. It is assumed that all super-resolution problems are under conditions that objects in LR frames are taken from a conventional orthographic projection camera, with constant luminous condition, negligible optical distortions, and only affine translations (only shift and rotate) between acquired frames. Although these assumptions cause some negligible uncertainties, they meet almost all requirements of super-resolution applications.

In this thesis, the super-resolution problem is defined as:

Reconstruct a high quality HR image $\mathbf{x} = \mu n_1 \times \mu n_2$, from a set of LR degraded frames

$\mathbf{y}_i(i = 1, \ldots, p)$ with $n_1 \times n_2$ pixels in dimension, where μ is the desired enlargement factor.
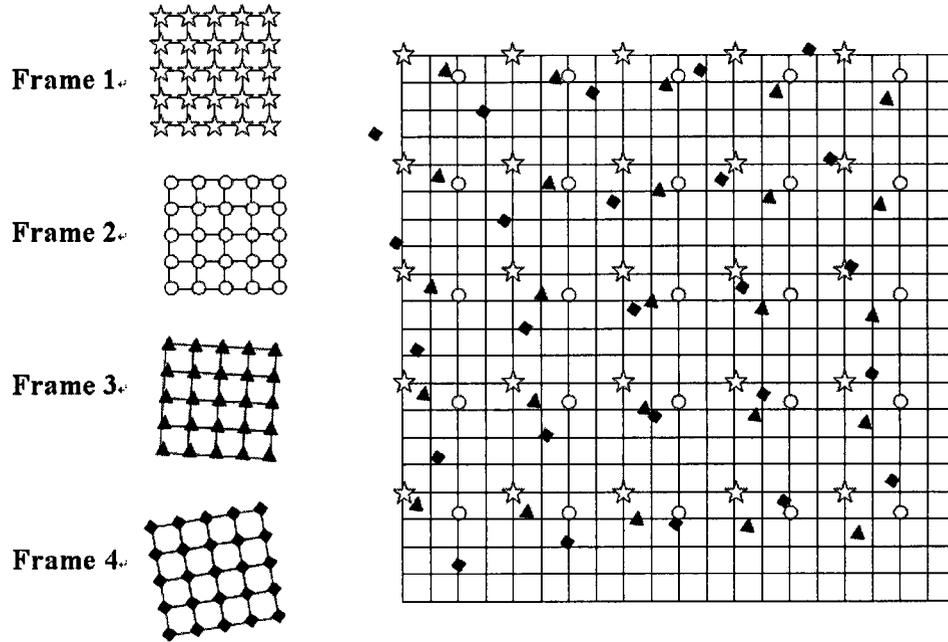
Figure 1.1: LR frame pixels projected on high-resolution grids

In Figure 1.1, a typical super-resolution problem model is described. There are four 5×5 pixels LR frames projected onto a 20×20 HR grid. The star, the circle, the filled triangle and the filled diamond symbols represent pixels in LR frames 1 to 4. In this figure, the first frame (star) is set as the reference frame. Pixels in frame 2 (circle) shift to right-down direction from frame 1. Frame 3, and frame 4 motion involves not only simple translation but also rotation with respect to the reference frame. In some applications, magnification motion such as camera zoom in and zoom out on input LR frame are also acceptable in most super-resolution model. The purpose of a super-resolution algorithm is to fill the unknown pixels value on the HR grids.

Inspired by many conventional image reconstruction problems [1] [2] [3] [4], most super-resolution problems are formulated as forward relationship. The forward relationship

4

means the image capturing process of a original scene in which the image information is degraded while forward to each step. Figure 1.2 depicts a situation of forward relationship between a real object, LR frames and HR images and explains how an ideal HR image of a real world object captured by an electronic camera is degraded.
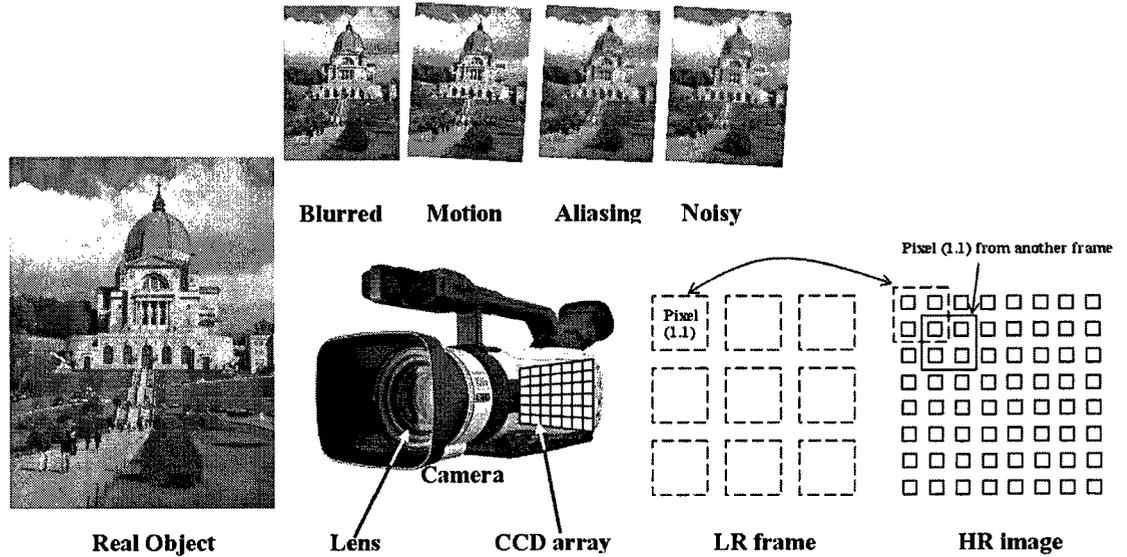
Figure 1.2: Forward relationship of an image capturing system

Figure 1.2 demonstrates a system where a real object is captured by the charge-coupled device (CCD) camera. The continuous photon energy first passes through the camera lens while the camera is shaking or moving, this creates a motion blur. The blurred version of a real object image hits the CCD array and the analog signal is transformed into a two dimensional discrete image. The density of the CCD array determines the aliasing parameter of the image capturing system and each photon charge-coupled device quantizing the light intensity then determines the quantization error. Moreover, image signals traveling in electronic camera are usually contaminated by additive noise. In other words, image degrading begins with blurring, followed by faulty translation, aliasing (or down-sampling), quantization and additive noise. The noisy image on the up-right hand side of Figure 1.2, the LR frame is considered as a blurred, translated, down-sampled and noisy version of

5

the HR image. The forward relationship can be described by the following mathematical formula:

$$\mathbf{y}_i = \mathbf{DM}_i\mathbf{B}_i\mathbf{x} + \eta_i \qquad 1 \leq i \leq p \tag{1.1}$$

where $\mathbf{D}$ is the down-sampling matrix, $\mathbf{M}_i$'s are the inter-frame motion (affine translation) matrices; this matrix maps the HR image grid to the individual LR frame grid, $\mathbf{B}_i$'s are the blurring and degradation matrix. $\mathbf{x}$ is the unknown ideal HR image and η_i's are the additive noise vectors, each vector element corresponding to noise on each LR frame. $\mathbf{y}_i$ are the given LR frames captured from the same scene in the vector form. In Equation (1.1) down-sampling operator $\mathbf{D}$ is known, since the magnification factor μ can be decided in SR reconstruction process. Motion operators $\mathbf{M}_i$'s are unknown. Then, the blurring and degradation operators $\mathbf{B}_i$'s in most SR applications are unknown but the characteristic of blurring operator is assumed to be linear spatial invariant (LSI) in each individual LR frame. The additive noise η_i, the last term in Equation (1.1), is assumed as Gaussian normal distribution due to the central limit theorem with unknown variance.

Consider a general SR reconstruction case with p available LR frames, where each LR frame $\mathbf{y}_i$ contains $n_1 \times n_2$ pixels. Applying these into the computer program, when examining the dimensions of each matrix symbol of Equation (1.1), each given LR frame $\mathbf{y}_i$ is vectorized into an $n_1 n_2 \times 1$ column vector. The desired HR image $\mathbf{x}$ is also written as a $\mu^2 n_1 n_2 \times 1$ column vector. The motion operator of each LR frame $\mathbf{M}_i$'s is a $(\mu^2 n_1 n_2)^2$ square matrix which is used to convert an ideal HR image $\mathbf{x}$ into a shifted and rotated version of itself. The blurring operator to each LR frame $\mathbf{B}_i$'s is also a $(\mu^2 n_1 n_2)^2$ square matrix, which degrades ideal HR image $\mathbf{x}$ into its blurred version, and the down-sampling matrix $\mathbf{D}$ is a $n_1 n_2 \times \mu^2 n_1 n_2$ matrix which decimates the ideal HR image into coarse LR frame. These operations are usually integrated into one complete system matrix $\mathbf{H}_i$ by multiplying all operator matrices in the forward process Equation (1.1). It is equivalent to

6

$$\mathbf{y}_i = \mathbf{H}_i\mathbf{x} + \eta_i \qquad\qquad (1.2)$$

where $\mathbf{H}_i = DM_iB_i$

In Equation (1.2), the size of the LR frame $\mathbf{y}$ and the additive noise η become $pn_1n_2 \times 1$ because the information of all the available frames are combined. The size of the complete system matrix $\mathbf{H}_i$ is described as $pn_1n_2 \times \mu^2 n_1n_2$. Note that the dimension of matrix $\mathbf{H}_i$ relies on the number of available frames, acquired LR frame size and down-sampling factor, which are typically very large. Therefore, the system of Equation (1.2) is difficult to solve or even to calculate with a powerful computer. Due to the high complexity of the SR reconstruction problem, many different approaches have been previously proposed and most of them were used in attempt to avoid solving the SR reconstruction problems directly from Equation (1.1) or Equation (1.2). Previous works and the evolution of SR reconstruction solutions are discussed in the following section.

1.3   Previous Work

Constructing a super-resolution image from multiple LR frames is a relatively new research topic in image processing. Historically, super-resolution techniques is evolved from image enlargement by simple interpolation methods to enhance high frequency components. The earlier approaches of super-resolution have mostly been focused on the frequency spectrum of a single acquired LR frame [5]. By constructing new frequency data points outside of the known frequency data points, a process known as "extrapolation", the high frequency information of the desired image can be built. In the 1960-70's, different methods were suggested, such as by Slepian et al. [6], they suggested obtaining high frequency information by extrapolating frequency information with a wave function in Prolate

7

spheroidal coordinates. Also, Papoulis [7] introduced an extrapolation algorithm by reducing error energy from LR frame spectrum. However, the results of extrapolation are often less meaningful, and are subject to greater uncertainty.

A multiple frame super-resolution approach was first proposed by Tsai and Huang [8]. In their proposed approach, LR frames are treated as down-sampled, shifted versions of the continuous ideal HR image, and all the acquired LR frames are converted into a frequency domain. They also discussed some Fourier transform properties such as shift property and the aliasing relation between Continuous Fourier Transform (CFT) of ideal HR image and Discreet Fourier Transform (DFT) of the observed LR frames to achieve super-resolution image reconstruction. This approach is used as a comparison method and more details will be described in Section 3.2.

An iterative approach was proposed by Irani and Peleg [9] [10] [1] called "inverse back-projection" (IBP), where the HR image is estimated by taking back projection of the simulated LR frames and the acquired LR frames. In this approach, the iteration is run until the difference between simulated LR frames and the acquired LR frames is less than a certain acceptable error. More details about this approach discussing how to generate the simulated LR frames and how a projection is performed will be provided in Section 3.3.

Another iterative approach, originally evolved from an incomplete data recovery technique known as "Convex Projections" derived by Youla and Webb [11] was used to recover missing image information. Sezan and Stark [12] further applied the convex projection technique to restore a single tomography image. Then, Stark and Oskoui [13] first introduced projections onto convex sets (POCS) which is an algebraic data recovery technique for multi-frame SR image reconstruction. In their work, the unknown ideal HR image is subject to several constraints, such as image signal being limited in both the frequency band and the spatial duration. Each constraint is considered as a convex set. By performing successive projection operations onto convex sets, the result of iterative projection can

converge to a data set that satisfies each constraint and gives the desired HR image. The origination and mathematical explanation of the POCS approach is detailed in Section 3.4.

Some researchers have attempted to solve the SR reconstruction problems by using linear algebra optimization methods. Elad and Feuer [14] as well as Nguyen et al. [15] presented a set of comprehensive numerical algorithms for SR reconstruction from motion estimation step to de-blurring step. Farsiu et al. [16] inherited the algorithm in [14], [15] and proposed a fast and robust super-resolution approach which combines previous optimization techniques and his own "Shift and Add" technique and produces really good quality HR images and requires less computational time than other iterative SR approaches.

These techniques and many other SR image reconstruction approaches that are beyond the scope of this thesis have been reviewed in detail in [17], [18] and [19].

1.4   Research Motivation

In the field of image or video signal processing, there are two methods to increase the image resolution. One involves increasing the elements of optical sensor of cameras, while the other involves image interpolation by computer software. The conventional image interpolation method by computer is utilized the most. Various image interpolation methods such as "near-neighbor interpolation", "bilinear interpolation", "bicubic interpolation" and "spline interpolation" are well known by many researchers. These traditional image interpolation methods simply increase the size of the images without considering the characteristics of the images. Normally, simple interpolation ignores the variance of the high and low frequency components of the image. As a result, the image becomes more blurry as the image increases in size. Therefore, the quality of the simple interpolated image is limited. In contrast, super-resolution reconstruction extracts the useful pixels from consecutive LR frames, taking the advantages of the relation of each frame and employs motion estimation algorithm to calculate spatial differences between each frame in order to reconstruct a HR

9

image.

Normally, the high frequency components of an image are the edges. The human eye system has higher sensitivity to the high frequency components when viewing an image. Hence, preserving the edges is very important in order to minimize human eye visual effects when increasing the image size. Eliminating saw-tooth effect on the edges, blurring and smoothing the image after increasing the resolution are the key points of developing a SR algorithm.

After examining some of the existing super-resolution reconstruction methods, we can conclude that the conventional POCS method is in general a good SR method, but it has some rooms to be improved. Especially on the subjective visual quality, the result of the conventional POCS method still has some noisy features such as grid-ish noise present on the final estimated HR image. These particular types of noise may not be overcame by the traditional noise reduction methods. Therefore, transforming the LR frame data into some other data space could be a solution to improve the conventional POCS method.

To date, very limited research in applying wavelet technique to super-resolution has been published and also wavelet has many beneficial features that are able to reduce the unwanted components from the image data. Considering the limited amount of research in this regard and the potential advantages it brings to super-resolution image reconstruction, the use of wavelet technique in the super-resolution algorithm has became the central subject of this thesis.

1.5   Scope and Organization of the Thesis

The main scope of this thesis is to present a wavelet-based POCS SR reconstruction algorithm. The proposed algorithm manipulates the LR frames data in the wavelet domain to obtain the HR image. The introduction chapter describes the requirements and applications of super-resolution techniques and the model of the super-resolution problem is

formulated as a matrix system in Equation (1.1). Subsequent chapters will describe the super-resolution process step by step. As introduced in Chapter 1, the super-resolution processes can be broken down into four major steps ,namely (1)image registration (2)degrading function estimation (3) SR reconstruction and (4) image de-blurring. The first, second and the last steps are presented in most of the previously proposed approaches and will be discussed in Chapter 2. The third step which may vary with each of the different proposed approaches, and will be discussed in Chapter 3. The main contributions of this thesis will be discussed in more detail in Chapter 4. The comparison and experimental results of the proposed super-resolution approach and other previously proposed approaches will be summarized in Chapter 5. The Chapter 6 concludes this thesis and provides some suggestions for further research. The organization of this thesis and the details are summarized briefly as follows.

In Chapter 2, a general review of the fundamental process steps of super-resolution image reconstruction is provided. It begins from the image registration process which includes the motion estimation and rotation estimation processes prior to aligning the consecutive frames with respect to the reference frame for SR process. The implemented image registration process algorithm will be discussed in detail in the first section. The following step after image registration is the degrading function estimation which estimates the blurring operator B_i presented in Equation (1.1). The SR reconstruction step will be briefly introduced in the third section because this SR reconstruction step varies greatly with each different approach, which will be particularly discussed in Chapter 3. The last section describes the method of the de-blurring process which is the last step in the SR reconstruction process in most of the proposed approaches.

In Chapter 3, five different super-resolution reconstruction approaches are described; four of them are also being used for simulation and comparison in this thesis. These different SR approaches are presented in chronological order according to their publishing

date. Section 3.1 presents three types of traditional interpolation methods. Although interpolation methods are not considered as SR approaches, they are, however, preliminary methods of SR. The nearest-neighbor, bilinear and bicubic interpolation are introduced in this section. Section 3.2 will present the classic and also the first SR reconstruction approach proposed by R. Y. Tsai and T. S. Huang in 1984 [8] which considers the reconstruction step of SR as a de-aliasing process of LR frames in frequency domain. Section 3.3 talks about an iterative SR reconstruction approach, namely "inverse back-projection" (IBP) which was introduced in [9]. The next approach is the conventional "projections onto convex sets" (POCS) approach, which is a well-known approach, but that has been rarely discussed in literatures. Thus, the conventional POCS approach will be reviewed in detail in Section 3.4, particularly. The fourth type of SR reconstruction approach introduced in the last section, is a group of regularized optimization approaches, which treat the SR reconstruction step as an inverse optimization problem of recovering missing data from the observed video frames. Three commonly used regularized SR optimization methods will be also discussed.

Chapter 4 presents the proposed wavelet-based POCS SR reconstruction approach. Section 4.1 introduces a framework of the wavelet-based POCS SR reconstruction process. Section 4.2 gives a basic presentation of the wavelet analysis of image data and the derivation of the wavelet domain convex set for wavelet-based POCS process. The wavelet coefficient refinement process is detailed in Section 4.3, which consists of a number of sub-processes as described below.

- The method for calculating the variance of wavelet coefficients for stochastic test

- Origination of stochastic F-Distribution

- Detailed information about the wavelet coefficient refinement method

Section 4.4 summarizes the proposed wavelet-based POCS SR reconstruction procedure.

The process flowchart is shown and explained step by step.

In Chapter 5, the proposed SR method is simulated and compared with some of the known methods. This chapter begins with the defined simulation environment, the process of the generating the degraded LR frames and the performance metrics for comparison. Then, the second section presents some intermediate results from the proposed algorithm. The last two sections present the subjective and objective simulation results, respectively. Comprehensive comparisons of both visual and numerical results are shown.

In Chapter 6, the first section summarizes the work of the thesis. Then, some possible directions for further research are suggested in the second section.

Chapter 2

The Fundamentals of Super-Resolution

The super-resolution techniques could be categorized under the image restoration class. However, many image or video signal processing books rarely describe super-resolution as a single technique, because it actually involves several sub-techniques such as image registration, degrading function estimation, SR image restoration and image de-blurring. This chapter details these essential preliminary steps in the super-resolution technique. Section 2.1 will describe the image registration step of the SR reconstruction process. In this step, the motion estimation and rotation estimation algorithms used in this thesis will be discussed. In Section 2.2, the algorithm of degrading function estimation which estimates the kernel of the burring operator, is explained. In Section 2.3, the restoration step of the SR reconstruction process is introduced. Since, this step varies from different approaches, it will then be further discussed in detail in Chapter 3. The image de-blurring step of the SR reconstruction process is introduced in Section 2.4. For a better understanding of these preliminary steps, Figure 2.1 shows the general steps of super-resolution reconstruction approach, namely (1) image registration, (2) degrading function estimation, (3) SR reconstruction and (4) image de-blurring.
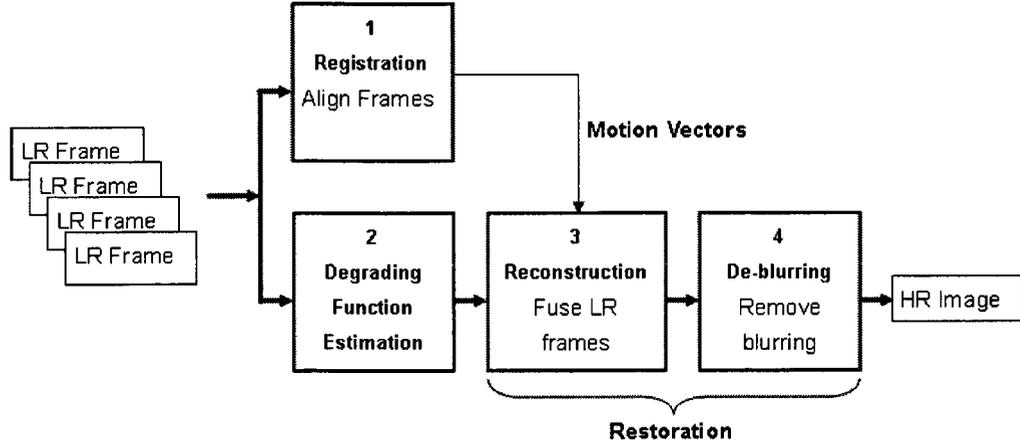
Figure 2.1: General steps of super-resolution reconstruction process

2.1   Image Registration

Image registration is a process that aligns two or more images of the same scene. It is very important to have motion estimation as accurate as possible to achieve image registration. Usually the image registration process is applied as a first step in super-resolution processing. Relative rotation angles of all the LR frames with respect to the reference LR frame are firstly estimated before the affine motion is estimated. A reasonable assumption is made that most acquired LR frames for SR image reconstruction problems are taken within very small time interval. Therefore a practical motion estimation algorithm considers only global motion between consecutive frames. According to the brief introduction of the image registration conditions mentioned above for SR image reconstruction process, the main goal of this section is to calculate the relative rotation angle $\Delta\theta$, relative x-direction motion vector Δx, and relative y-direction motion vector Δy.

2.1.1   Rotation Estimation

The rotation estimation algorithm used in this thesis is a method originated from an energy correlation. Firstly, vectors through the center of LR frame vary by different angles

are formed and in order to calculate the average energy from every component in each vector. Vectors containing average energy at each angle of reference frame and rotated frame are calculated first. The relative rotation angle between two LR frames can be found by seeking the maximum correlation of the average energy vectors. The seeking procedure is that (1) set an average energy vector on the reference frame, (2) set an average energy vector on the rotated frame, (3) matching these two vectors for a maximum correlation while changing the angle of the vector on the rotated frame. A graphical illustration of the relative rotation angle estimation method is shown in Figure 2.2.
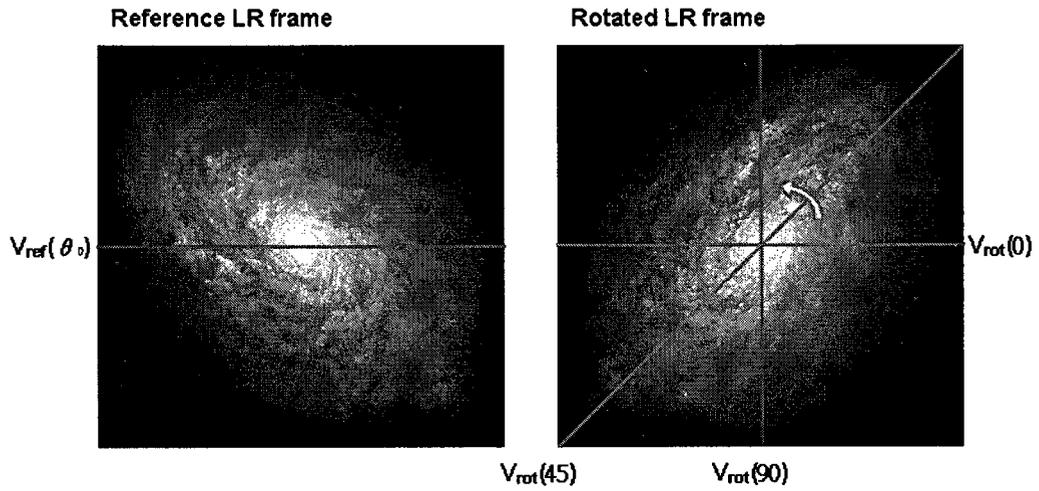
Figure 2.2: Relative rotation estimation of two images

$$r_{max} = \text{max}(\text{Correlation}(V_{ref}(\theta_0), V_{rot}(\theta_i)))$$

$$\Delta\theta = \theta_0 - \theta_i \tag{2.1}$$

On the left hand side of Figure 2.2, the maximum correlation seeking task begins from the vector containing average energy at 0 degrees on the rotated frame. By increasing the scanning angle by 0.1 degrees or 0.5 degrees, depending on the precision requirement of

the SR application, the maximum correlation of the average energy vectors between V_{ref} and V_{rot} can be found. In this example, the maximum correlation happens when $\Delta\theta = 90\,^{\circ}$. Sometimes multiple vectors are compared across the reference and rotated frames, in order to guarantee the accuracy of the estimated angle.

2.1.2 Motion Estimation

The motion estimation algorithm implemented in this work is a hierarchical model-based motion estimation technique [20] [21]. The framework of the hierarchical motion estimation and its pyramid structure are shown in Figure 2.3 and Figure 2.4, respectively.
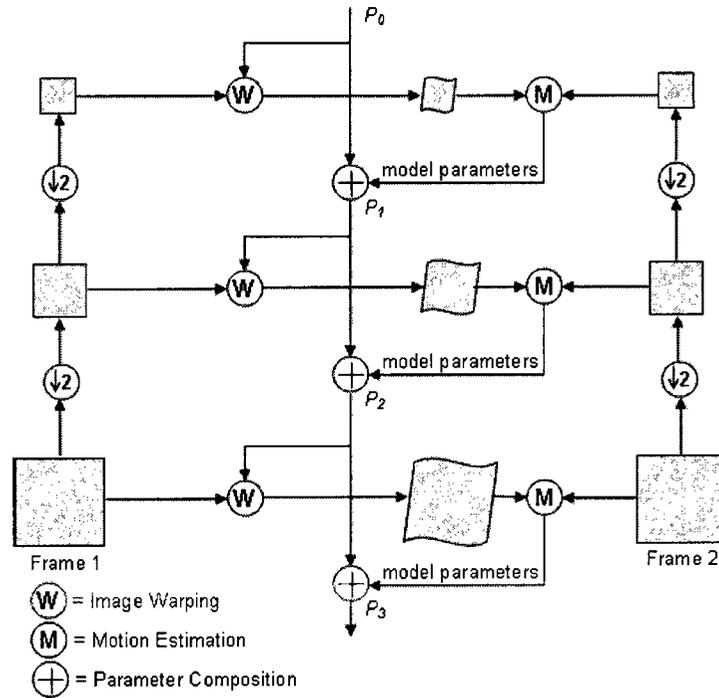
Figure 2.3: Framework of the hierarchical motion estimation

17

There are four basic elements in this framework. (1) pyramid decomposition, (2) motion estimation, (3) image warping and (4) coarse-to-fine refinement. This motion estimation algorithm is capable to estimate global motion and multiple local motions on the consecutive frames by choosing different motion models such as "affine flow model", "planar surface flow model" [22], "rigid body model" and "general flow field model". However, in the super-resolution process, global motion estimation is sufficient for most cases. In other words the "affine flow model" will be employed for our motion estimation process.
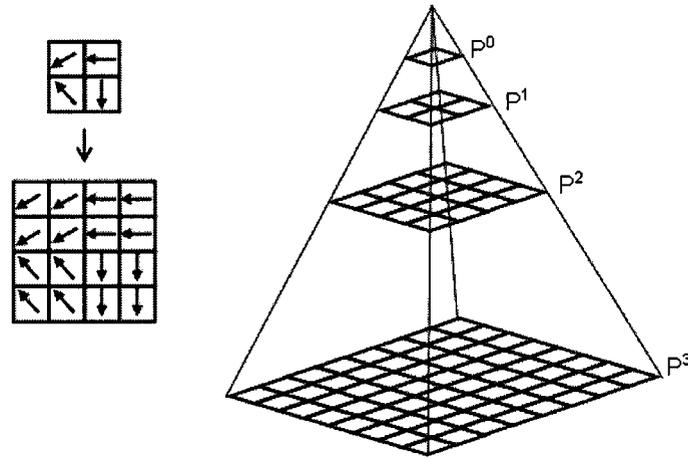
Figure 2.4: Pyramid structure of the hierarchical motion estimation

Considering the procedure of this motion estimation algorithm, the reference frame and the frame to be estimated are first down-sized by the Laplacian pyramid method [23]. The estimation process starts from the most coarse version of the two frames (see node P_0 in Figure 2.3). Frame 1 in the left hand side, set as a reference frame which is firstly pseudo-warped. The reason of not preforming the first warping is that in this stage there is no motion parameters provided from the previous level. Secondly, on the nodes with symbol M, the motion estimation method is performed according to the "affine flow model", and in these nodes the sum of squared differences (SSD) minimization is involved. The estimated motion parameters are passed to the finer level for the motion parameter composition and the warping process. As the frame resolution increases, the coarse-to-fine refinement

18

method is implemented by using the Gauss-Newton minimization technique. The estimated final motion parameters are provided at the node P_3 in Figure 2.3.

The mathematical deviation of the second element, motion estimation, usually starts from the generic image intensity function $I_g(x, y, t)$ in the space and time domain. The conservation of intensity of video frames can be expressed as follows:

$$I_g(x, y, t) = f(x + u(x, y)\triangle t, y + v(x, y)\triangle t, t + \triangle t) \tag{2.2}$$

where $u(x, y)$ and $u(x, y)$ are the frame velocity which are defined as

$$u(x, y) = p_1 + p_2 x + p_3 y$$
$$v(x, y) = p_4 + p_5 x + p_6 y \tag{2.3}$$

To describe the affine flow field with respect to frame origin, the Equation (2.3) can be rewritten in vector form as:

$$\mathbf{f}(x, y) = \mathbf{A}\mathbf{p} \tag{2.4}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{bmatrix}$$
$$\mathbf{p} = (p_1, p_2, p_3, p_4, p_5, p_6)^T \tag{2.5}$$

The motion in complete region is specified by the parameter $\mathbf{p}$, which is unknown and needs to be estimated.

To solve Equation (2.4), the Laplacian pyramid image intensity is applied, Thus,

$$I(x, y, t) = I(x - u(x, y), y - v(x, y), t - 1)$$

19

where I represents Laplacian pyramid image intensity. Therefore, the flow field in a specific region is estimated by SSD error minimization.

$$E(\{\mathbf{f}\}) = \sum_{x,y} ((I(x,y,t) - I(x - u(x,y), y - v(x,y), t - 1))^2 \qquad (2.6)$$

where $\{\mathbf{f}\} = \{u(x,y), v(x,y)\}$ denotes the entire flow field within a region specified by the overall points (x,y). Practically, in Equation (2.6) the sum of individual error is not a quadratic function in view of the unknown entire flow field within the region of interest $\{\mathbf{f}\}$, and it is a non-linear minimization problem.

This kind of problem in this motion estimation algorithm can be solved by applying the Gauss-Newton's method. If $\{\mathbf{f}\}_i$ denotes the flow field estimated at the ith iteration, the incremental estimate $\{\delta\mathbf{f}\}$ can be obtained by minimizing the quadratic error function as given below,

$$E(\{\delta\mathbf{f}\}) = \sum_{x,y} (\Delta I + \nabla I \cdot \delta\mathbf{f})^2 \qquad (2.7)$$

where

$$\Delta I(x,y,t) = I(x,y,t) - I(x - u(x,y)_i, y - v(x,y)_i, t - 1)$$

considering the flow field estimated from each iteration, which is different between the two frames at corresponding pixels. The minimization problem described in Equation (2.7) can be reformulated in terms of the unknown incremental affine flow model parameters. Let $\mathbf{p}_i$ denote the current estimated affine flow parameters, then the incremental parameters $\delta\mathbf{p}$ of the affine flow model can be achieved by substituting $\delta\mathbf{f}$ in Equation (2.4) by Equation (2.7), and the error measured is a function of $\delta\mathbf{p}$.

$$E(\{\delta\mathbf{p}\}) = \sum_{x,y} (\Delta I + (\nabla I)^T \mathbf{A}\delta\mathbf{p})^2 \qquad (2.8)$$

20

Minimizing Equation (2.8) with respect to $\delta\mathbf{p}$, we obtain

$$[\sum \mathbf{A}^T (\nabla I)(\nabla I)^T \mathbf{A}]\delta\mathbf{p} = -\sum \mathbf{A}^T (\nabla I)(\Delta I) \qquad (2.9)$$

The motion parameter $\mathbf{p}$ is obtained from Equation (2.9). Applying the estimated motion parameter $\mathbf{p}$ back to Equation (2.4), the displacement of frame 2 on both horizontal and vertical direction can be obtained.

The third element, frame warping, is done by computing a flow field according to the affine flow model parameters and then based on flow field to warp $I(t-1)$ toward the reference frame $I(t)$. The warping algorithm imposes bilinear interpolation to estimate some missing pixels. The computation of error ΔI is the warped frame. The intensity gradient ∇I is computed from the reference frame.

The last element, coarse-to-fine refinement, generates the estimated motion from one level to the next, which are then used as an initial frame. For the global motion model, parameters are directly forwarded to the next level to perform the initial warping.

2.2   Degrading Function Estimation

The purpose of estimating the degrading function of acquired frames is to provide blurring operation information for the SR image reconstruction step. The problem of restoring a single degraded image to its original version is similar to the SR reconstruction problem to restore multiple degraded LR frames. The main difference is that restoring a single degraded image to its original version is at the same resolution scale, and reconstructing an SR image from several degraded LR frames is at a higher resolution scale. A single image degradation model can be mathematically described as follows:

$$\mathbf{g} = \mathbf{Bx} + \eta \qquad (2.10)$$

In Equation (2.10), B is the degrading or blurring matrix which blurs the original image x into a blurring state and g is an acquired degraded noisy image. In most cases, blurring is considered as a linear spatial invariant system. The original image x can be recovered by taking the "deconvolution" operation of B from the acquired degraded noisy image g. Comparing Equation (2.10) and Equation (1.2), it is known that super-resolution is a special case of image restoration. According to the super-resolution framework, Equation (2.10) can be rewritten as

$$\mathbf{y}_i = \mathbf{DM}_i\mathbf{Bx} + \eta_i \qquad 1 \le i \le \mu^2 \tag{2.11}$$

where μ is an arbitrary scaling factor and the LR frames $\mathbf{y}_i$ are obtained from the original image x by shifting from frame to frame motion operator $\mathbf{M}_i$, degraded by the blurring operator B and down-sampled by a factor μ.

In most degrading function estimation problems, the blurring operator B is unknown or partially known within a set of parameters. Problems of the estimating degrading function from observed image with incomplete degraded information are called "blind blur identification" in some image processing literatures. Similarly, blind blur identification can be used broadly to handle multiple observed frames and applied on super-resolution problems to estimate the degrading function on each observed LR frame. To judge the quality of the degrading function estimator, a good SR blind blur identification method should posses characteristics such as the ability to suppress the noise, to minimize the edge artifacts and to prevent degrading function from converging to the trivial delta function.

The method of estimating the degrading function used in this thesis uses N. Nguyen's [2] "parametric blurring model blind blur identification". The concept of this method is to formulate the blurring operator B as a function of point spread function(PSF) parameter set σ and reformulate Equation (2.11) to be

$$\begin{aligned} \mathbf{y}_i &= \mathbf{DB}(\sigma)\mathbf{M}_i\mathbf{x} + \eta_i \qquad 1 \le i \le p \\ &= \mathbf{H}_i(\sigma)\mathbf{x} + \eta_i \end{aligned} \qquad (2.12)$$

where $\mathbf{B}$ is the blurring operator and it is now a function of SPF parameter set σ. Using the least squares minimization method to solve Equation (2.12), we get

$$\mathrm{argmin}_{\mathbf{x}} \sum_{i=1}^{p} \left\| \mathbf{y}_i - \mathbf{H}(\sigma)\mathbf{x} \right\|_2^2 + \lambda \mathbf{x}^T \mathbf{Q}\mathbf{x} \qquad (2.13)$$

where λ is the regularization parameter controling the smoothness and the stabilization of the solution. The matrix $\mathbf{Q}$ is usually a symmetric positive definite matrix, and here $\mathbf{Q}$ is set to be the identity matrix in order to stay for the generality. The minimization of Equation (2.13) can be expressed in the following format:

$$\mathbf{x}(\sigma, \lambda) = (\mathbf{H}(\sigma)^T \mathbf{H}(\sigma) + \lambda \mathbf{I})^{-1} \mathbf{H}(\sigma)^T \mathbf{y} \qquad (2.14)$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \mathbf{H} = \begin{bmatrix} H_1 \\ \vdots \\ H_p \end{bmatrix} \qquad (2.15)$$

The goal of the degrading function estimation here is to find the PSF parameters σ and the regularization parameter λ. There are some available algorithms such as "preconditioning for conjugate gradient" (PCG) and "generalized cross-validation" (GCV), are used to solve the minimization problem like Equation (2.14). The detailed procedure of solving Equation (2.14) is beyond the scope of this section. The algorithms of PCG and GCV have been discussed in detail in [2].

23

2.3 SR Image Reconstruction

The SR Image Reconstruction step is the core of the whole super-resolution process. Researchers try to solve the SR reconstruction problem from different directions, therefore the algorithms vary with different approaches. In this step, the LR frames are joined together to contribute certain amount of useful information to build the HR image. In other words, sub-pixels from each LR frame are fused to provide corresponding HR pixels. Since there are various SR image reconstruction approaches, each SR image reconstruction approach will not be specifically discussed in this section. This section is going to provide the classification of recent existing SR image reconstruction processes and a brief introduction of each class of the SR image reconstruction approach. A graphic representation of the classification of SR image reconstruction approaches is shown in following tree graphic:
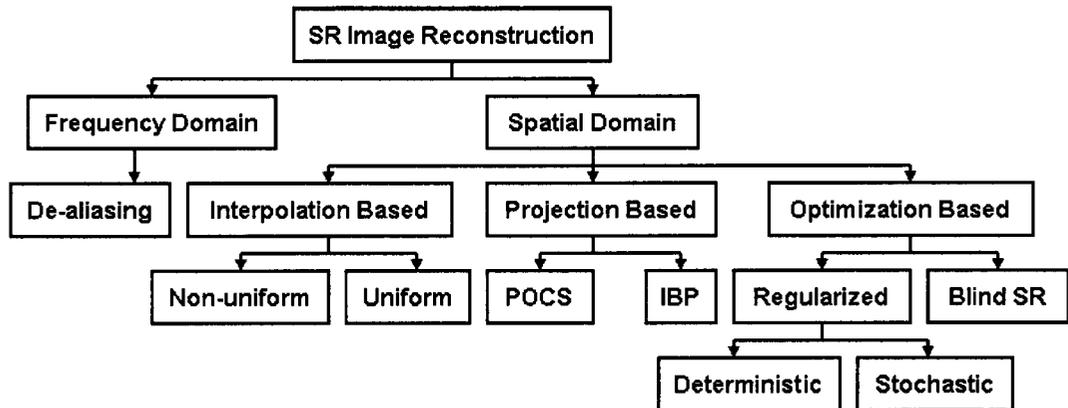
Figure 2.5: Classification of SR image reconstruction process

Considering Figure 2.5 from top to bottom, left to right, the SR image reconstruction processes can be roughly divided into two big classes which are the frequency domain and the spatial domain reconstruction approaches. Intuitively, frequency domain SR image reconstruction process converts the acquired LR frames into frequency domain and the approach mainly focuses on the de-aliasing of acquired LR frames by utilizing Fourier transform shift property. The frequency domain SR image reconstruction was the first

24

proposed multi-frame super-resolution approach which has lower computational complexity and is easier to perform parallel processes. However, frequency domain methods are mostly based on ideal theories; it can only be used when the cases have simple shift motion and linear spatial invariant (LSI) degrading models. Since they are more difficult to apply on most applications, no further methods have been extended under this category.

The spatial domain SR image reconstruction approaches can accommodate more spatial constraints and have more variations. There are three sub-classes classified under spatial domain SR image reconstruction process. (i) Interpolation based SR image reconstruction processes consume less computational resources and are fast enough to apply for real-time SR applications, however they are hard to apply with priori knowledge for a better degradation effects removal. (ii) The projection based SR image reconstruction processes are iterative processes. Both POCS and IBP SR Image reconstruction processes allow users to add priori knowledge to ensure that the final SR image has the desired characteristics. The proposed SR image reconstruction process is categorized under POCS approach and is extended from conventional POCS by adding a wavelet domain convex set and a wavelet coefficient refinement process to improve the final HR image quality. (iii) The optimization based SR image reconstruction processes treat the SR image reconstruction problems as ill-pose or ill-conditioned problem, which estimates the desired HR image by imposing various engineering inverse optimization algorithms such as the least square minimization. These inverse optimization algorithms can be further classified into regularized optimization and blind SR optimization [24] approaches. Regularized optimization approaches reduce the computational complexity of the optimization problems. There are two types of regularized optimization methods. (1) The deterministic types of regularization methods, which optimize the solution by using pre-determined unknown parameters according to priori knowledge such as constrained least square (CLS) method. (2) The stochastic type

of regularization methods apply stochastic estimator methods such as the maximum a posteriori (MAP) estimator to solve the objective function. The other inverse optimization SR image reconstruction approach, the Blind SR Optimization, uses almost only acquired LR frames information without additional assumptions to recover original SR image.

Some of the previously mentioned SR image reconstruction approaches such as uniform interpolation, frequency domain method, IBP, conventional POCS, and regularized optimization SR image reconstruction approaches will be further discussed in detail in Chapter 3.

2.4 Image De-blurring

Image de-blurring is usually classified under image restoration process in most image processing textbooks and is defined as a process that removes the degrading effects; The task to "remove" usually refers to filtering or compensating the unwanted operation. In some SR reconstruction approaches, the de-blurring process is separated from the SR image reconstruction step in order to reduce the computational complexity and the loading in one step. It was found that performing the de-blurring process prior to presenting the final HR image improves most SR image reconstruction algorithms. However, performing the de-blurring process at the last step of super-resolution algorithm is optional.

There are many factors that could cause image blurring or image degradation. The most common reason of blurring is when the camera uses long exposure time when capturing moving objects or capturing steady object while camera is in motion. In super-resolution applications, normally consecutive LR frames are usually taken within very short period of time. As a result, motion blurring can be prevented or minimized by taking high capturing rate. Other factors that cause degradation of images such as out-of-focus optics, air turbulence or lack of exposure time causes insufficient photons to be captured, these conditions of degradation can be treated as blurring images.

Considering the image degradation model described in Equation (2.10), the concept of removing blurring is intuitively simple, and can be treated as a filtering process. The difficult step is to identify the blurring matrix B which has been discussed in detail in Section 3.2. Once the blurring matrix B is estimated by blur identification methods, the operation $x = B^{-1}g - n$ brings back the original image x.

To date, the de-blurring technology is relatively advanced compared to other image processing problems. The simulation program Matlab includes many reliable de-blurring tools such as "Wiener filtering", "regularized filtering" and "recursive Kalman filtering". Among these de-blurring methods, the proposed SR algorithm uses the Matlab function DECONVWNR() which implements the Wiener filtering technique [25] [26], to de-blur the estimated HR image.

The program statement "g = DECONVWNR(x,PSF)", deconvolves image x using the Wiener filtering algorithm, returning a deblurred image g. The assumption is that the image x was created by convolving a true image with a point-spread function (PSF) and possibly by adding noise. The algorithm is optimal in the sense of least mean square errors between the estimated and the true images, and utilizes the correlation matrices of image and noise. If there is no noise present, the Wiener filter is reduced to the ideal inverse filter.

2.5 Conclusion

In this chapter, four fundamental SR reconstruction steps are independently explained. The first step, image registration process, is described by explaining rotation estimation and frame to frame motion estimation methods. In the motion estimation method, the Pyramid motion estimation model is discussed in detail. The second step, the degrading function estimation process uses the parametric blurring model to identify the blurring point spread function (PSF) kernel. The third step, different types of SR image reconstruction process are classified by their approach. In the classification, SR image reconstruction processes

are roughly classified into two big classes which are frequency domain and spatial domain approaches. The spatial domain approaches can be further categorized into several sub-classes. The proposed SR reconstruction algorithm falls into the projection based spatial domain approach. The last step of SR reconstruction process is the image de-blurring process. The de-blurring process is optional in some other SR reconstruction approaches. The proposed SR reconstruction algorithm uses the Wiener filtering method to remove the blurring effects on the estimated HR image.

Chapter 3

Super-Resolution Reconstruction Approaches

There are various types of SR reconstruction algorithms. With reference to the classification of different SR reconstruction approaches shown in Figure 2.5, this chapter discusses in detail some well-known interpolation methods, namely, the "nearest-neighborhood", "bilinear", and "bicubic interpolation". As well, the "frequency domain approach", IBP, "conventional POCS" and the "regularized inverse optimization" will be discussed.

3.1 Traditional Image Interpolation

To obtain a visually clearer image, a very fundamental process is to increase the total number of pixels. The image interpolation process does the work. Image interpolation is a very basic process that almost all SR reconstruction approaches perform during the SR reconstruction procedure. However, image interpolation alone is not considered as a super-resolution technique. This section introduces three commonly used interpolation methods; one of them, the bilinear interpolation method will be employed in the proposed SR reconstruction algorithm.

3.1.1 Nearest-neighborhood Interpolation

When increasing the size of an image, some new unknown pixels may appear on non-integer coordinates, due to the various incremental ratios. Therefore, it is necessary to insert new pixel values, according to neighbor pixels' value. There are many interpolation methods which use neighboring pixels to interpolate new pixel values, in which the simplest one is the nearest-neighborhood interpolation [27]. Its main idea is to find the known integer pixel which has the shortest distance to the unknown non-integer new pixel. The value of each non-integer's new pixels duplicates from its nearest known neighbor pixels.
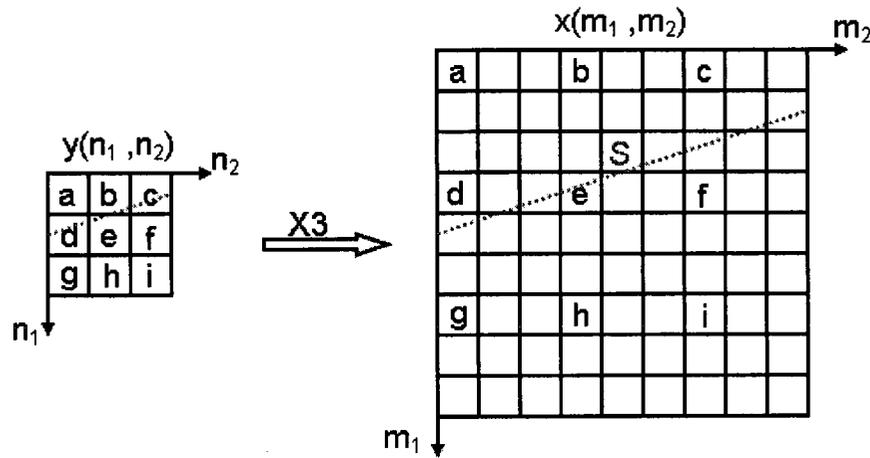
Figure 3.1: Illustration of nearest-neighborhood interpolation

In Figure 3.1, pixel **S** in the magnified image $x(m_1, m_2)$ is located between pixels **b**, **c**, **e** and **f**. The nearest-neighborhood interpolation algorithm estimates the distance from the pixel **S** to its neighbor pixels and chooses the pixel with shortest distance to the pixel **S** which is pixel **e** in this case. Then the value of pixel **S** is assigned by copying the value from the pixel **e**.

The nearest-neighborhood interpolation algorithm is the simplest, and most efficient and computationally frugal image magnification method. However, images interpolated by nearest-neighborhood interpolation are of low quality because the jagged and blocky

30

effects are usually noticeable.

3.1.2 Bilinear Interpolation

Bilinear interpolation is a similar idea to nearest-neighborhood interpolation. It estimates new unknown pixels by referencing its four known neighboring pixels. According to the distance ratio of the new unknown pixels to its four known neighboring pixels, each neighboring pixel obtains a weighing factor which is larger when the neighboring pixel is closer to the unknown pixel, in contrast, the further neighboring pixel receives smaller weighing factor. Figure 3.2 and Equation (3.1) describe how an unknown pixel value is determined by bilinear interpolation.
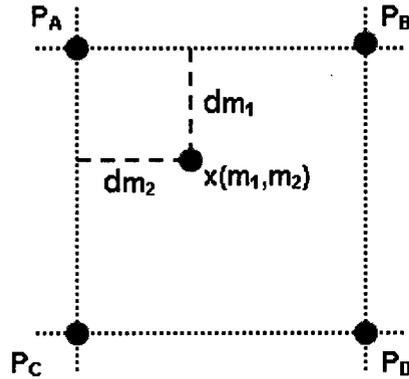
Figure 3.2: Illustration of bilinear interpolation

$$x(m_1, m_2) = (1-d_{m_1})(1-d_{m_2})P_A + d_{m_2}(1-d_{m_1})P_B + (1-d_{m_2})d_{m_1}P_C + d_{m_1}d_{m_2}P_D \quad (3.1)$$

In Figure 3.2, $x(m_1, m_2)$ is the pixel to be interpolated, P_A, P_B, P_C and P_D are the values of four known neighboring pixels of $x(m_1, m_2)$. The known neighboring pixel with shorter distance to the unknown pixel $x(m_1, m_2)$ means that the pixel has greater influence

31

to the new pixel and vice versa.

The images interpolated by the bilinear interpolation method show better continuity and smoother quality compared to the nearest-neighborhood interpolation. There are some enhanced versions of bilinear interpolation [28] which extend the reference pixels from 4 points to 8 points or more. Those enhanced bilinear interpolation methods provide sharper edges presentation, but the complexity of the algorithms are dramatically increased.

In the simulation experiment of this work, bilinear interpolation is included as one of the reference SR approach for comparison. In the SR reconstruction experiment, all the acquired LR frames are first interpolated by bilinear interpolation according to the required enlargement factor. Then, all the interpolated images are aligned by the image registration process introduced in Section 2.1. Finally, the SR image is generated by averaging the sum of each interpolated frame. The purpose of using bilinear interpolation is to build a comparison case and show that the interpolation alone can not reconstruct a good quality SR image, specially when the enlargement factor is big.

3.1.3 Bicubic Interpolation

Bicubic interpolation [29] is also a well known interpolation method. Because of its transfer function, it has a better approximation method to the Sinc function and to any previously mentioned interpolation methods. Hence, the images interpolated by bicubic interpolation have very good quality. The origination of bicubic interpolation is derived from cubic polynomial which is shown in Equation (3.2)

$$
\text{cubic}(d) = \begin{cases} (a+2)|d|^3 - (a+3)|d|^2 + 1, & 0 \le |d| < 1 \\ a|d|^3 - 5a|d|^2 + 8a|d| - 4a, & 1 \le |d| < 2 \\ 0, & \text{elsewhere} \end{cases} \qquad (3.2)
$$

where d is the pixel distance, and a is a parameter that controls the quality of the interpolation. The range of variable a is typically (-0.5,-1) , and in many bicubic interpolation applications a is set to be -1. Therefore, the coefficient of bicubic interpolation is formulated by substituting a=-1 into Equation (3.2) then

$$C(d) = \begin{cases} |d|^3 - 2|d|^2 + 1, & 0 \leq |d| < 1 \\ -|d|^3 + 5|d|^2 - 8|d| + 4, & 1 \leq |d| < 2 \\ 0, & \text{elsewhere} \end{cases} \tag{3.3}$$

In Figure 3.3, the mechanism of bicubic interpolation is shown. It requires 16 neighboring pixels to estimate one new unknown pixel by a convolution operation. During the convolution operation, the 16 neighboring pixels need to multiply by two interpolation coefficients individually. The interpolation coefficients are obtained by substituting the distance between neighboring pixels and the new unknown pixel into Equation (3.3). The value of the new unknown pixel is then obtained by performing the convolution operation as shown below,

$$S = \sum_{m=-1}^{2} \sum_{n=-1}^{2} P_{(i+m,j+n)} C(n-x) C(m-y) \tag{3.4}$$

where S is the value of the unknown pixel, $P_{(i+m,j+n)}$ represents the value of 16 neighboring pixels, $C(n-x)$ and $C(m-y)$ are the convolution coefficients calculated from Equation (3.3).

The bicubic interpolation method produces a smoother texture and causes much less interpolation artifacts on interpolated images than the bilinear interpolation method does. However, it takes longer computational time to complete the same amount of work. This method is often chosen when speed is not an issue.
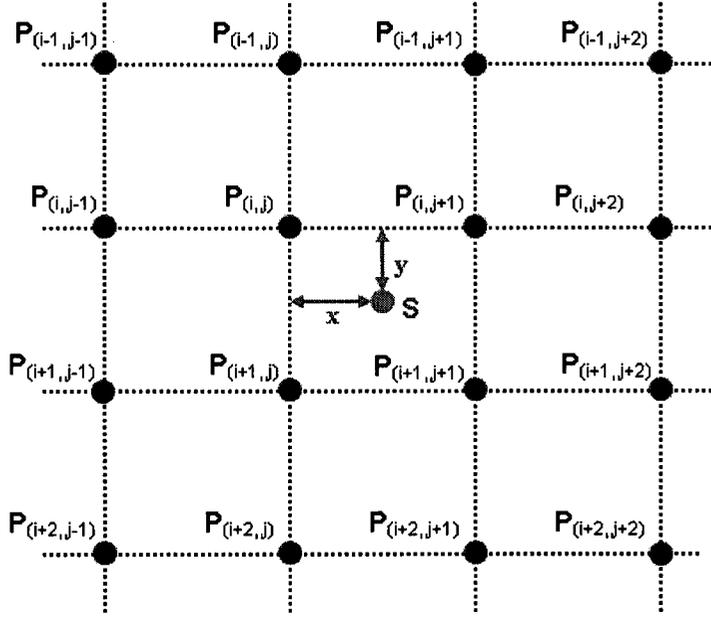
Figure 3.3: Illustration of bicubic interpolation

3.2 Frequency-Domain Approach

The frequency domain approach for super-resolution image reconstruction is a de-aliasing process. An assumption is made that in each acquired LR frame, there is aliasing. The algorithm and system equation were first derived by Tsai and Huang [8]. In their published paper, the relationship between the LR frames and the desired HR image is described. The achievement of the frequency domain approach is based on three basic principles. 1) the original HR image is bandlimited, 2) the shifting properties of Fourier transform, 3) the aliasing relationship between the discrete Fourier transform (DFT) of the acquired LR frames and the continuous Fourier transform (CFT) of the original HR image. These properties relate the aliased DFT coefficients of the acquired LR frames to a sample of the CFT of an unknown HR image. For instance, let two 1-D signals represent an under-sampled LR frame and one of the 1-D signals is transformed into frequency domain. According to the principles described above, the relation between the DFT coefficients of

34

the aliased LR frame and the CFT samples of the desired HR image can be depicted as the frequency domain signals shown in Figure 3.4.
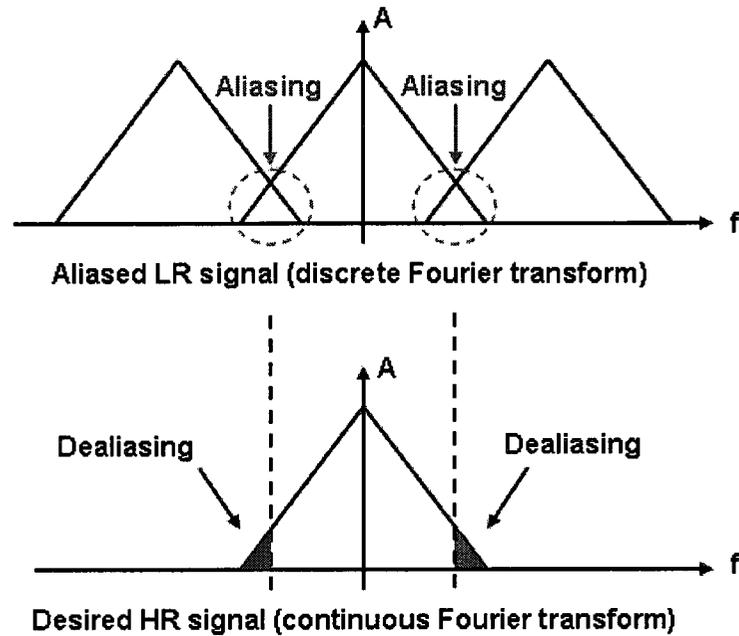
Figure 3.4: Relationship of LR frame and HR image

In Figure 3.4, the waveform at the top is the DFT coefficient of the aliased LR frame and the bottom waveform is the CFT sample of the desired HR image. The goal of the frequency domain approach is to remove the aliasing from the DFT coefficients of the aliased LR frame and to obtain information in the high frequency portion. The mathematical derivation of this method is described in detail as follows. First, the notation of each of the variables maintains the same as those used in the original paper and these notations are valid only in this section.

Variable notations:

(x, y) = real-space coordinates

(u, v) = frequency-space coordinates

$f(x, y)$ = ideal image

$f_k(x, y)$ = ideal image shifted by $(\delta_{k,x}, \delta_{k,y}) = f(x + \delta_{k,y}, y + \delta_{k,y})$

p = number of available LR frames

$\{f_{i,j}{}^k\}$ = (i, j)th sample in the kth frame; $f_{i,j}{}^k = f(iT_x + \delta_{k,x}, jT_y + \delta_{k,y})$

M = number of pixels in LR frame in x-direction

N = number of pixels in LR frame in y-direction

T_x = sampling interval in x-direction

T_y = sampling interval in y-direction

$\delta_{k,x}$ = spatial offset of kth frame in x-direction

$\delta_{k,y}$ = spatial offset of kth frame in y-direction

$F(u, v)$ = Continuous Fourier Transform (CFT) of continuous signal $f(x, y)$

$F_k(u, v)$ = CFT of shifted continuous signal $f_k(x, y)$

ω_{sx} = sampling frequency in x-direction = $\frac{1}{T_x}$

ω_{sy} = sampling frequency in y-direction = $\frac{1}{T_y}$

$F^k{}_{m,n}$ = (m, n)th component of DFT of kth frame $(f_k(x, y))$.

L_u = resolution improvement factor in u

L_v = resolution improvement factor in v

J = complex number $\sqrt{-1}$

Note that the LR frame counter k is defined as $k \in [0, p - 1]$. Then, the 0th frame is defined as the reference frame without shifting in both x and y directions. Therefore $\delta_{0,x} = \delta_{0,y} = 0$. The pixel's counter (i, j) is defined as $i \in [0, M - 1]$ and $j \in [0, N - 1]$.

According to the continuous Fourier transform theory, the relation between the shifted and non-shifted ideal images can be described as

$$F_k(u, v) = F(u, v) * e^{J2\pi(\delta_{k,x}u, \delta_{k,y}v)} \tag{3.5}$$

36

where $F(u, v)$ is the continuous Fourier transform of the continuous signal $f(x, y)$. Relating Equation (3.5) to the discrete space by discrete Fourier transform, then

$$F^k{}_{m,n} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f_{ij}^k * e^{-J2\pi(\frac{im}{M}, \frac{jn}{N})} \tag{3.6}$$

There are p number of available LR frames, thus there are p DFT matrices of size $M \times N$. Relating DFT Equation (3.6) to the CFT Equation (3.5) by the "aliasing relationship", the DFT of kth frame becomes

$$F^k{}_{m,n} = \frac{1}{T_x T_y} \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} F_k(\frac{a}{MT_x} + m\omega_{sx}, \frac{b}{NT_y} + n\omega_{sy}) \tag{3.7}$$

Equation (3.7) relates the shifted (m, n)th frequency component of $F^k{}_{m,n}$, and the representation of the shifted continuous signal. According to the first basic principle, the $F(u, v)$ is assumed to be bandlimited in both directions. Therefore, $F(u, v) = 0$ for all ω satisfies $|u| > L_u * \omega_{sx}$ and $|v| > L_v * \omega_{sy}$. This enables another matrix equation to be formulated.

$$G_{mn} = \Phi_{mn} * F_{mn} \tag{3.8}$$

where G_{mn} is defined as

$$G_{mn} = \begin{bmatrix} F^0{}_{m,n} \\ F^1{}_{m,n} \\ \vdots \\ F^k{}_{m,n} \\ \vdots \\ F^{p-1}{}_{m,n} \end{bmatrix} \tag{3.9}$$

thus, the components inside the vector G_{mn} are the known DFT of $f_k(x, y)$ described in

37

Equation (3.7) and it has p components.

$$\Phi_{mn} = \begin{bmatrix} (\phi_{mn})_{1,1} & (\phi_{mn})_{2,1} & \cdots & \cdots & \cdots & (\phi_{mn})_{4L_uL_v,1} \\ (\phi_{mn})_{1,2} & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & (\phi_{mn})_{l,k} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ (\phi_{mn})_{1,p} & \cdots & \cdots & \cdots & \cdots & (\phi_{mn})_{4L_uL_v,p} \end{bmatrix} \qquad (3.10)$$

with

$$(\phi_{mn})_{l,k} = e^{J2\pi[\delta_{k,x}(\frac{m}{MT_x} - \frac{\gamma}{T_x}) + \delta_{k,y}(\frac{n}{NT_y} - \frac{s}{T_y})]}$$

where $\gamma = i \bmod (2L_u) - L_u$ and $s = \frac{iL_v}{2L_u} - L_v$. The matrix Φ_{mn} is of size $p \times 4L_uL_v$. The Equation (3.9) and Equation (3.10) can be calculated at this stage. However, it is best to avoid the direct calculation of both matrices G_{mn} and Φ_{mn} because direct calculations are computationally expensive. Hence, it is desirable to find the value of $(F_{mn})_i$ first, which can be obtained easier by noting symmetries of the vector G_{mn} and the matrix Φ_{mn}.

The unknown vector F_{mn} at the last term of Equation (3.8) is defined as

38

$$
F_{mn} = \begin{bmatrix} (F_{mn})_1 \\ (F_{mn})_2 \\ \vdots \\ (F_{mn})_i \\ \vdots \\ (F_{mn})_{4L_u L_v} \end{bmatrix}
\qquad\qquad (3.11)
$$

where

$$
(F_{mn})_i = F(\frac{m}{MT_x} + \gamma\omega_{sx}, \frac{n}{NT_y} + s\omega_{sy})
$$

where $\gamma = i \bmod (2L_u) - L_u$ and $s = \frac{iL_u}{2L_u} - L_v$. The vector F_{mn} is noted as the relation between the CFT at a given point in continuous space and values from the DFT. The vector F_{mn} contains $4L_u L_v$ components and now each value $(F_{mn})_i$ located in the continuous (u, v) coordinate can be exactly calculated.

Once Equation (3.8) is solved, it means the values in vector F_{mn} are available and those are the values of the continuous Fourier transform at a given (u, v) point. Assigning (u, v) coordinates to the $4L_u L_v$ values in the frequency domain as sample point of the frequency representation of the continuous ideal image, then

$$
(u, v) = (\frac{m}{MT_x} + \gamma\omega_{sx}, \frac{n}{NT_y} + s\omega_{sy})
$$

The matrix Equation (3.8) considers only the (m, n)th component of all the available LR frames p of the DFTs. There are totally $N \times M$ samples in each frame, thus we have $(N - 1)(M - 1)$ more matrix equations to solve. Between each sample, we want to insert $4L_u L_v$ new samples, therefore there are $4L_u L_v(N - 1)(M - 1)$ more points to add to the frequency representation of the continuous ideal image. In total, there will be $4L_u L_v MN$ points in the representation, whereas if it starts with MN, then the most points could be

39

obtained from any image. Hence, the resolution is improved by a factor of $4L_uL_v$. If the desired image resolution improvement factor is 2 in each direction (ie: 4 in total), then let $L_u = L_v = 1$. In general, the values of L_u and L_v state small values in the frequency domain approach.

In order to recover the HR image as perfect as possible, sufficient points are needed to get beyond the Nyquist sampling frequency with respect to different types of the real scene. There is no doubt that the more points added to the band-limited frequency representation, the more accurate the signal is obtained in the space domain. Finally, the HR image is obtained by taking the 2-D inverse Fourier transform after all points from each (m, n) have been mapped into the frequency space.

The major advantage of the frequency domain approach is the simplicity of its theory, in which the relationship between the LR and the HR images is clearly defined in the frequency domain. Its frequency domain operations are also not difficult to implement either by software or hardware. However, the degrading model of this approach is restricted to only global frame to frame motion. Moreover, the linear spatial invariant (LSI) blurring model is restricted due to the difficulty of applying the spatial domain data correlation and priori knowledge constraint processes.

3.3   Inverse Back-Projection (IBP)

Inverse back-projection (IBP) is an iterative SR reconstruction method proposed by M. Irani and S. Peleg in the early 1990's. In this approach, the HR image is obtained through iteratively reducing the difference of the simulated LR frames and the observed LR frames. The concept of the IBP approach is the basis of many different iterative image reconstruction algorithms, for instance the idea of POCS was inspired from the IBP concept. In the recent years, many SR reconstruction algorithms have been developed by referencing the back-projection related concepts.

The IBP SR reconstruction is an efficient and simple SR reconstruction method. The HR image is reconstructed by minimizing the errors between the simulated LR frames and the acquired LR frames during the iteration. This method can be combined with the desired constraints such as smoothness of HR image or specific characteristic of other restrictive conditions.

The mathematical derivation is described in detail in the following by assuming that there are p available LR frames $y_1, y_2, \ldots, y_p$, each of which has $n_1 \times n_2$ pixels in dimension, and the desired HR image x has $m_1 \times m_2$ pixels in dimension. Let the magnification factor be μ, then the HR image size can be obtained by extending the LR frame y both in horizontal and vertical direction by μ times. Thus, $m_1 \times m_2 = \mu n_1 \times \mu n_2$. The degrading model of the IBP SR reconstruction is described as

$$y_i = (M_i Bx + \eta_i) \downarrow \mu \qquad 1 \leq i \leq p \qquad (3.12)$$

where M_i is the frame to frame relative motion transfer function, B is the blurring or degrading transfer function, η_i is the additive noise, and $\downarrow \mu$ is the down-sampling operation.

Assuming that the effects of the motion transfer function can be compensated by image registration and the blurring or degrading transfer function can be estimated, then the estimated "LR" frame $\tilde{y}$ can be described by

$$\tilde{y}_i^n = (M_i B(\tilde{x}_i^n) + \eta_i) \downarrow \mu \qquad 1 \leq i \leq n \qquad (3.13)$$

where $\tilde{y}_i^n$ is the estimated LR frame generated from the ith frame at nth iteration, $\tilde{x}_i^n$ is the estimated HR image simulated from the ith frame at the nth iteration.

According to the IBP SR reconstruction algorithm, the estimated HR image can be presented as

$$\tilde{x}^{n+1} = \tilde{x}^n + \frac{1}{i} \sum_{i=1}^{i} M_i^{-1}(((\tilde{y}_i^n - y_i) \uparrow \mu) * h^{BP}) \qquad (3.14)$$

41

where M_i^{-1} is the motion compensation operation, $\uparrow \mu$ is the up-sampling operation and h^{BP} is the inverse projection transfer function, h^{BP} can be chosen arbitrarily and the choice of h^{BP} affects the characteristics of the reconstructed HR image. Therefore, h^{BP} can be utilized as an additional constraint which represents the desired properties of the solution. For a better explanation, the process of the IBP SR reconstruction method is shown in Figure 3.5.

Projection:
Produces simulated LR images Hx

Hx   y   (y - Hx)   **Back-projection:**
Update HR image by
adding   h^{BP} (y - Hx)

$\otimes$ h^{BP}

Simulated LR   **Acquired LR**
images   **frames**

Additional constraint   h^{BP}
(desire property of solution)
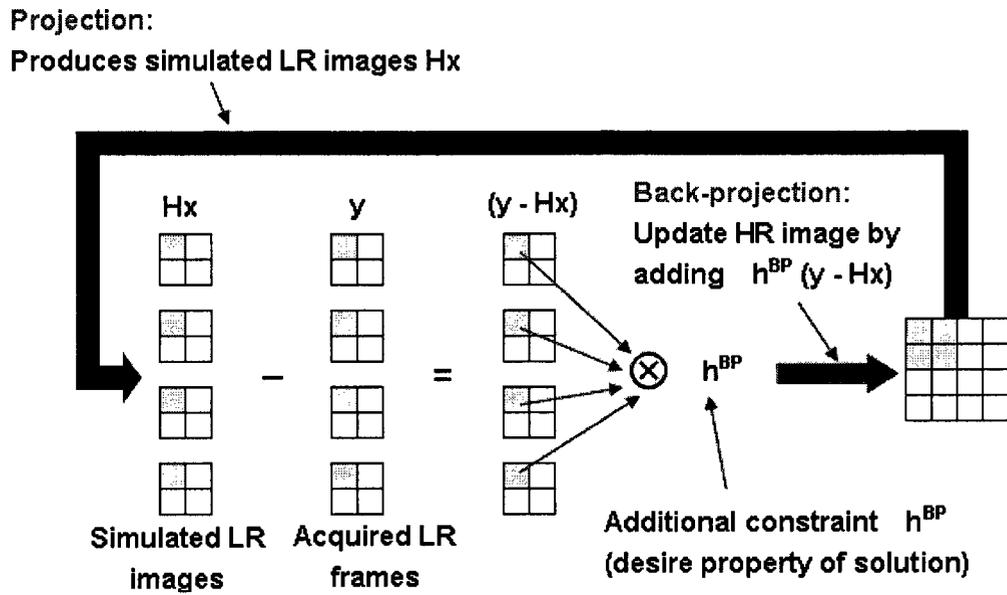
Figure 3.5: The framework of inverse back-projection of SR image reconstruction

In Figure 3.5, H represents the complete degrading operation which includes motion operation, blurring operation, and down-sampling operation.

The advantage of IBP is that it can be intuitively understood. The disadvantage, however, is that this method does not guarantee a unique solution because of the nature of ill-posed inverse problems. Therefore, in some situations it is diffcult to choose the inverse projection kernel h^{BP}.

3.4 Projections onto Convex Sets (POCS)

In many super-resolution research literatures, the technical term "Projection onto Convex Sets" (POCS) is also frequently encountered. POCS is a method that can be applied to different signal recovery applications. This section provides the basics of the POCS method, introduces the commonly used signals that can be formulated as convex sets and explains the conventional POCS SR reconstruction algorithm. The proposed SR reconstruction algorithm is an extended version of this conventional POCS approach, therefore some procedures explained in this section will be used in the proposed SR reconstruction algorithm in the next chapter.

3.4.1 Basics of POCS

Iterative projections onto convex sets (POCS) [30] is a useful tool for image recovery and restoration. The basic idea behind this is that the desired incomplete image can be defined by several convex signal sets and these convex sets work as constraint parameters. A signal to be recovered that contains all the desired properties can be reconstructed by iteratively projecting onto these convex constraint sets. Practically, some convex constraint parameter sets are commonly used in image reconstruction. Those convex constraint sets include the sets of band-limited signals, duration-limited signals, the signals that are the same (e.g., zero) on some given interval, bounded signals, signals of a given area, and complex signals with a specified phase. More information about signal convexity properties will be provided in sub-section 3.4.2.

POCS was first proposed by Bregman [31] and Gubin et al. [32] and was later popularized by Youla and Webb [11]. However, Sezan and Stark [12] initialized the application of POCS on image signal recovery. There have been a big variety of applications imposing the POCS method such as sampling theory, signal recovery, deconvolution and extrapolation, tomography, and time-frequency analysis.

43

To understand projections onto convex spaces, first, the concept and the definition of convex sets have to be explained. A convex set is defined simply as a bounded space where any points on a straight line between any two points in this space is always within this bounded space. In other words, any point in a convex set can be seen by other points without any obstacle in between including the set boundary. A geometric demonstration of convex sets and none-convex sets are shown in Figure 3.6.
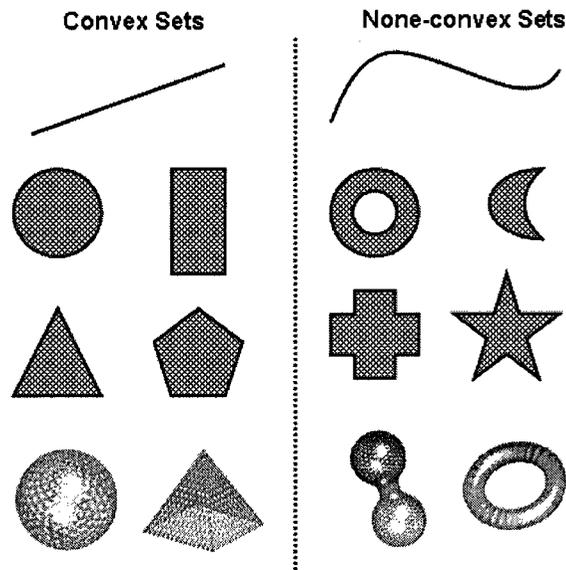
Figure 3.6: An example of convex sets and none-convex sets

Examples of convex sets and non-convex sets are separated by the central dash-line in Figure 3.6. In a 1-D scenario, points in a straight line are considered as a convex set., In contrast, points in a curve line are not. In a 2-D scenario, every point in the filled circle, rectangle, triangle and pentagon are convex sets. However the shape of the ring, the crescent moon, star and cross shapes are non-convex sets, due to there being points in these shapes that can not be seen by some other points. The same method can be applied to examine data sets in 3-D objects. For the elements in multi-dimensional objects, it is difficult to identify the convexity visually. Therefore, the convexity must be examined by

44

the convexity definition which is described by Equation (3.15) in Section 3.4.2.

Convex set has an interesting property in that the projection of a given vector $\vec{V}$ onto a convex set denoted A, results in a unique point which is the closest, and most importantly projects $\vec{V}$ on A, as shown in Figure 3.7.
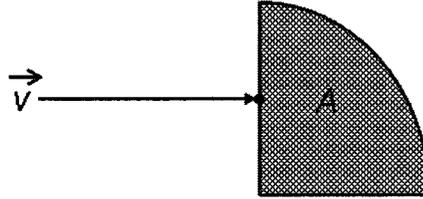
Figure 3.7: Vector V projected onto convex set A

There are three cases of operating projections onto convex sets. The first case is that of POCS convergence to a point, meaning that convex sets have an intersection. Figure 3.8 illustrates the first case of convex set projection, where the convex set A is a straight line and the other convex set B is the shaded area. In this case a given starting point $\vec{O}^{(0)}$ consecutively projects onto the convex sets and results in a point which is the intersection of convex sets. This case is desired when employ POCS to solve SR problems, because it means the process is convergence. In other word, the unknown HR image can be found. Figure 3.9 illustrates the second case of convex set projection which the POCS solves for the unique shortest distance between the two sets for the two convex sets do not have an intersection. In the third case, three independent convex sets are subject to the projection operation which does not have an intersection. In this case, the projection operation can not convergence. However, this case is rarely encountered in super-resolution applications, where the POCS solution will oscillate among the respective nearest distances between each set when there is no intersection among multiple convex sets and this case is geometrically displayed in Figure 3.10.

These geometrical illustrations briefly explain the properties of convex sets and POCS. However, POCS does not seem to be a simple tool for signal processing. In fact, the

45
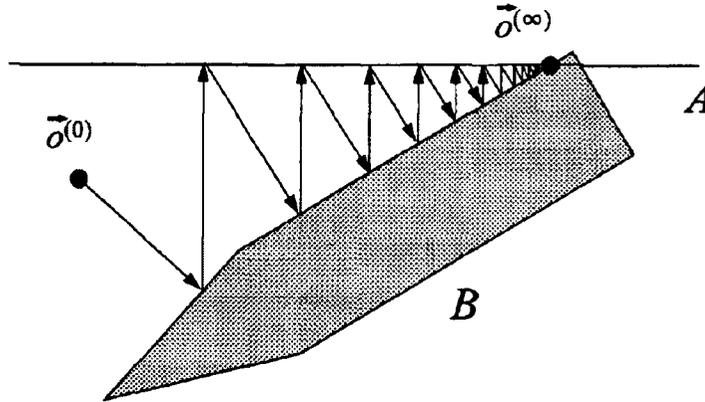
Figure 3.8: Case 1: An intersection of convex sets
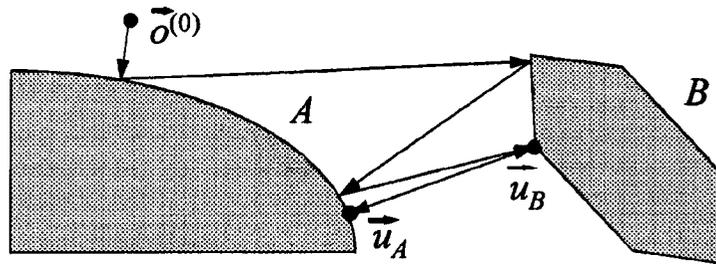
Figure 3.9: Case 2: An unique shortest distance between the two sets

most simple case of POCS application is signal filtering. For example, a low-pass filtering process where the filtered signal denotes one convex set (B) and the input signal denotes another convex set (A) are defined. The POCS method projects set (A) onto set (B) resulting simply in set (B) being included in set (A). Another early application of POCS was the recovering of a random 1-D phase-only (not zero-phase) signal [33]. This is done by calculating the Fourier transform of the acquired signal, and setting some specific interval of the amplitude spectrum to a constant. In this case, the constraint convex sets are the phase spectrum in the frequency domain and the signal series length in the time domain. During the recovery procedure, successive Fourier transforms/inverse Fourier transform are performed, then the subsequent projections will converge to the point of intersection of those sets. As a result, a close form of the original signal can be obtained.
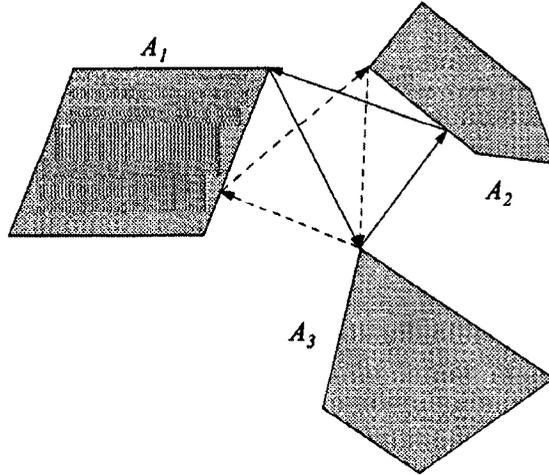
Figure 3.10: Case 3: No intersection in the POCS solution

3.4.2 Commonly Used Signals as Convex Sets

In the video and image signal processing, it is very common to deal with different types of signals. In order to impose the POCS technique for the signal recovery processes, it is very important to know the definition of convex signal and understand the convex properties of signals. Hence, the test of the convexity of different types of signals is explained in the following.

A set, **A**, is a convex set if for every vector $\vec{u} \in$ **A** and every $\vec{v} \in$ **A**, it follows that $\lambda\vec{u} + (1 - \lambda)\vec{v} \in$ **A** for all $0 \leq \lambda \leq 1$. According to the statement above, the convexity test for a set is defined as

$$\{\lambda S_1 + (1 - \lambda)S_2\} \in \mathbf{S} \quad \forall(0 \leq \lambda \leq 1) \tag{3.15}$$

where **S** is the set to be tested, S_1 and S_2 denote arbitrary sub-sets in **S**. Then **S** is said to be a convex set if for all $0 \leq \lambda \leq 1$ the corresponding sub-sets are in the set **S**.

The following list of signals have been proved to possess convex properties [34].

1. Band-limited signals

2. Signals with limited duration

3. Signals with real and positive Fourier transform

4. The constant area of a signal

5. Signals with bounded energy

6. Signals with constant phase

7. Signal bounded by a real function

The above-listed signals can be treated as a priori knowledge when recovering a partially-known signal. In the POCS process, a priori knowledge can be defined as a convex set when the desired signal has the properties listed above.

The POCS method is not only used for 1-D signal recovery; the same manner can also be applied on 2-D image recovery. The conventional POCS algorithm for super-resolution image reconstruction is provided in Sub-section 3.4.3

3.4.3 Conventional POCS Super-Resolution Algorithm

The POCS method is an iterative approach combining priori knowledge about the desired high quality image into the reconstruction process. When converting priori knowledge to the convex constraint sets, the POCS method can estimate the SR image by performing restoration and interpolation simultaneously.

Stark and Oskoui [13] initially formulated the SR reconstruction by using POCS method. After Stark and Oskoui's first publication, Tekalp et al. [3] extended their formulation by including additive noise. According to the method of POCS [35], a SR image x is assumed to have m cases of priori knowledge, then it reflects that the m cases of priori knowledge in Hilbert space exists m constraint sets $C_1, C_2, \ldots, C_m$, we denote that C_s is the intersection of these constraint sets. Combining those priori knowledge into the POCS

process can be explained as constraining the SR result to be a member of a closed convex set $C_k(k = 1, 2, \ldots, m)$ that is defined as a set of vectors which satisfies a particular property. If all the constraint sets have an intersection, then a SR result that belongs to the intersection set $C_s = \cap_{k=1}^{m} C_k$, and is also a convex set, which can be found by iteratively projecting onto these convex sets. The resulting SR image is a feasible solution, since any SR estimated image in the intersection of all the constraint sets is consistent with the prior constraints. Therefore, the POCS can be applied to find a vector that belongs to the intersection by iterating the projection process.

$$x^{n+1} = P_m P_{m-1} \ldots P_2 P_1 x^n \qquad (3.16)$$

The POCS method starts with an arbitrary initial HR image x^0 in the Hilbert space. P_k $(k = 1, 2, \ldots, m)$ are the projection operators. Each of them projects signal x onto a corresponding convex set C_k $(k = 1, 2, \ldots, m)$. Assuming that the motion information is accurate, the constraint set of a data consistency based on the observation model shown in Equation (1.2) is formulated to represent each pixel in the LR images $y_i(n_1, n_2)$. The data consistency constraint is given by

$$C_D(n_1, n_2) = \{x(m_1, m_2) : |r^{(x)}(n_1, n_2)| \leq \delta_i(n_1, n_2)\} \qquad (3.17)$$

Where the m_1 and m_2 are the pixel indexes of HR image, that are defined as $m_1 = \mu n_1$, $m_2 = \mu n_2$, and the associated residual $r^{(x)}(n_1, n_2)$ is defined as: [36]

$$r^{(x)}(n_1, n_2) = y_i(n_1, n_2) - H_i(n_1, n_2; m_1, m_2)x(m_1, m_2) + \eta_i(n_1, n_2) \qquad (3.18)$$

is bounded in the magnitude of $\delta_i(n_1, n_2)$. The ideal image $x(m_1, m_2)$ is a number of the sets that satisfy a certain statistical confidence, where the boundary $\delta_i(n_1, n_2)$ is determined

by the statistics of noise. For example, the additive Gaussian noise η_i in the observation model has a standard deviation σ_η, then $\delta_i(n_1, n_2) = C\sigma_\eta$, $C \geq 0$ that is determined by an appropriate statistical confidence bound so that $\delta_i(n_1, n_2)$ is large enough to guarantee that the intersection of the solution sets with respect to different constraints is not empty. The projector P_D projects an initial estimated HR image $x^0(m_1, m_2)$ onto convex set C_D that is defined by [3]

$$x^{n+1}(m_1, m_2) = P_{Di}[x^n(m_1, m_2)] = x^n(m_1, m_2) +$$

$$\begin{cases} \dfrac{(r^{(x)}(n_1, n_2) - \delta_i(n_1, n_2))\mathbf{H}_i(m_1, m_2; n_1, n_2)}{\sum_{o_1} \sum_{o_2} \mathbf{H}_i^2(o_1, o_2; n_1, n_2)}, & r^{(x)} > \delta_i(n_1, n_2), \quad x \notin C_D \\ 0, & |r^{(x)}| \leq \delta_i(n_1, n_2), \quad x \in C_D \\ \dfrac{(r^{(x)}(n_1, n_2) + \delta_i(n_1, n_2))\mathbf{H}_i(m_1, m_2; n_1, n_2)}{\sum_{o_1} \sum_{o_2} \mathbf{H}_i^2(o_1, o_2; n_1, n_2)}, & r^{(x)} < -\delta_i(n_1, n_2), \quad x \notin C_D \end{cases}$$

$$(3.19)$$

where n represents the number of iterations and i represents the counter of available LR frames. $\mathbf{H}_i$ is the degradation function defined in Equation (1.2). The projection operation shown in Equation (3.19) is successively performed until the information of all available LR frames have been used. A single convex constraint projection operation is not enough to reconstruct a good quality HR image. Typically, a conventional POCS SR reconstruction algorithm consists of at least two convex set projections. Hence, it is necessary to add another constraint, which is the amplitude constraint C_A. For a 8-bit gray level image system, the amplitude constraint C_A is defined as

$$C_A(n_1, n_2) = \{x(m_1, m_2) : 0 \leq x(m_1, m_2) \leq 255\} \qquad (3.20)$$

The projector P_A projects an estimated HR image $x^n(m_1, m_2)$ onto the convex set C_A

50

is defined by

$$x^{n+1}(m_1, m_2) = P_A[x^n(m_1, m_2)] = \begin{cases} 0, & x(m_1, m_2) < 0, & x \notin C_A \\ x(m_1, m_2), & 0 \geq x(m_1, m_2) \leq 255, & x \in C_A \\ 255, & x(m_1, m_2) > 255, & x \notin C_A \end{cases}$$

(3.21)

According to the multiple projection operation described in Equation (3.16), the final estimated HR image $x(m_1, m_2)$ is obtained by combining the two projection operators P_D and P_A:

$$x^{n+1} = P_A[P_{Di}[x^n]]$$

(3.22)

3.5   Regularized Inverse Optimization Problem

Thanks to the high computational ability of the state of the art computers, solving huge data set and complicated engineering inverse optimization problems, such as SR image reconstruction problem, is becoming more and more achievable. Conventionally, solving an inverse optimization problem usually starts with firstly determining the cost function (objective function), then minimizing the error of the cost function to get the optimal result. However, most practical inverse optimization problems are not that simple to solve. Inverse problems typically lead to mathematical models that are ill-posed. Specially, their solution is unstable due to the data perturbations, so that special numerical methods that can be overcomed with these instabilities, so-called "regularization", need to be applied.

In order to stabilize the inversion and reduce the computational complexity, regularization methods such as Tikhonov regularization, constrained least square (CLS), probability regularization estimator, maximum a posteriori (MAP) and maximum likely-hood (ML)

estimator are commonly used to solve an inverse optimization problems. The following sub-sections provide the derivation of each of the above mentioned regularization methods for the super-resolution inverse problems.

3.5.1 Tikhonov Regularization

Recall the complete system of super-resolution problem $\mathbf{y}_i = \mathbf{H}_i\mathbf{x} + \eta_i$, where $\mathbf{x}$ is the desired unknown HR image. $\mathbf{y}_i$ represents the acquired LR frames, and $\mathbf{H}_i$ is the ill-conditioned complete degrading function.

If the SR problem is an over-determined system, the standard approach to solve the linear equations is given as

$$\mathbf{H}_i\mathbf{x} = \mathbf{y}_i \tag{3.23}$$

The method to solve Equation (3.23) is known as linear least squares and its solution is to seek the minimum residual of following equation:

$$\sum_{i=1}^{p} \left[\|\mathbf{H}_i\mathbf{x} - \mathbf{y}_i\|^2 \right] \tag{3.24}$$

where $\| \cdot \|$ represents the Euclidean norm. However, in practice the SR complete degrading function $\mathbf{H}_i$ is an ill-conditioned matrix. Therefore, when solving Equation (3.24) directly, even a small change in $\mathbf{y}_i$ can result in big oscillations when approximating $\mathbf{x}$. In order to obtain a reasonable estimate for $\mathbf{x}$, the regularization term is added in this minimization problem:

$$\sum_{i=1}^{p} \left[\|\mathbf{H}_i\mathbf{x} - \mathbf{y}_i\|^2 + \|\Gamma\mathbf{y}_i\|^2 \right] \tag{3.25}$$

where Γ is Tikhonov's matrix, which serves as a stabilization matrix. In many cases, this matrix is chosen as the identity matrix $\Gamma = \mathbf{I}$. However, Γ can also incorporate some priori knowledge of the problem such as the degree of smoothness. This regularization also

52

improves the conditioning of the problem, thus enabling a numerical solution. Minimizing Equation (3.25), the relation of $\tilde{x}$ and y_i can be derived as: [37]

$$\sum_{i=1}^{p} \left[(H_i^T H_i + \Gamma^T \Gamma) \right] \tilde{x} = \sum_{i=1}^{p} \left[H_i^T y_i \right]$$

Then, an explicit solution of the estimated HR image $\tilde{x}$, is given by

$$\tilde{x} = \sum_{i=1}^{p} \left[(H_i^T H_i + \Gamma^T \Gamma)^{-1} H_i^T y_i \right] \tag{3.26}$$

The Tikhonov regularization is the most commonly used method in many engineering inverse optimization problems. The algorithm is very simple and easy to implement. However, in its approximation, the process very often encounters huge matrix inversion operation. Consequently, the result will either not be found if the matrix inverse does not exist, or consumes longer computational time.

3.5.2 Constrained Least Squares

The constrained least squares (CLS) regularization method shares a very similar scheme idea with the Tikhonov's regularization method. The goal of applying CLS regularization to solve an ill-conditioned system is also to stabilize the inversion of the optimization problem. Therefore, the inverse problem in Equation (3.23) can be reformulated by the constrained least squares regularization as:

$$\sum_{i=1}^{p} \left[\|H_i x - y_i\|^2 + \lambda \|C y_i\|^2 \right] \tag{3.27}$$

where λ is the Lagrange multiplier serving as a regularization parameter, and C is the stabilization matrix. Generally, C represents the desired properties and a priori knowledge of the HR image usually, a high-pass filter function is used because most images are naturally smooth. The regularization parameter λ controls the tradeoff of the HR image to be more

depending on the acquired data or higher smoothness. A smaller value of λ relatively increases the portion of the term $\|\mathbf{H}_i\mathbf{x} - \mathbf{y}_i\|^2$ which makes the result more dependent on the acquired data. In contrast, a larger value of λ weighs more on the term $\|\mathbf{Cy}_i\|^2$ which increases the effects of the high-pass operation, so that the result becomes smoother. By minimizing the cost function described in Equation (3.28), a basic deterministic iterative technique is used, hence, the Equation (3.28) is reformulated as

$$\Big[\sum_{i=1}^{p} \Big(\mathbf{H}_i{}^T\mathbf{H}_i + \lambda \mathbf{C}^T\mathbf{C} \Big) \Big] \widetilde{\mathbf{x}} = \sum_{i=1}^{p} \Big(\mathbf{H}_i{}^T\mathbf{y}_i \Big)$$

Then, an unique solution of $\widetilde{\mathbf{x}}$ can be found as given below,

$$\widetilde{\mathbf{x}}^{n+1} = \widetilde{\mathbf{x}}^n + \beta \Big[\sum_{i=1}^{p} \Big(\mathbf{H}_i{}^T(\mathbf{y}_i - \mathbf{H}_i\widetilde{\mathbf{x}}^n) \Big) - \lambda \mathbf{C}^T\mathbf{C}\widetilde{\mathbf{x}}^n \Big] \qquad (3.28)$$

where β is a convergence parameter, and the transpose of the complete degrading function $\mathbf{H}_i{}^T$ represents an up-sampling operation and a type of blurring and motion operation.

3.5.3   Bayesian Maximum a Posteriori (MAP)

In Bayesian statistics, a maximum a posteriori (MAP) [38] probability estimation is a statistic mode of the posterior distribution. In SR image reconstruction applications, the MAP can be used to obtain unobserved pixels based on acquired LR frames. The MAP estimation method is closely related to another probability estimation method known as maximum likelihood (ML). An ML estimation is a special case of MAP estimation with no priori item. The MAP probability estimation is more favored than the ML estimation to be used for stochastic regularization on SR image reconstruction.

The SR image reconstruction methods of this family consider the HR image $\mathbf{x}$ and degrading parameter $\mathbf{h}$ as random fields described by the joint priori probability density function (PDF) $P(\mathbf{x}, \mathbf{h})$. The Bayesian MAP estimation methods are used when a posteriori

PDF of the desired HR image can be established. The procedure of a typical Bayesian MAP reconstruction method [39] is described below.

The first step is to form a joint MAP estimation of the HR image x and the degrading parameter h given by the observed LR frames y, thus the joint posterior probability of x and h are $P(x, h|y_i)$ conditioned on the observed y. Then the estimation can be computed as:

$$x = \arg \max P(x, h|y_1, y_2, \cdots, y_p) \tag{3.29}$$

Applying the Bayes theorem on Equation (3.29), the estimation of x and h can be rewritten as:

$$\widetilde{x}, \widetilde{h} = \arg \max P\left(\frac{P(y|x, h)P(x, h)}{P(y)} \right) \tag{3.30}$$

Since x and h are statistically independent, $P(x, h)$ can be separated as $P(x)P(h)$. Moreover, $P(y)$ is not a function of x or h , so it could be omitted from the optimization process with respect to x and h . Therefore, the estimation is given by:

$$\widetilde{x}, \widetilde{h} = \arg \max P\big(P(y|x, h)P(x)P(h)\big) \tag{3.31}$$

Taking the natural logarithmic function of Equation (3.31) and minimizing the negative logarithm of Equation (3.31), the MAP optimization problem can be expressed as:

$$\widetilde{x}, \widetilde{h} = \arg \min P\big(- \ln[P(y|x, h)] - \ln[P(x)] - \ln[P(h)]\big) \tag{3.32}$$

In Equation (3.32), $P(y|x, h)$ is the PDF of the prior image. $P(h)$ is the PDF of the prior degrading parameter, and the conditional PDF is $P(y|x, h)$. These three PDFs will be defined by a priori knowledge according to the HR image x and the noise statistical information.

The priori model for the degrading parameters h is highly application specific. In general, P(h) may be dropped from the cost function, implying that we do not have any priori knowledge about the degrading parameters.

A stable and effective solution to the regularized SR inverse problem can be estimated accurately because of the MAP optimization in Equation (3.32) which includes priori knowledge represented by P(x). By using Bayesian MAP estimation method, possible solutions can be distinguished either using a priori image or Markov random fields (MRF) priors [40]. The MRF priors [41] [42] are often adopted in order to provide mighty methods for image prior modeling. Utilizing the MRF prior, the PDF of the prior image P(x) is described by a Gibbs prior [43] [44], so the prior image PDF which preserves convexity can be expressed as:

$$P(x) = \frac{1}{C}e^{-\alpha f(x)} = \frac{1}{C}e^{-\sum_{g \in S} \varphi_g(x)} \tag{3.33}$$

where C is a normalizing constant ensures that P(x) is a probability, α is a positive constant, and $f(x)$ is a function of the HR image x. This function is usually called energy function and may be chosen to encourage the neighboring pixels to have similar values so that the first derivative of the HR image function is continuous, $\varphi_g(x)$ is a potential function of the derivative of the image, g is the pixel group and S is the set of group g. The result of function $\varphi_g(x)$ only depends on the value of pixels located within the group g. Normally, an image is assumed to be globally smooth, therefore, the estimation problem is incorporated into a Gaussian prior.

If the noise elements of η_i in Equation (1.2) are assumed to be independent and identically distributed (iid), Gaussian distribution with zero mean, and the error between each LR frame is assumed to be independent. Then, the optimization problem can then be expressed as:

$$\tilde{\mathbf{x}} = \arg\min \left[\sum_{i=1}^{p} \|\mathbf{H}_i\mathbf{x} - \mathbf{y}\|^2 + \lambda \sum_{g \in S} \varphi_g(\mathbf{x}) \right] \qquad (3.34)$$

where λ is the regularization parameter.

Compared to other regularized optimization methods, the Bayesian Maximum a Priori estimation is more convenient to control the use of priori knowledge such as Gaussian prior, MRF prior, and edge-preserving image prior. However, this method heavily relies on probability theories, thus the implementation has more challenges.

3.6   Conclusion

Five different types of SR reconstruction approaches have been discussed in this chapter. The first type is the uniform image interpolation method such as nearest- neighborhood, bilinear and bicubic interpolation. Although, interpolation methods are not considered as SR reconstruction approaches, the interpolation methods still play an important role in SR reconstruction process. This is because most SR reconstructions perform interpolation during the reconstruction process. The second SR reconstruction approach is the frequency domain approach which was firstly introduced as a multi-frame SR reconstruction approach. The frequency domain approach is explained and will be employed as one of the comparison algorithms. The third SR reconstruction approach is the inverse back-projection approach, and its framework and mathematical derivation are provided. The fourth type of the SR reconstruction approach is the conventional projections onto convex sets approach. This approach is the foundation of the proposed SR reconstruction approach. There are two constraint sets employed as part of the proposed SR reconstruction process. The basic concept of convex sets and its projection operations are clearly explained. The fifth type of SR reconstruction approach is the regularized inverse optimization approach. This type of SR reconstruction approach can be further classified into deterministic regularization and

57

stochastic regularization processes. Two deterministic methods: Tikhonov's regularization and constrained least squares, and one stochastic regularization method, the Bayesian maximum a posteriori (MAP), have been briefly introduced.

Chapter 4

Wavelet-based POCS and Stochastic Refinement Process for Super-Resolution

Compared to "Fourier" signal analysis, "wavelet" analysis is a relatively new tool in the video and image signal processing areas. Also, there is no doubt that many signal processing methods strongly rely on various stochastic concepts. This chapter proposes a wavelet-based POCS super-resolution algorithm which formulates a wavelet convex set and uses the stochastic process in the conventional POCS approach for wavelet coefficient sub-image refinement. This chapter first addresses the framework of the proposed algorithm in Section 4.1. Then, in Section 4.2 the wavelet analysis of the estimated HR image and how wavelet transformation is used to form a convex set in POCS process, are introduced. Section 4.3 explains the stochastic refinement process which corrects wavelet coefficients according to a stochastic distribution in order to distinguish useful and faulty estimated pixels. The complete procedure of the proposed wavelet-based POCS algorithm is summarized in Section 4.4.

4.1 The Framework of Wavelet-Based POCS

The proposed SR reconstruction approach is based on the conventional POCS discussed in Section 3.4. According to the result of the conventional POCS SR method, we want to improve its quality and reduced the particular noise. Also, based on the concept of the POCS process, if we can use more priori knowledge to form a convex set into the POCS process, the final result of the estimated HR image would be very close to the original scene. As a result, a new wavelet-based convex set and a wavelet domain coefficient sub-images modification process are added into the POCS iterative process, in order to produce a better HR image and rapidly remove the noise. For a better understanding of the proposed approach, a graphical illustration of it is provided in Figure 4.1 to show the block diagram of the proposed wavelet-based POCS SR image reconstruction.
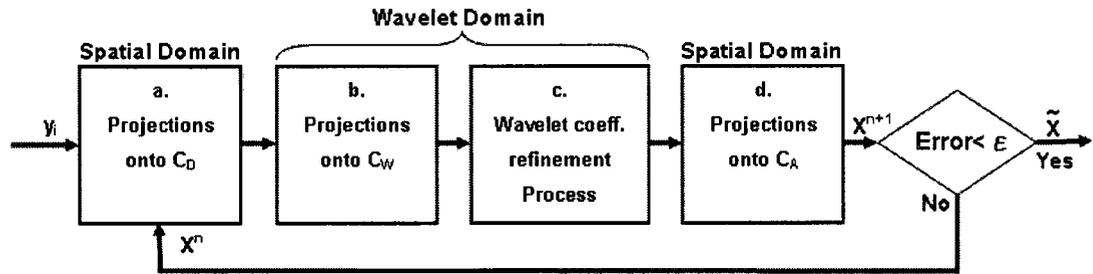
Figure 4.1: Framework of proposed wavelet-based POCS SR image reconstruction steps

From the left hand side of Figure 4.1, the LR frames y_i are input to first projection process onto the constraint set C_D by performing the projection operation described in Equation (3.19). After projection onto C_D convex set, the estimated HR image is forwarded to the next projection operation. In the second block (b.), the estimated HR image is projected onto wavelet constraint C_W. In this block, the previously estimated HR image is transformed into the wavelet domain. After the wavelet convex set projection, the continuing process (block c.) is a wavelet coefficient refinement process. The detailed operation of the wavelet domain projection and wavelet coefficient refinement process will be detailed

60

in Section 4.2 and Section 4.3, respectively. The last projection process of the proposed POCS SR image reconstruction is the projection onto amplitude constraint C_A convex set which has been described in Equation (3.21). It can be clearly seen that the proposed SR image reconstruction is an iterative process. The estimated HR image at each iteration is compared to the previous one to make sure the difference from each estimated HR image converges under a certain pre-determined error criterion, ε. If the error of the estimated HR image is less than the preset error criterion ε, according to the major SR process steps shown in Figure 2.1, the reconstructed HR image is further de-blurred by Wiener filtering process to generate the final estimated HR image.

4.2   Wavelet Domain Convex Set for POCS

There are a larger number of data/signal transforms were invented by mathematicians. Among those transformation methods, "Fourier transform" is the most commonly used one in signal processing research. However, the wavelet transform has powerful analysis ability for data compression and signal recovery applications. This section introduces wavelet analysis and defines the wavelet convex set for the proposed wavelet-based POCS SR reconstruction approach.

4.2.1   Wavelet for Super-Resolution

For one-dimensional cases, "Fourier transform" converts time domain signals into the frequency domain. In contrast, the result of the wavelet transform is the product of the time domain signal and a wavelet function. After the transformation, the wavelet coefficients of a time domain signal are still in the time domain. In the case of two dimensional signal, the wavelet coefficient sub-images of an image can be treated as spatial domain data.

A major advantage of the wavelet analysis in image processing is the ability to perform

61

$$F(\omega) \;=\; \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt$$

$$C(scale,\ position) \;=\; \int_{-\infty}^{\infty} f(t)\psi(scale,\ position,\ t)\,dt$$
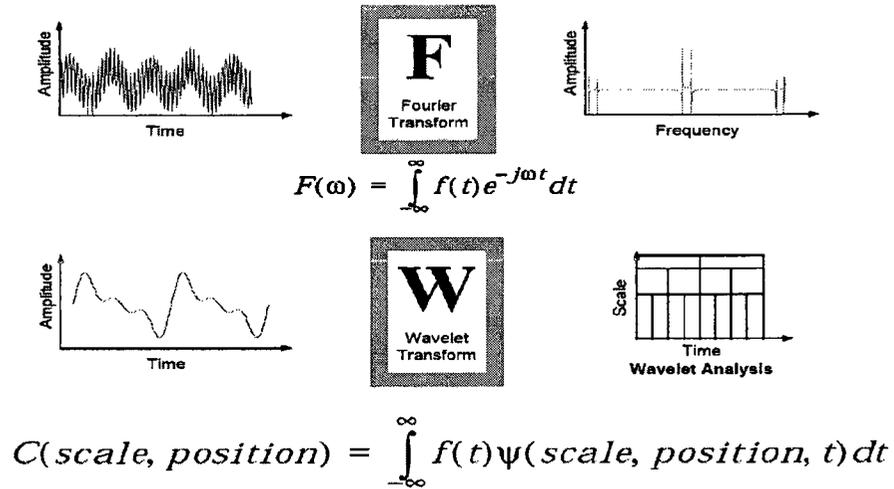
Figure 4.2: Comparison of Fourier transform and wavelet transform

local analysis; meaning that it can analyze a localized area of a larger image. wavelet analysis is capable of discovering the image piece wise regions, edges and textures that other signal analysis techniques cannot. A simple example is illustrated in Figure 4.3 where a one-dimensional sinusoid waveform with a very short period of discontinuity is separately transformed by Fourier and wavelet transform. The results clearly show that wavelet coefficient presents the discontinuity duration, while "Fourier transform" does not provide information about the discontinuity.

Sinusoid with a small discontinuity

Fourier Coefficients

Wavelet Coefficients
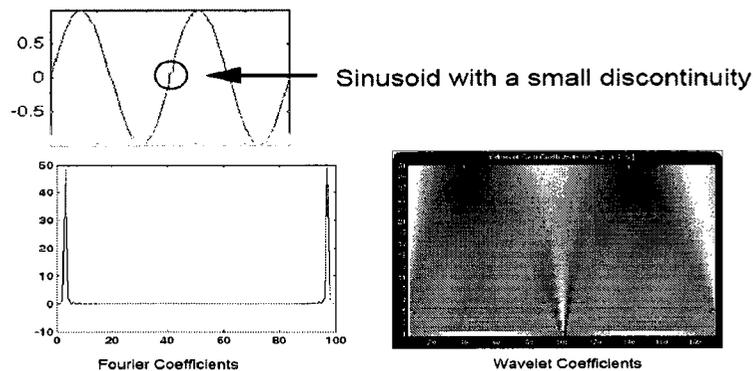
Figure 4.3: Sinusoid waveform transformed by Fourier and wavelet transforms

In order to perform the POCS process in the wavelet basis, it is essential to represent

an image in wavelet domain. Basically, wavelets are functions, in which their dilations and translations form a basis of $L^2(R)$. More specifically, there is a function $\psi(t)$, namely "mother wavelet", which defines a set of wavelet functions.

$$\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n) \tag{4.1}$$

There is a function $\phi(t)$, namely, "scaling function", which is defined as

$$\phi(t) = \sum_{k=0}^{N} \sqrt{2} h_k \phi(2t - k) \tag{4.2}$$

where h_k, $0 \le k \le N$ are coefficients of the low-pass filter. For orthonormal wavelet function, ψ and its dual pair $\overline{\psi}$ can be derived from Equation (4.1)

$$\psi(t) = \sum_{k=0}^{N} \sqrt{2} g_k \phi(2t - k) \tag{4.3}$$

$$\overline{\psi}(t) = \sum_{k=0}^{N} \sqrt{2\overline{g_k}} \phi(2t - k) \tag{4.4}$$

where g and $\overline{g}$ are the corresponding high-pass filter coefficients of ψ and $\overline{\psi}$, respectively.

In the 2-D case, the components of the wavelet approximation to an image, can be separated into one scaling function and three wavelet functions

$$\phi(t, s) = \phi(t)\phi(s) \tag{4.5}$$

$$\psi^h(t, s) = \psi(t)\phi(s) \tag{4.6}$$

$$\psi^v(t, s) = \phi(t)\psi(s) \tag{4.7}$$

$$\psi^d(t, s) = \psi(t)\psi(s) \tag{4.8}$$

where $\psi^h(t, s)$, $\psi^v(t, s)$ and $\psi^h(t, s)$ are the 2-D wavelet functions which retain the detailed

63

information of input image in horizontal, vertical and diagonal directions, respectively.

An HR image $\mathbf{x}(m_1, m_2)$ is represented in wavelet domain as a sum of the wavelet approximation component in the LL band and three detailed components in the LH, HL and HH bands:

$$
\mathbf{x}(m_1, m_2) = \sum_{k,l \in Z} a_{k,l} \overline{\phi}_{k,l}(m_1, m_2) + \sum_{k,l \in Z} b^h{}_{k,l} \overline{\psi}^h{}_{k,l}(m_1, m_2) +
$$
$$
\sum_{k,l \in Z} b^v{}_{k,l} \overline{\psi}^v{}_{k,l}(m_1, m_2) + \sum_{k,l \in Z} b^d{}_{k,l} \overline{\psi}^d{}_{k,l}(m_1, m_2) \qquad (4.9)
$$

where $\phi_{k,l}(m_1, m_2)$ is the (k, l) coarse scaling function; $\psi^h{}_{k,l}(m_1, m_2)$, $\psi^v{}_{k,l}(m_1, m_2)$ and $\psi^d{}_{k,l}(m_1, m_2)$ are wavelets at the next coarse level that retain the detailed information in the horizontal, vertical and diagonal directions, respectively. The top bar on each function denotes their corresponding inverse functions used to reconstruct the image back to spatial domain. The approximation component and the detailed wavelet coefficients are given by

$$
a_{k,l} = \sum_{m_1, m_2} x(m_1, m_2) \phi_{k,l}(m_1, m_2) \qquad (4.10)
$$

$$
b^h{}_{k,l} = \sum_{m_1, m_2} x(m_1, m_2) \psi^h{}_{k,l}(m_1, m_2) \qquad (4.11)
$$

$$
b^v{}_{k,l} = \sum_{m_1, m_2} x(m_1, m_2) \psi^v{}_{k,l}(m_1, m_2) \qquad (4.12)
$$

$$
b^d{}_{k,l} = \sum_{m_1, m_2} x(m_1, m_2) \psi^d{}_{k,l}(m_1, m_2) \qquad (4.13)
$$

In view of the above mathematical operations, we can treat the process of 2-D wavelet transformation as a filtering process. An image is filtered and split into four sub-images in wavelet domain. Figure 4.4 shows a block diagram of a one-level wavelet decomposition of an image.

Practically, the wavelet image decomposition can be further divided into several levels

64
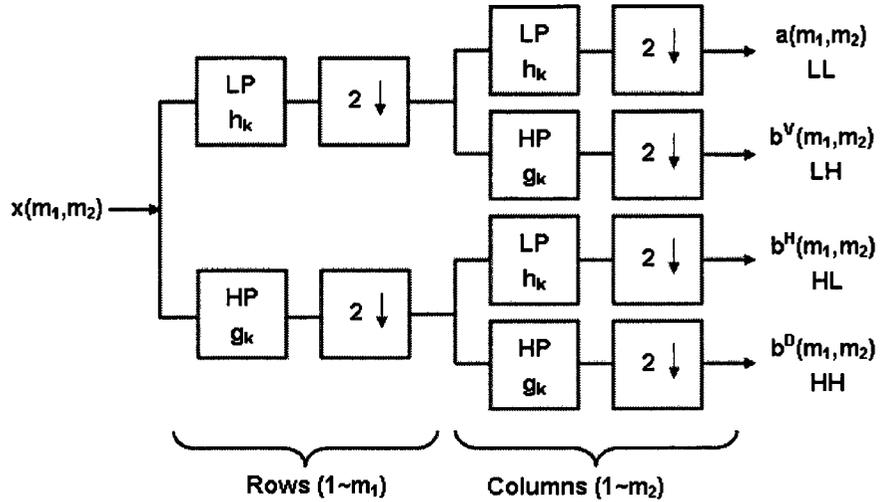
Figure 4.4: A filter bank block diagram for 2-D wavelet transform of image

of wavelet coefficient sub-images; the first level, wavelet coefficient sub-images have the same size as the original image. The second level refers to wavelet coefficient sub-images that are down-sampled by a factor of 2. Since, an SR image reconstruction procedure does not requare an image compression process, wavelet coefficients do not need to be down-sampled. Therefore, the wavelet transform used in the proposed algorithm is set on the first level.
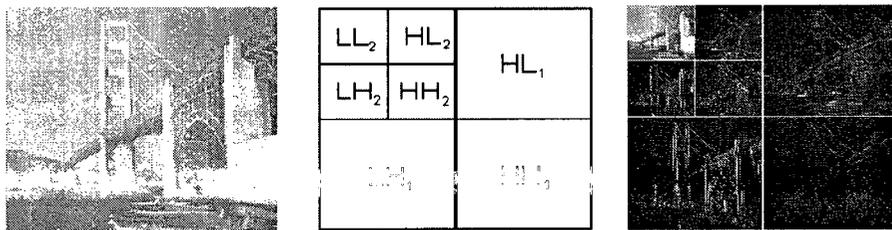
Figure 4.5: A two-level wavelet transform result of an image

4.2.2 Projections onto Wavelet Convex Set

A wavelet image decomposition described in Equation (4.9) contains both high and low frequency information separately. According to the idea of wavelet multi-resolution analysis (MRA), to meet the requirements of SR applications, the high frequency information preserved in the detailed coefficients should be extended. In contrast, the low frequency information contained in the approximation coefficients should be also preserved. In order to formulate the wavelet convex set, the spatial domain of the SR image reconstruction model described in Equation (1.2) is re-formulated by using the wavelet domain representation.

$$\mathbf{y}_i = S(a_{k,l}(m_1, m_2)) + \eta_i = S\left(\sum_{m_1, m_2} \mathbf{x}(m_1, m_2)\phi_{k,l}(m_1, m_2) \right) + \eta_i \qquad (4.14)$$

where $S()$ is a function that generates appropriate scaling factor and converts the wavelet domain coefficient to the spatial domain; the $a_{k,l}$ are the approximation coefficients of the desired HR image described in Equation (4.10). Therefore $S(a_{k,l}(m_1, m_2))$ represents the appropriate scaled approximation coefficients of the desired HR image which is made equivalent to LR frames. The degrading function $\mathbf{H}_i$ in Equation (1.2) is now represented by h_k shown in Equation (4.2). According to the SR image reconstruction model in Equation (4.14), the LR frames are obtained from the down scaled wavelet approximation coefficients of the idea HR image.

Based on the wavelet domain representation of the SR image reconstruction model, the convex set of the estimated HR image $\tilde{\mathbf{x}}(m_1, m_2)$ in POCS process is then defined as:

$$C_W = \left\{ \tilde{\mathbf{x}}^{n+1}(m_1, m_2) : \sum_{m_1, m_2} \tilde{\mathbf{x}}^n(m_1, m_2)\phi_{k,l}(m_1, m_2) = S(\mathbf{y}_i(n_1, n_2)) \right\} \qquad (4.15)$$

By simply applying Equation (3.15), it can be shown that set C_W is a convex set. The projection operation of the estimated HR image $\tilde{\mathbf{x}}(m_1, m_2)$ onto C_W is defined as [45]

66

$$\widetilde{\mathbf{x}}^{n+1}(m_1, m_2) = P_W[\widetilde{\mathbf{x}}^n(m_1, m_2)] = \sum_{k,l \in Z} S(\mathbf{y}_i(n_1, n_2))\overline{\widetilde{\phi}}_{k,l}(m_1, m_2)$$

$$+ \sum_{k,l \in Z} \widetilde{b}^h_{k,l}\overline{\widetilde{\psi}}^h_{k,l}(m_1, m_2)$$

$$+ \sum_{k,l \in Z} \widetilde{b}^v_{k,l}\overline{\widetilde{\psi}}^v_{k,l}(m_1, m_2)$$

$$+ \sum_{k,l \in Z} \widetilde{b}^d_{k,l}\overline{\widetilde{\psi}}^d_{k,l}(m_1, m_2) \qquad (4.16)$$

Note that the equivalently scaled LR frame $S(\mathbf{y}_i(n_1, n_2))$ instead of the LL band scaling coefficients $a_{k,l}(m_1, m_2)$ is used in the POCS iteration.

4.3 Stochastic Wavelet Coefficient Refinement Process

The stochastic wavelet coefficient refinement process is performed after the wavelet convex set projection operation. This section explains the wavelet coefficient refinement process for the purpose of eliminating noisy components, adjusting wrongly estimated pixels and refining the estimated HR image. This process is a variance dependent processing. A general framework of the wavelet coefficient refinement process is provided first. The process framework is depicted in Figure 4.6 and the procedure and image data flow is detailed step by step in the following sub-sections.

In Figure 4.6, a previously estimated HR image is first decomposed by wavelet transformation. The approximation component of the HR image contained in the LL band is retained without any modification. Secondly, the variance of each of the three high frequency sub-images is computed. Thirdly, three variances for LH,HL,HH sub-images are cross-referenced to determine whether the wavelet coefficient sub-images should be changed or preserved according to the stochastic F-distribution. Finally, the refined HR image is reconstructed from the modified three high frequency sub-images and the LL band approximation sub-image by the inverse wavelet transformation.

67
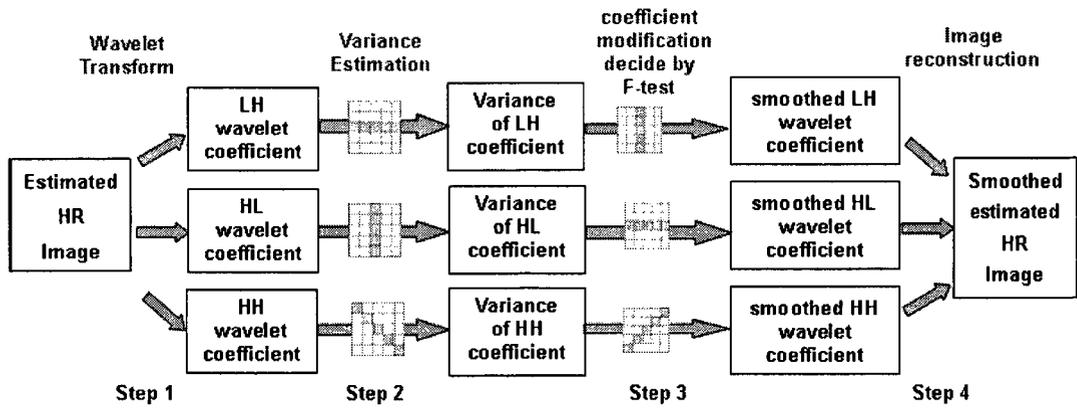
Figure 4.6: Framework of stochastic wavelet coefficient refinement process

4.3.1 Estimating the Variance of Wavelet Coefficients

Variance estimation is very important for the wavelet coefficient refinement process. Accurate variance estimation would yield better refinement and noise reduction results. This step estimates the variance of wavelet coefficient sub-images. The input of variance estimation is generated from step 1 wavelet decomposition of the HR image. It is shown as an $d \times d$ 2-D window on a wavelet coefficient sub-image where d is an odd integer in Figure 4.7. To estimate the variance of the coefficient at sub-image location (i,j), the largest and smallest coefficient in the window is first eliminated. In this case, coefficient values 20 and 0 are eliminated. The variance at location (i,j) is then calculated from the remaining coefficients within $d \times d$ window using Equation (4.17).
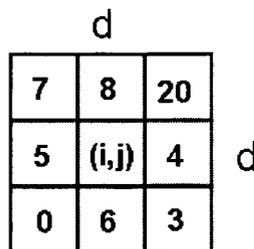
Figure 4.7: 2-D variance calculation window

68

$$\sigma^2{}_{i,j} = \frac{1}{d^2 - 2} \sum_{l=2}^{d^2-1} [W^l_{(i,j)}]^2 \qquad (4.17)$$

In Equation (4.17), d is the window size , $W_{(i,j)}$ is the wavelet coefficient at the location(i.j). The summation counter l starts at "2", since the largest and smallest coefficients have already been eliminated. In practice, a one-dimensional window is used instead of a two-dimensional window; therefore this operation is processed only in one dimension. Note that for an one dimensional window, the window size d should be an odd integer and at least 5 because two wavelet coefficients have to be eliminated before calculating the variance. The windows corresponding to each wavelet coefficient sub-image is shown in Figure 4.8.
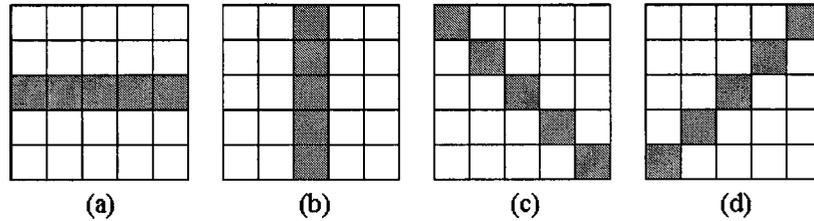
Figure 4.8: 1-D windows for variance estimation (a) LH wavelet coefficient sub-image, (b) HL sub-image, (c) and (d) HH sub-image

The result of variance estimation of three coefficient sub-images is called "variance sub-images". For example, the LH variance sub-image contains horizontal variances which are estimated from the LH wavelet coefficient sub-image. All variance sub-images are prepared for the next step to enhance the noisy wavelet coefficient sub-images by applying an appropriate stochastic threshold determination.

4.3.2   F-Distribution

The F-distribution is a continuous probability distribution and it is used frequently as the null distribution of a test statistic, especially in the analysis of variance. A random

variable which has F-distribution can also be seen as the ratio of two Chi-squared random variables [46].

$$\text{F-distribution random variable} = \frac{C_1/d_1}{C_2/d_2} \qquad (4.18)$$

In Equation (4.18), C_1 and C_2 have Chi-square distributions with d_1 and d_2 degree of freedom, respectively. The cumulative distribution function (CDF) curve of the F-distribution is determined by degrees of freedom d_1 and d_2. Figure 4.9 shows the CDF curves, and Equation (4.19) describes the CDF of the F-distribution.
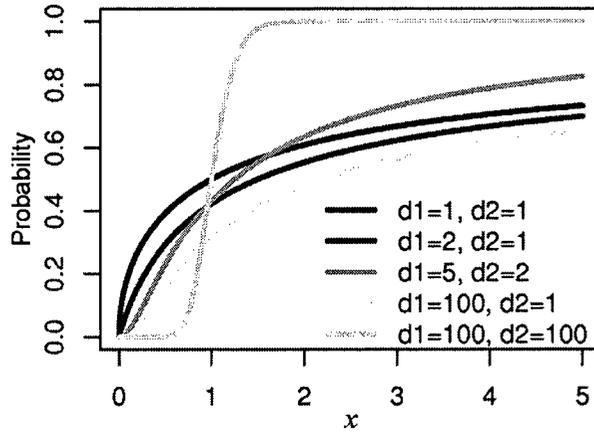
Figure 4.9: Cumulative distribution function of F-distribution

$$G(x) = I_{\frac{d_1 x}{d_1 x + d_2}}\left(\frac{d_1}{2}, \frac{d_2}{2}\right) \qquad (4.19)$$

where $G(x)$ is the CDF of the F-distribution and I is the regularized incomplete beta function.

The reason for introducing F-distribution is to impose the F-test as a tool to determine a critical value which serves as a threshold to examine the variance of wavelet coefficients. Before explaining the F-test, some assumptions are made to model the outcome from each

70

step. Firstly, suppose that each wavelet coefficient $W_{i,j}$ generated from the wavelet decomposition has Gaussian distribution with zero mean and variance $\sigma^2{}_{i,j}$, although some researches declared that the wavelet coefficient follows the "generalized Gaussian distribution". however, the reason is that one could use regular Gaussian distribution to approximate the generalized Gaussian distribution. Therefore, we still can declare that the variance of wavelet coefficients estimated by taking its local neighborhood within a 1-D window follows a Chi-square distribution. From the probability theory, we know that the F-distribution is formed by a ratio of two Chi-square distribution random variables. Therefore, the ratio of two variance coefficients R from the variance sub-images will follow the F-distribution. The stochastic assumptions described above can be rewritten by the following notations for better understanding.

$$W_{i,j} \sim N(0, \sigma^2{}_{i,j}): \text{Wavelet coefficient follows the Gaussian distribution} \qquad (4.20)$$

$$\widetilde{\sigma}^2_{i,j} \sim X^2(d^2): \text{Variance of wavelet coefficient follows the Chi-square distribution}$$
$$(4.21)$$

$$R = \frac{\widetilde{\sigma}^2_{k,l}}{\widetilde{\sigma}^2_{i,j}} \sim F(d^2, d^2): \text{Ratio of two variance coefficients follows the F-distribution}$$
$$(4.22)$$

In the F-Test and the coefficient modification step, the F-Test results in a critical value C which is determined by two parameters, here we use d^2 for both parameters. Thus, d in Equation (4.22) is the size of the window for variance evaluation. The critical value C will be detailed in Section 4.3.3.

4.3.3 F-Test and Coefficient Modification

Since the ratio of two variances of wavelet coefficients follows the F-distribution, the very useful variance analysis technique, it can be imposed to determine an appropriate threshold for the wavelet coefficient refinement process. The value of the ratio of two variance coefficients R represents the difference of any two neighboring variances caused by the randomness of the wavelet coefficient or caused by different image feature distribution. If the value of R is close to unity, it means that the two variance coefficients are more likely to be in the same smooth region. Otherwise, the variance coefficient reflects an image feature. However, only considering the value R is not enough for us to distinguish the features from the noisy image. We need to employ the F-test to verify whether any two variance coefficients are statistically different.

In order to distinguish the image features and preserve edges, the F-test processing also uses the 1-D window, which is orthogonal with respect to the window used in step 2. Hence, a vertical window is used to apply the F-test on the LH variance sub-image; a horizontal window is used to apply the F-test on the HL variance sub-image and the window for HH F-testing should use another orthogonal diagonal direction. The windows used for F-test are shown in Figure 4.10 which can be compared to Figure 4.8 for better understanding.
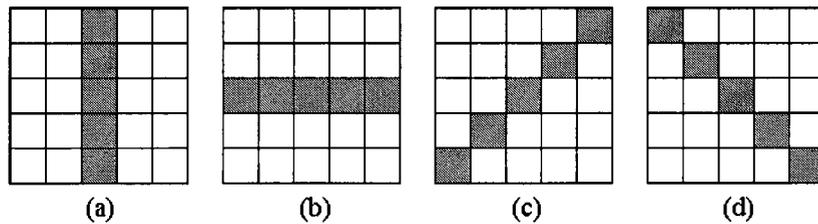
Figure 4.10: 1-D windows for F-Test. (a) LH variance sub-image, (b) HL variance sub-image, (c) and (d) HH variance sub-image

The detailed procedure of the F-test is described as follows: first, the center variance coefficient is compared to the maximum of its two touching neighbors. If the comparison according to the F-distribution indicates that variance coefficients are statistically different

then the corresponding feature is preserved. Otherwise, the center variance coefficient needs to be compared again with the minimum of the median of variance coefficients in the two neighboring sub-windows and checked again. If the F-test indicates that the center variance coefficient is different from this minimum of medians, then the corresponding feature is preserved; otherwise that feature should be modified. Figure 4.11 gives a graphic representation of the above explanation.
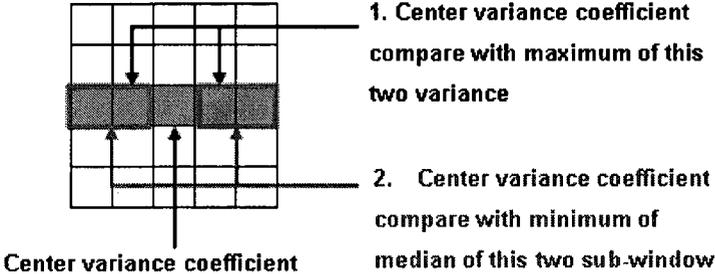
Figure 4.11: Graphical presentation of F-test comparison

In practice, the critical value C of the F-test function provided in Matlab needs three parameters to determine whether another comparison is needed or not. The first parameter is the level of confidence γ which is a real value between zero and one $0 < \gamma \leq 1$. Any γ determines an F-test look-up table according to F-distribution CDF. The other parameter is the degree of freedom d which is the size of the window in this refinement process, as previously mentioned. The mathematical notation for the F-test can be written as:

$$\text{Critical Value } C = F(\gamma, d^2, d^2)$$

or

$$\text{Critical Value } C = F_\gamma(d^2, d^2) \tag{4.23}$$

By comparing the ratio of two variances R, with the critical value C, it can be determined whether the wavelet coefficient is mainly contributed by the noise or image feature. If the comparison indicates that the variance is due to noise, the corresponding wavelet coefficient is modified by multiplying an attenuating factor α. The following paragraphs explain in detail the wavelet coefficient modification scheme, the purpose of attenuating the noisy wavelet coefficients, and correcting wrongly estimated pixels.

The third step of the stochastic refinement process framework in Figure 4.6 modifies the noisy and wrongly estimated wavelet coefficients after the F-Test identifies which wavelet coefficient should be suppressed. According to [46], the wavelet coefficient modification scheme namely "variance dependent attenuation" uses the attenuating factor α. This attenuating factor α is set to be adaptive, and its value is determined by Equation (4.24) and Equation (4.25).

$$\widetilde{W}_{i,j} = \alpha W_{i,j}; \begin{cases} \alpha = 1, & \text{if } R \leq F_{\gamma_1}(d,d) \\ \alpha = \beta, & \text{if } R > F_{\gamma_2}(d,d) \\ \alpha = f(F_{\gamma_1}, F_{\gamma_2}), & \text{otherwise} \end{cases} \qquad (4.24)$$

where

$$f(F_{\gamma_1}, F_{\gamma_2}) = \beta + (1-\beta)\left\{ \frac{1 - e^{\xi(\frac{R-F_{\gamma_1}}{F_{\gamma_1}-F_{\gamma_2}})}}{1 - e^{\xi}} \right\} \qquad (4.25)$$

Note that in the wavelet coefficient modification scheme in Equation (4.24), two different levels of confidence γ_1 and γ_2 are used. Also the degree of freedom here is d, not d^2, β is a predetermined attenuating factor which modifies wavelet coefficients when the F-test indicates the variance coefficient is due to the noise. Typically, β is set to be 0.5 but it can be adjusted to obtain a better PSNR compared to the original HR image. Two parameters of the function f are critical values obtained from F-test Equation (4.23) by choosing two different values for the levels of confidence γ, ξ is an integer weighting factor between the -5 and +5. Function f provides more weights toward preserving features when ξ=-5 ;in

74

contrast, function f provides more weights toward smoothing wavelet coefficients when $\xi=+5$.

This variance dependent attenuating scheme divides the attenuating factor α into three threshold regions. Figure 4.12 represents a better visual description to the three threshold regions of attenuating factor α where the horizontal axis is the ratio R of two variances.

Wavelet coefficient preserved	Wavelet coefficient adjusted by adaptive factor	Wavelet coefficient adjusted by fixed factor
$\alpha = 1$	$\alpha = f(F_{\gamma_1}, F_{\gamma_2})$	$\alpha = \beta$

$$F_{\gamma_1}(d,d) \qquad F_{\gamma_2}(d,d) \qquad R$$
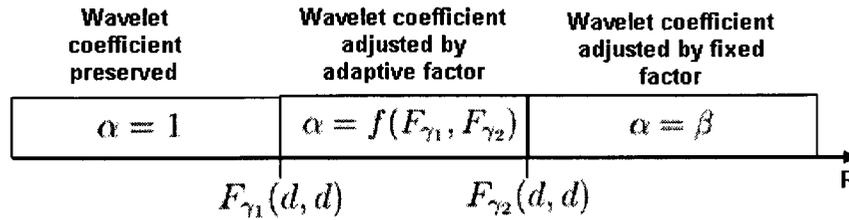
Figure 4.12: Three threshold regions of attenuating factor α

Examining the ratio of the wavelet coefficient variance $\tilde{\sigma}_{k,l}^2$ and $\tilde{\sigma}_{i,j}^2$, R by providing 2 different levels of confidence γ_1 and γ_2 to the F-test function, the value of R can be separated into 3 different regions. The value of R is located at the region in the most left hand side of Figure 4.12 meaning that the wavelet coefficient should be kept unchanged; thus, the attenuating factor α is set to be 1. Then, if the value of R is located at the region between critical value $F_{\gamma_1}(d,d)$ and $F_{\gamma_2}(d,d)$, it means that the wavelet coefficients need to be adaptively attenuated by the factor α determined by Equation (4.25). Moreover, if the value of R is located at the region greater than the critical value $F_{\gamma_2}(d,d)$, the attenuating factor α is set to a pre-determined attenuating factor.

To summarize this stochastic refinement process (please refer to Figure 4.6), firstly, the estimated HR image from the wavelet projection process is decomposed by wavelet transformation into LL, LH, HL, and HH bands. Secondly, the LL band approximation data is kept unchanged. The Variance coefficients of LH band wavelet coefficient sub-image are calculated by using 1-D window as shown in Figure 4.8(a). Thirdly, F-Test is performed to

examine LH band variance coefficients by using the orthogonal window as shown in Figure 4.10(a). According to the the result of the F-test, the coefficients of LH band wavelet coefficient sub-image is modified by using Equation (4.24). The same procedures apply on HL, and HH band wavelet coefficient sub-images with corresponding variance estimation window and F-Test examination window to obtain the smoothed version of LH, HL, and HH wavelet coefficient sub-images. Finally, The refined HR image is obtained by an inverse wavelet transformation from the LL band and the three modified LH, HL, and HH band sub-images.

4.4 The Wavelet-Based POCS Iterative Algorithm

The overall procedure of the proposed wavelet-based POCS plus stochastic wavelet coefficient refinement process for SR image reconstruction is explained as follows.

By referencing Figure 4.13, the process starts from the top of the flow chart, and the acquired LR frames are input after the simulation program is initialized. The following gives the step-by-step procedure.

- Step 1. Image registration: Firstly, a LR frame is set as reference frame and all other consecutive LR frames are compared to the reference frame. The rotation and motion estimation algorithm discussed in Section 2.1 is applied to obtain global motion vector of each LR frame. According to the motion vector of the LR frames, the algorithm aligns all the LR frames to the reference frame. Then the aligned LR frames are passed to the next step.

- Step 2. Degrading function estimation: The degrading function is usually an image blurring effect on each LR frame which is estimated in this step. The blind blur identification algorithm discussed in Section 2.2 is used to find the kernel of the blurring filter B of each LR frame. The blurring filter B is one of the degrading

76
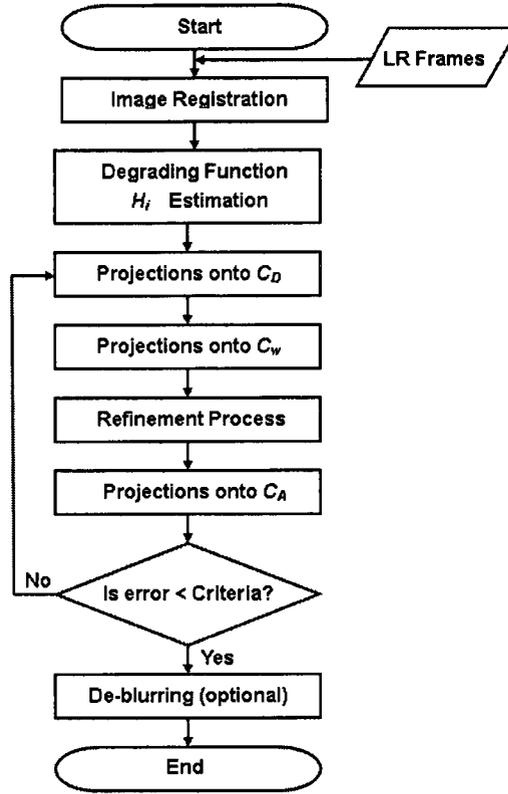
Figure 4.13: Proposed wavelet-Based POCS SR reconstruction procedure

components needed to form the complete degrading function H_i defined in Equation (1.2). This Equation (3.19) will be used in the first POCS convex set projection.

- Step 3. Projections on to C_D: The iteration of the proposed wavelet-based POCS starts at this point. The first LR frame is scaled by down-sampling factor μ. The bilinear interpolation method is used as the first initial guess of HR source image $x^0(m_1, m_2)$ and then the residual $r^{(x)}(n_1, n_2)$ is computed according to Equation (3.18). Followed by the data consistency constraint C_D, the projection operation is then performed to update the estimated HR image $\tilde{x}$ according to Equation (3.19). Residual calculation and C_D projection operation is repeated until all LR frames are used. The estimated HR image is then passed to the wavelet convex set projection.

- Step 4. Projections on to C_W: This step performs wavelet convex set projection. The estimated HR image from Step 3 is decomposed by wavelet transform. Then the wavelet convex set projection P_W is performed by replacing scaling coefficients with LR frames $\mathbf{y}_i(m_1, m_2)$ as given in Equation (4.16).

- Step 5. Wavelet sub-image coefficient refinement process: This step refines wavelet sub-image coefficients by applying stochastic F-distribution. Four wavelet coefficient sub-images of an estimated HR image received from Step 4 is refined according to the proposed algorithm detailed in Section 4.3 (See refinement process framework in Figure 4.6).

- Step 6. Projections on to C_A: The pixel value of an estimated HR image $\tilde{\mathbf{x}}$ from Step 5 is then restricted by the amplitude constraint C_A according to Equation (3.21). After performing at least three times of POCS iterations, the estimated HR image from the previous iteration is then subtracted from the current estimated HR image to check the convergence, and the difference of these two estimated HR images is then compared to the previous difference in the sense of l_1 norm by a pre-determined convergence criterion (typical value 10^{-4}). If the difference of two estimated HR images is greater than the pre-determined criterion, then the current estimated HR image is passed on to Step 3 as a new initial HR image and the POCS process is repeated until the convergence criterion is achieved.

- Step 7. De-blurring process: This step performs the De-blurring process to remove blurring effects either originally acquired from LR frames, or caused by the artifacts from previous steps. The de-blurring method introduced in Section 2.4 is used. After the de-blurring process, the final SR image is reconstructed.

4.5 Conclusion

The proposed SR image reconstruction algorithm is explained in detailed in this chapter. The framework of the proposed algorithm is firstly addressed by extending the conventional POCS SR reconstruction algorithm. The properties of the wavelet transformation and the basics of wavelet analysis are briefly discussed. The new idea, wavelet convex set is derived by replacing the LL band approximation coefficient with properly scaled and accumulated LF frames. The procedure of stochastic wavelet coefficient refinement process is also explained from the coefficient variance calculation to the F-distribution derivation and the wavelet coefficient attenuation scheme. An overall process of the proposed SR image reconstruction algorithm has been step-by-step explained.

Chapter 5

Experimental Results

Several experiments have been done to show the performance of the proposed wavelet-based POCS SR reconstruction algorithm. This chapter presents the complete experimental and simulation work. The setup of the simulation environment is provided in Section 5.1, in which the degrading procedure of the LR frames, the sample HR images used for simulation and the methods for the objective performance evaluation are presented. Some intermediate results of the proposed SR algorithm are provided in Section 5.2 to show how the frame data is manipulated. In the Section 5.3, experimental results of the proposed SR algorithm and some other SR approaches with respect to different types of image are compared subjectively. In Section 5.4, numerical experimental results such as the PSNR measurement and the convergence speed of the proposed SR algorithm and other SR approaches are compared objectively. Later in this section, the HR images reconstructed by the proposed method with different parameter setups are presented.

5.1 Simulation Environment

5.1.1 Simulation Setup

Hardware and Software Setup:

- Hardware: Desktop PC, Intel CPU 1.8GHz, 2GB RAM

- Software: Microsoft Windows XP professional SP2, Mathworks Matlab Version: 2007b

Generating the degraded LR frames:

In the experiments, a separate program is developed to generate the degraded LR frames. Several LR frames were obtained by degrading the original HR images according to Equation (1.1). A graphical presentation of the degrading procedure is shown in Figure 5.1.

In Figure 5.1, the HR image is used as the desired HR image. Ten LR frames are generated from the corresponding HR image. The original HR image is firstly blurred by a Gaussian low-pass filter with size 4×4 and a standard deviation of 1. The nine randomly generated motion vectors are applied to all the frames except the first one. The motion vector does not apply on the first frame in order to use it as a reference frame for the image registration process. These 10 degraded frames were then down-sampled in both the horizontal and vertical directions with a factor of 4, and in the end a Gaussian noise was then added on these LR frames.

5.1.2 Performance Evaluation

To evaluate the performance of the proposed algorithm for SR image reconstruction algorithm, the HR image resulting from the proposed method are simulated and compared subjectively and objectively with those obtained from the bilinear interpolation, T. Tsai
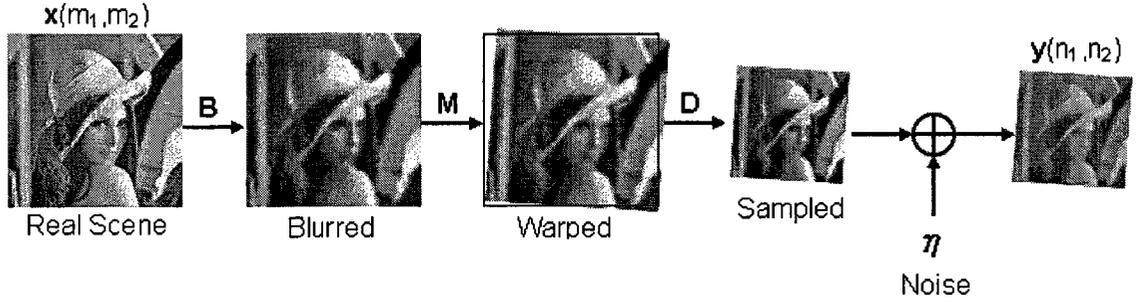
Figure 5.1: Degrading procedure of LR frames

and R. Huang's frequency domain method [8], M. Irani's IBP method [9] and A. Tekalp's conventional POCS [3].

The parameters selected in each experiment for different algorithms were chosen to produce the most visually appealing results. For a fair evaluation, the algorithms in companion are performed several times with different parameters in order to obtain the best result of each algorithm for the same experiment.

Method for Objective Comparison:

The 2-D peak signal to noise ratio (PSNR) defined in Equation (5.1) is used for objective comparison.

$$ \text{PSNR} = 10\log\left\{ \frac{255^2}{\|\tilde{x} - x\|^2} \right\} \quad \text{(dB)} \tag{5.1} $$

where $\tilde{x}$ is the estimated HR image, x is the original HR image.

The final estimated HR images may appear as black pixels on each side of the estimated HR image edges, due to the motion compensation process filling up the unknown pixels with black pixels on each degraded LR frames. In order to obtain fair PSNR measurements, the pixels close to the image edges are trimmed off, therefore only the center
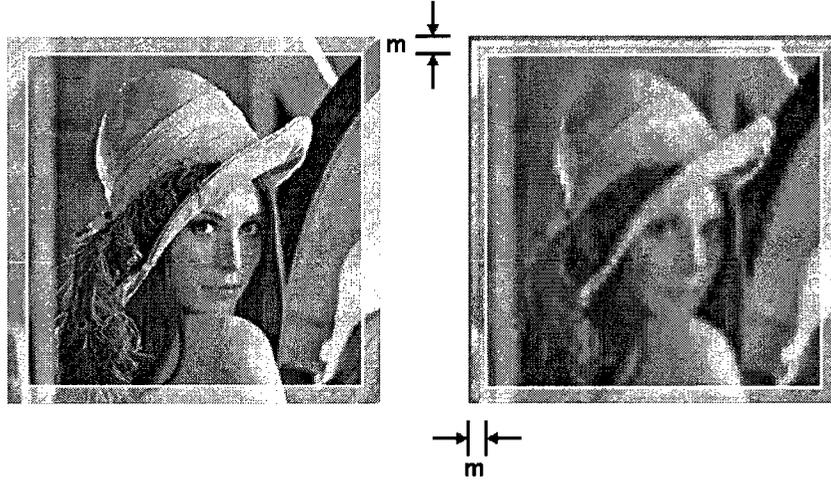
Figure 5.2: Image edges removal manner of PSNR measurement

part of the original HR image and the estimated HR image are taken to the PSNR calculation. Figure 5.2 presents the manner of PSNR measurement and it shows that pixels outside of the square frame are not included for PSNR calculation. The number of pixels **m** to be trimmed off is determined by the biggest displacement of the inter-frame motion.

Method for the Convergence Comparison:

During the of iterative process of the SR reconstruction algorithm, the HR image obtained from the previous iteration is subtracted from the current estimated HR image to check the convergence. The difference of each two consecutive estimated HR images is compared in the sense of l_1 norm by a pre-determined convergence criterion ε, as shown in the following equation.

$$\varepsilon = \frac{\|\widetilde{\mathbf{x}}^n - \widetilde{\mathbf{x}}^{n-1}\|}{\|\widetilde{\mathbf{x}}^n\|} \tag{5.2}$$

where $\widetilde{\mathbf{x}}^n$ is the HR image estimated at the current iteration and $\widetilde{\mathbf{x}}^{n-1}$ is the HR image from

83
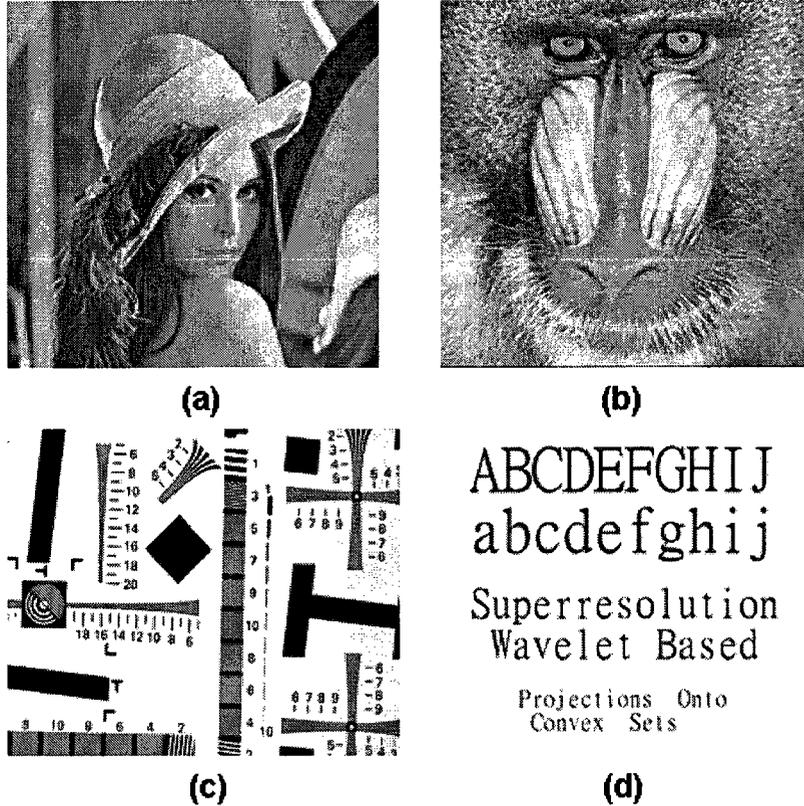
Figure 5.3: The HR images for experiment (a) Lena, (b) Baboon, (c) Chart, (d) Letters.

the last iteration.

A typical value of the convergence criterion ε is set to 10^{-4}. The iteration stops when the difference of two consecutive estimated HR images is less than 10^{-4}. The improvement is considered not significant to the human visual system if the difference of the consecutive estimated HR images is less than the convergence criterion.

5.1.3 Sample Images for Experiments

For the experimental study of the proposed SR algorithm, four typical 256-gray-level HR images, i.e., Lena, Baboon, Chart, and Letters, are selected to show both subjective and objective performance evaluation.

84

The four chosen HR images contain different characteristics of image features. The image Lena represents the typical natural real scene which contains a mix of low, medium and high frequency image components. The image Baboon represents a scene that mostly contains high frequency components. The image Chart represents the pictures that contain artificial textures and small font words. The image Letters represent the images with sharp contour objects and strong contrastive transitions.

5.2 Intermediate Results of Proposed SR Reconstruction Algorithm

This section provides the intermediate results of two important processes of the proposed SR reconstruction algorithm. These two important processes are the projections onto wavelet constraint set process and the stochastic wavelet coefficient refinement process. The intermediate results of these two processes are shown individually and the visual results will show their effectiveness.

5.2.1 Intermediate Result of Projections onto Wavelet Constraint Set

This sub-section presents the intermediate result of the projections onto the wavelet constraint set which was explained in Section 4.2. The type of the chosen wavelet transformation is introduced and the demonstration of this projection process will be visually presented.

There are many different types of wavelet functions such as "Haar", "Daubechies" and "Biorthogonal" wavelet. Among those wavelets, the Daubechies wavelet function is employed for the decomposition of the estimated HR image. The reason is that the "compactly supported orthogonal wavelets" invented by Dr. Ingrid Daubechies makes discrete wavelet

analysis practicable. Therefore, Daubechies wavelet is much appropriate for image processing. The names of the Daubechies family wavelets are written as dbN, where N is the order of wavelet function, and db is the "nickname" of the wavelet. The proposed algorithm mainly uses db2 or db3 of the Daubechies wavelet function.

Four wavelet coefficient sub-images will be obtained by decomposing an estimated HR image by the Daubechies db2 wavelet. The LL band approximation is one of the four wavelet coefficient sub-images which will be replaced by the scaled acquired LR frames in this projection process. The mathematical expression of this process was previously given in Equation (4.16). The other three sub-images, the LH, HL and HH band wavelet coefficients, are unchanged during this wavelet constraint set projection process. A visual example of projections onto wavelet constraint set process is shown in Figure 5.4.

In Figure 5.4, the image on the left hand side is the temporary HR image estimated by projections onto data consistency constraint C_D. By carefully checking this image, there is a grid-ish horizontal and vertical artifact present. This particular type of noise is not an additive noise, and it can not be easily removed by simple noise reduction methods. The images on the right hand side are the wavelet components decomposed by Daubechies db2 wavelet from the temporary estimated HR image. From the top to the bottom, they are the LL approximation and the LH, HL and HH wavelet coefficients. The LL band approximation image, which contains low frequency components of the original scene, will be replaced by the scaled and accumulated acquired LR frames. The non-redundant information contained in each different LR frame contributes certain useful features to the new estimated HR image. This process is done when all available LR frames are used, then the four wavelet components are forwarded to the stochastic wavelet coefficient refinement process.
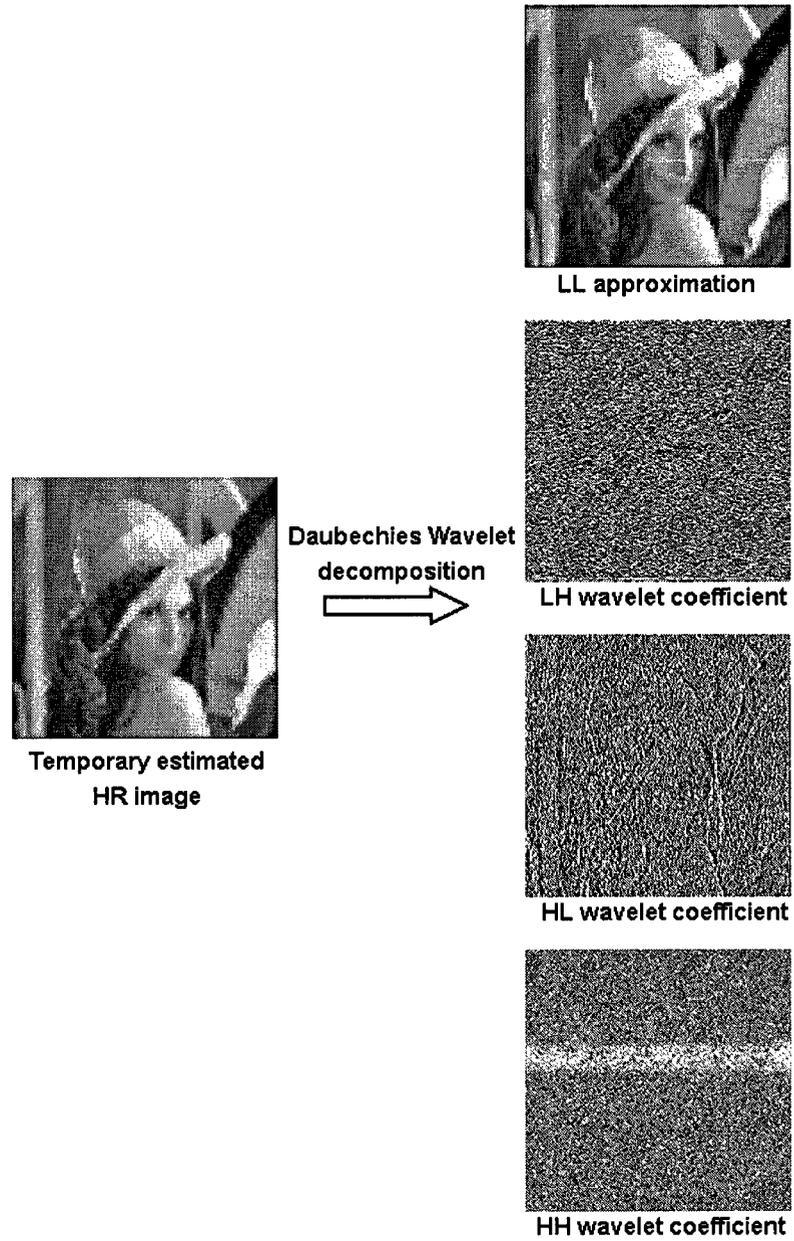
Figure 5.4: A visual example of projections onto wavelet constraint set process
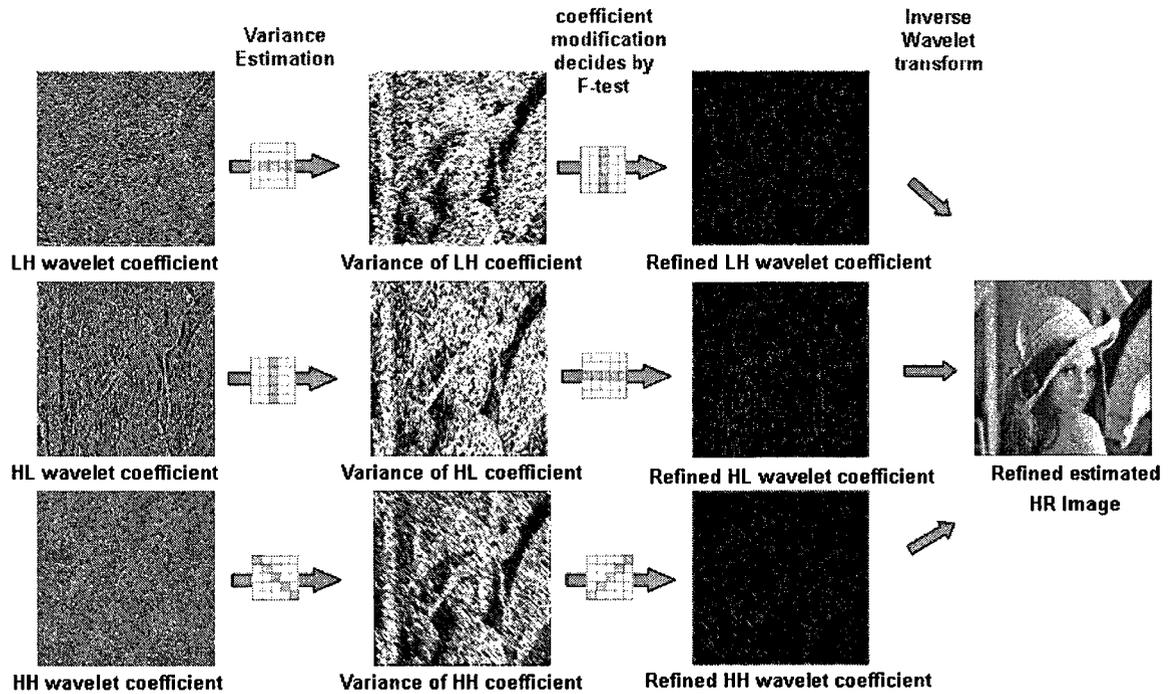
Figure 5.5: Visual demonstration of stochastic wavelet coefficients refinement process

5.2.2 Intermediate Result of Stochastic Wavelet Coefficient Refinement Process

This sub-section provides the demonstration of the stochastic wavelet coefficient refinement process which was explained in Section 4.3. The outcome of each step of this refinement process is shown in sequential order.

In this stochastic wavelet coefficient refinement process, there are four wavelet coefficient sub-images received from the previous process. They are the LL band approximation sub-image and the LH, HL and HH band wavelet sub-image coefficients. In this process, the LL band sub-image remains unchanged. The other three wavelet coefficient sub-images, are modified according to the stochastic wavelet coefficient refinement algorithm. The visual demonstration of this refinement process is presented in Figure 5.5.

Figure 5.5 contains four columns of images and the images on the most left hand side

are considered as the first column. The images on the second column are the variances of the wavelet coefficient sub-images which were calculated according to the corresponding 1-D window on their left hand side. The images on the third column are the refined wavelet coefficient sub-images. The refinement process modifies the wavelet coefficient sub-images regarding the F-test thresholds applied on their coefficient variance on the second column. The refined LH, HL and HH band wavelet coefficient sub-images and the LL band approximation, which is not shown in Figure 5.5, are then synthesized by the inverse wavelet conversion to reconstruct the estimated HR image. The comparison of the temporary estimated HR image shown in Figure 5.4 and the refined HR image shown in Figure 5.5, show that the horizontal and vertical grid-ish artifact appearing on the temporary estimated HR image has been removed by the refinement process.

According to the wavelet coefficient attenuation scheme described in Equation (4.24) and Equation (4.25), the simulation parameters for this coefficient refinement process is given as

Wavelet Type=$db2$: Daubechies wavelet type for wavelet transform.

$d_1 = 5$: window size for wavelet coefficient variance estimation.

$d_2 = 7$: window size for F-distribution test.

$\gamma_1 = 0.01$: level of confidence 1.

$\gamma_2 = 0.05$: level of confidence 2.

$\beta = 0.5$: predetermined attenuating factor.

$\xi = 5$: integer weighting factor.

5.3 Subjective Experimental Results

This section presents the simulation work by evaluating four different types of still images. The HR images reconstructed by Bilinear interpolation, frequency domain approach, IBP approach, conventional POCS approach and the proposed wavelet-based POCS plus wavelet coefficient refinement SR reconstruction algorithm are presented subjectively.

Figure 5.6 shows the SR reconstruction results based on the proposed algorithm as well as various existing algorithms. Figure 5.6(a) shows the original HR image Lena of size 256×256. Figure 5.6(b) shows one of the LR frames of size 64×64. There are 10 LR frames in total, which are generated by degrading the HR image with 9 randomly generated motion vectors and blurred by a Gaussian low-pass filter with size 4×4 and a standard deviation of 1. These 10 degraded images were then down-sampled both horizontally and vertically with a factor of 4. Gaussian noise was then added to these LR frames. In order to have a better visual comparison, the LR frame show in Figure 5.6(b) is magnified by nearest-neighborhood interpolation. In the actual SR reconstruction process however, the LR frames are in the size of 64×64. Figure 5.6(c) shows the SR image reconstructed by using the bilinear interpolation. Figure 5.6(d) gives the result obtained from the frequency domain method. Figure 5.6(e) shows the result obtained by using the IBP method. Figure 5.6(f) is from the conventional POCS and Figure 5.6(g) shows the SR image reconstructed by the proposed wavelet-based POCS method. The HR images reconstructed by different SR approaches are all of the same size as the original HR image.

In Figure 5.6,it is clearly seen that the proposed SR reconstruction presents more detailed features such as "lips and feathers" textures than other SR reconstruction approaches.

The features contained in HR image Lena are mixed with middle and low frequency components. Figure 5.7 presents the results of image Baboon which contain mostly high frequency texture components. Figure 5.7(a) shows the original HR image Baboon with

(a) Original (256x256)

(b) One of the LR frames (64x64)

(c) Bilinear Interpolation

(d) Frequency Domain

(e) IBP

(f) Conventional POCS

(g) Proposed
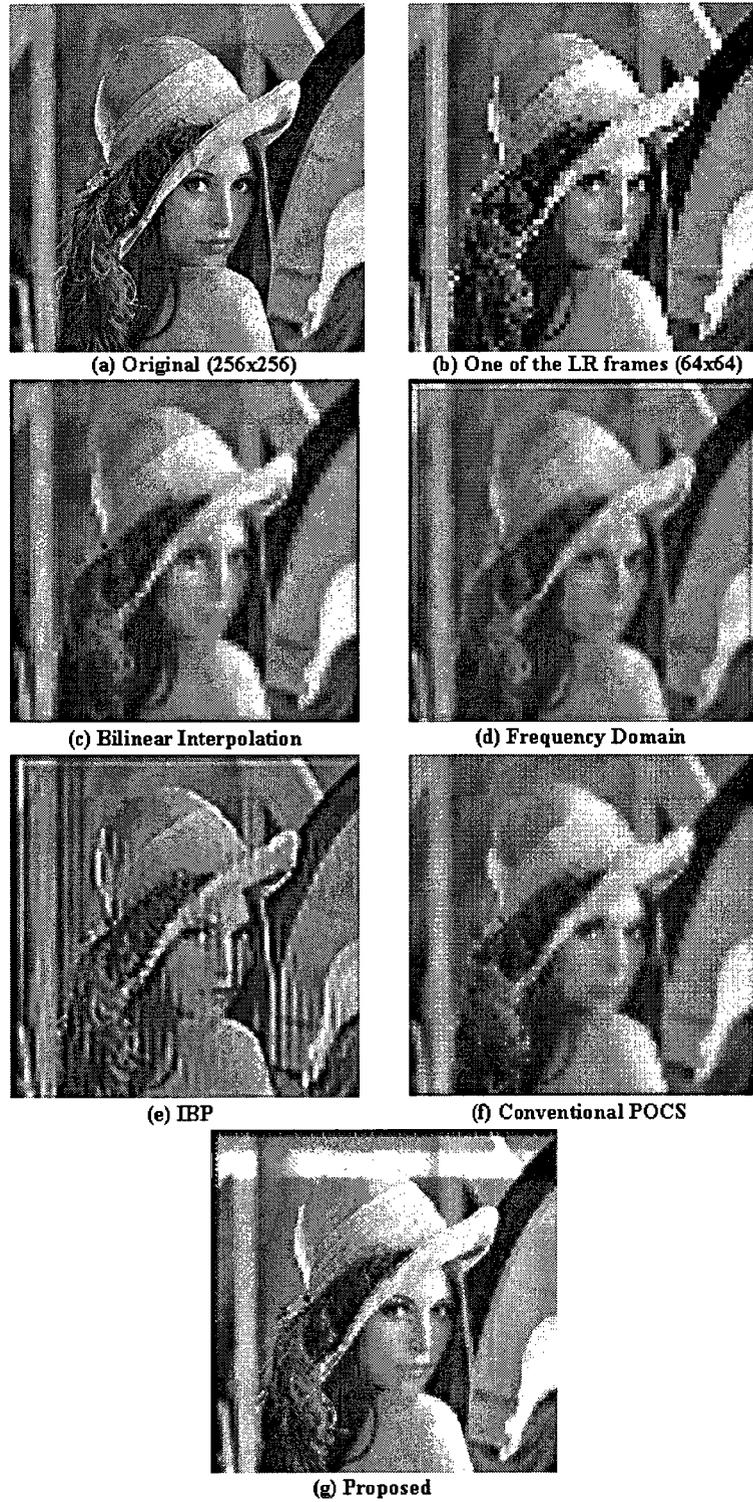
Figure 5.6: Super-resolution HR image reconstruction: Lena

91

size 256×256. Figure 5.7(b) shows one of the LR frames with size 64×64. Figure 5.7(c)-(g) show the SR images reconstructed by using the Bilinear interpolation, the frequency domain method, the IBP method, the conventional POCS and the proposed wavelet-based POCS method, respectively. Again, the estimated HR images reconstructed by different SR approaches all have the same size as the original Baboon HR image. It is seen that the HR image reconstructed by the frequency domain approach has fairly good subjective results, but the proposed SR reconstruction process shows improvements on both image contrasty and the resolution.

The HR image Lena and Baboon are natural images. Normally, the human eye has similar perception to natural images. Therefore, an artificial HR image is used as a different type of image for comparison. In the following SR simulation, the HR image Chart is employed. This image contains small font size numbers and different texture charts. Thus, a different visual perception is provided by this type of image. Figure 5.8(a) shows the original HR image Chart which contains artificial textures of size 256×320. Figure 5.8(b) shows one of the LR frames with size 64×80. Figure 5.8(c) shows the SR image reconstructed by using the Bilinear interpolation. Figure 5.8(d)-(g) show the results obtained from the frequency domain method, the IBP method, the conventional POCS and the proposed wavelet-based POCS approach, respectively. The HR images reconstructed by different SR approaches all have the same size as the original HR image.

From Figure 5.8, the IBP approach seen to have better than the frequency domain and the conventional POCS approaches. The HR image reconstructed by the IBP approach presents more information on numbers such as character "2" in the middle bottom and shows more textures on the vertical chart in the middle of the image. But overall, the HR image reconstructed by the proposed algorithm has a better quality.

An image with only letters is also used to test the performance of the different SR reconstruction approaches on a class of images that have sharp transition edges. In the

(a) Original (256x256)

(b) One of the LR frames (64x64)

(c) Bilinear Interpolation

(d) Frequency Domain

(e) IBP

(f) Conventional POCS

(g) Proposed
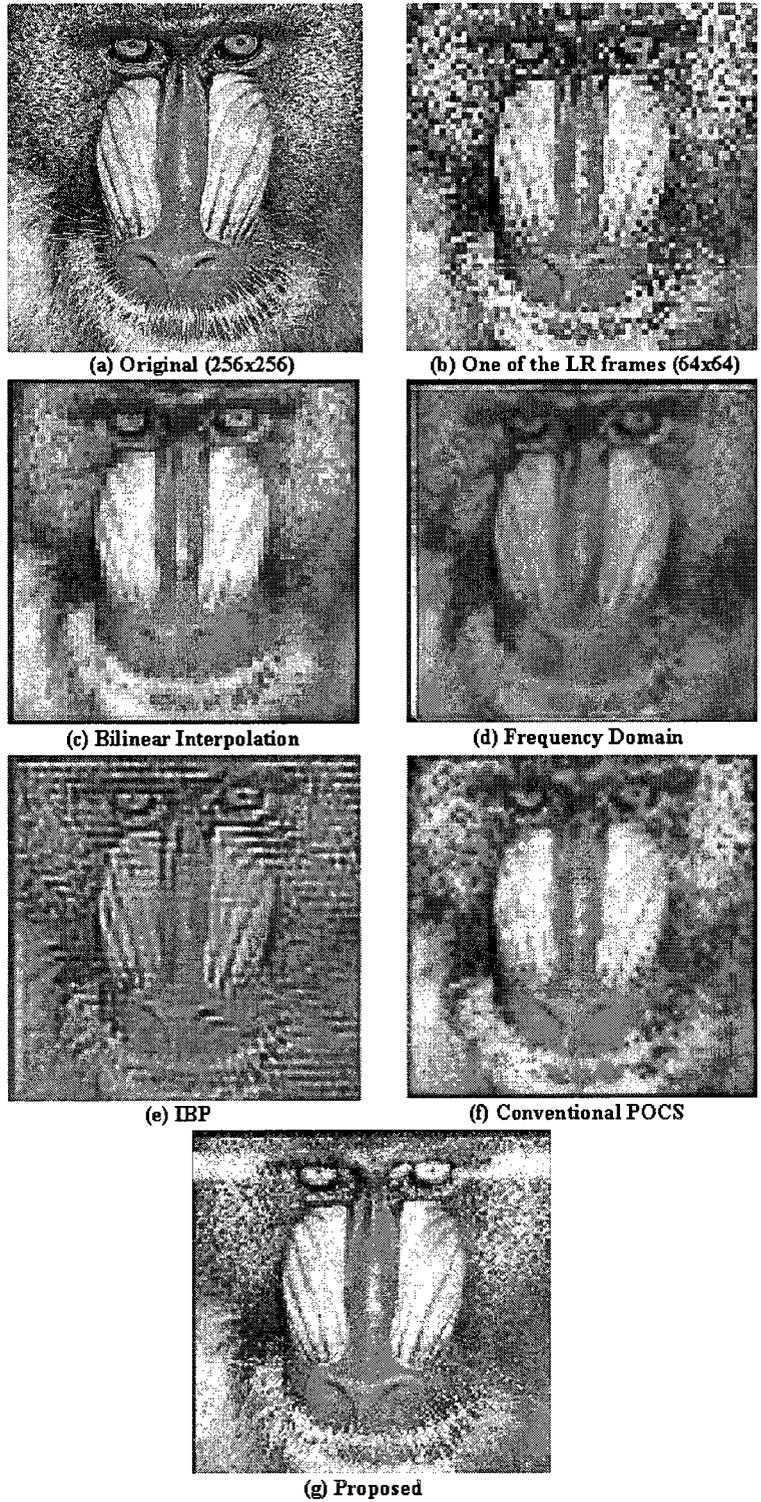
Figure 5.7: Super-resolution HR image reconstruction: Baboon

93

(a) Original (256x320)

(b) One of the LR frames (64x80)

(c) Bilinear Interpolation

(d) Frequency Domain

(e) IBP

(f) Conventional POCS

(g) Proposed
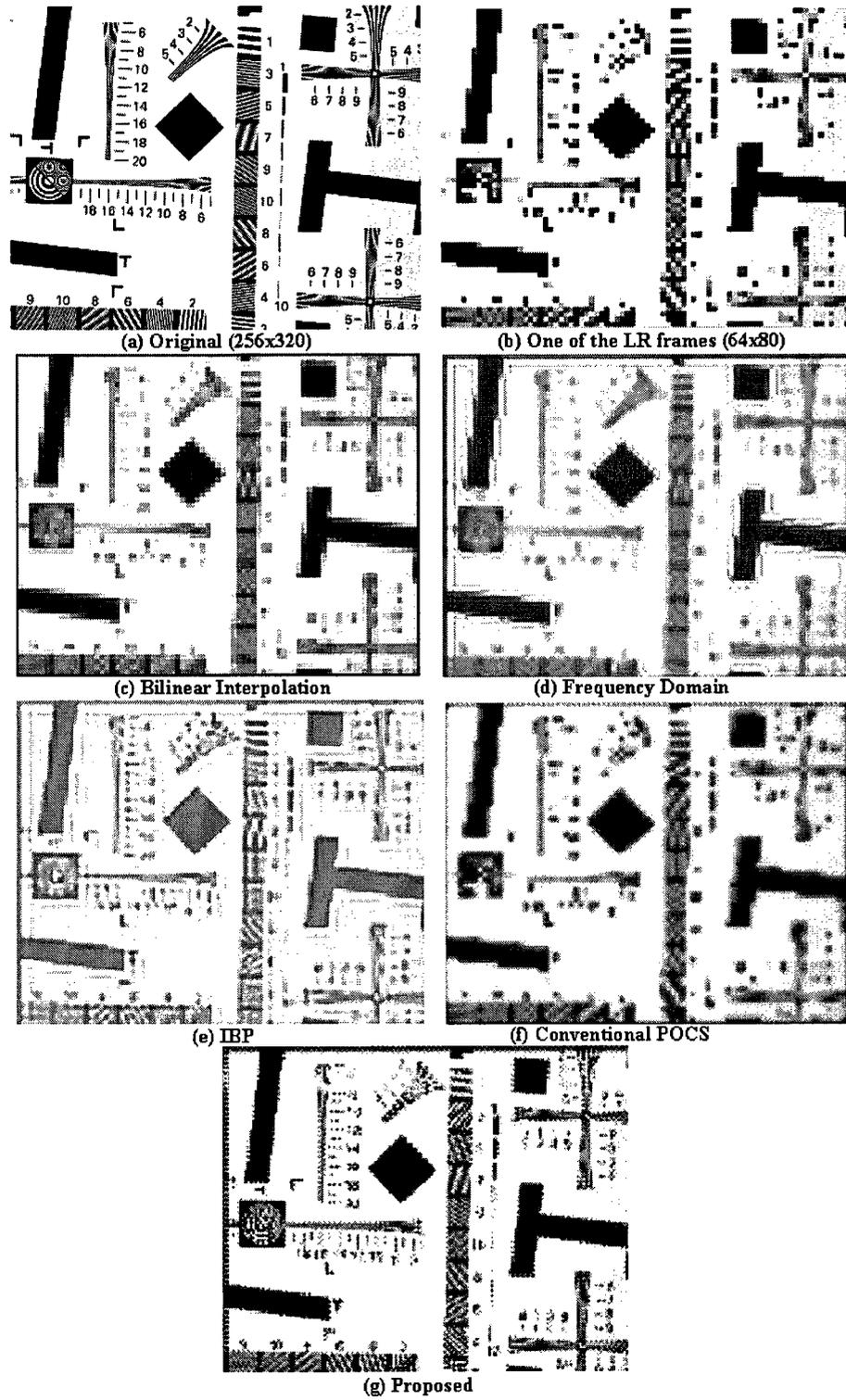
Figure 5.8: Super-resolution HR image reconstruction: Chart

94

HR image, letters with different font sizes are shown, hence the limitation of each SR reconstruction approach can be presented. Figure 5.9(a) shows the original HR image Letters which contains only white background color and black English characters with size 256×256. Figure 5.9(b) shows one of the LR frames with size 128×128. Different from the previous three experiments, the LR frames in the experiment are down-sampled from the original HR image in both horizontal and vertical directions with only a factor of 2. Figures 5.9(c)-(g) show the SR images reconstructed by using the bilinear interpolation, the frequency domain method, the IBP method, the conventional POCS and the proposed method, respectively. It is noted that the estimated HR images have the same size as the original HR image, the smallest words in the picture are hardly recognizable from the images reconstructed by the frequency domain and the conventional POCS approaches. However, most characters still can be easily read in the images constructed by the IBP approach and the proposed SR reconstruction algorithm.

5.4 Objective Experimental Results

From the subjective comparison of the five different SR reconstruction approaches in Figure 5.6- Figure 5.9 it is seen that the HR images reconstructed by the proposed SR algorithm has a better visual appearance. In the following sub-section the objective numerical results on the proposed SR method in comparison with other SR methods are shown. Also, the results of the proposed SR reconstruction method are compared to itself with different parameter setups.

ABCDEFGHIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(a) Original (256x256)

ABCDEFGIIIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(b) One of the LR frames (128x128)

ABCDEFGHIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(c) Bilinear Interpolation

ABCDEFGHIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(d) Frequency Domain

ABCDEFGHIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(e) IBP

ABCDEFGIIIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(f) Conventional POCS

ABCDEFGHIJ
abcdefghij

Superresolution
Wavelet Based

Projections Onto
Convex Sets

(g) Proposed

Figure 5.9: Super-resolution HR image reconstruction: Letters

96

Table 1: PSNR comparison of different SR approaches (Unit:dB)

	Bilinear	Freq. Domain	IBP	conv.POCS	Proposed
Lena	22.5875	21.6269	17.9043	22.2776	22.1960
Baboon	18.1108	16.7039	16.0115	17.5673	17.5065
Chart	13.8239	12.4857	11.0832	13.3915	13.3675
Letters	14.0205	13.1972	13.7533	13.0228	12.8450

5.4.1 Results of Standard Parameters Setup

The objective comparison is conducted by measuring the peak signal to noise ratio (PSNR) of each HR image reconstructed by different SR approaches. The PSNR measurements of four HR images obtained by the five SR approached are shown in Table 1.

By looking at the PSNR values shown in Table 1, most SR reconstruction approaches provide better measurement on the reconstructed image Lena but poor PSNR measurement on image Chart. By comparing each individual SR reconstruction approach, the Bilinear interpolation approach provides better PSNR measurement on most images. In contrast, the proposed algorithm provides lower PSNR values on most images. By looking at the subjective and objective comparison results, it can be concluded that a higher PSNR value does not necessarily mean a better SR reconstruction result. For example, the HR image Baboon reconstructed by the bilinear interpolation has the highest PSNR measurement. However, in Figure 5.7(b) the visual quality is not the best. A brief conclusion can be drawn that the PSNR is a very good measurement for image de-noising and single image restoration but it may not properly reflect SR reconstruction performance due to the nature of SR reconstruction process.

The convergence speed is an important index for determining whether a SR reconstruction algorithm can be performed rapidly or not. The IBP and conventional POCS approaches are used as reference algorithms to compare with the proposed wavelet-based POCS in the convergence performance. The reason is that, the IBP, conventional POCS and the proposed wavelet-based POCS SR algorithms are classified as projection-based SR

Table 2: Convergence comparison of different iterative SR approaches (Unit:Iterations)

|	IBP	conv.POCS	Proposed
Lena	29	49	20
Baboon	28	49	25
Chart	33	49	16
Letters	34	33	18

approaches. Also IBP, conventional POCS and the proposed SR are iterative algorithms. The convergence scheme is defined in Equation (5.2). Figure 5.10 and Figure 5.10 show the convergence speed comparison of IBP, conventional POCS and the proposed SR reconstruction algorithm.

In each plot of Figure 5.10, the value on the x-axis is the mean error, and the value on the axis y is the number of iterations. Figure 5.10(a)-(d) gives the convergence plots for images Lena, Baboon, Chart and Letters, respectively. It is seen that the conventional POCS needs more iterations to reach the preset convergence criterion ε, because conventional POCS does not have strong sensitivity on distinguishing the noise pixels and useful pixels. In contrast, the convergence speed drops significantly in the proposed wavelet-based POCS SR reconstruction algorithm. The reason is that the wavelet coefficient refinement process has the ability of removing the noise much faster that the projection operation in the convectional POCS algorithm. In Figure 5.10, we can notice that the convergence plots of the IBP method do not have a smooth convergence curve because the HR image estimated from each iteration of the IBP method is degraded again as compared to the acquired LR frames. During some of the iterations, the estimated complete degrading function H_i in the back-projection process (see Figure 3.5) may not be very accurate. Therefore, we can see the error value increases at some iterations. Table 2 summarizes the convergence plot by recording the number of iterations for different approaches.

It is to be noted that the proposed SR reconstruction algorithm consists of more processes than the IBP and the conventional POCS approaches, thus the execution time for
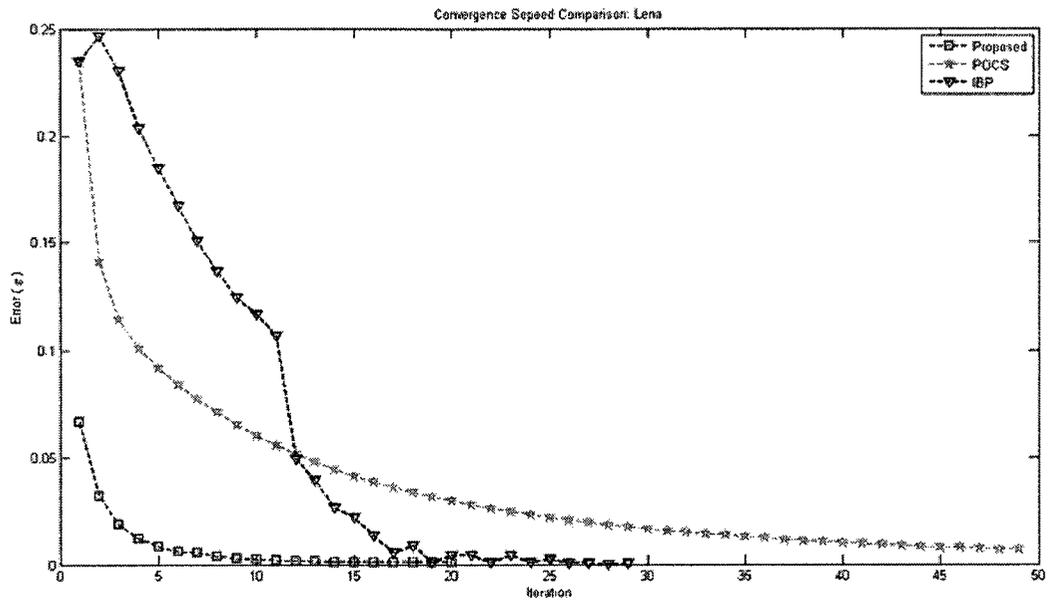
(a) Lena
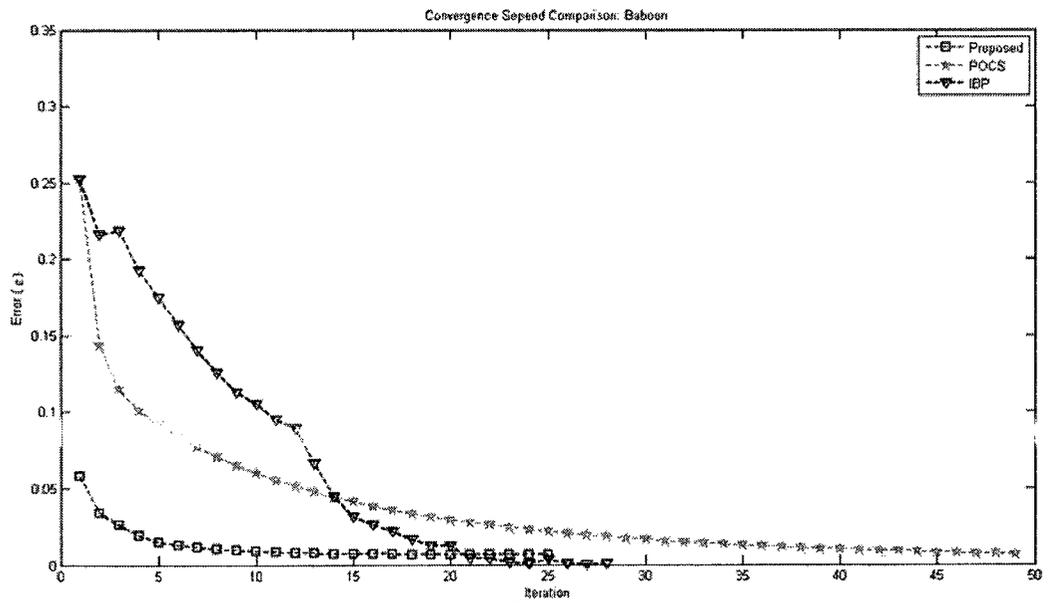
(b) Baboon

Figure 5.10: SR algorithm convergence speed comparison (1)

99

(c) Chart

(d) Letters
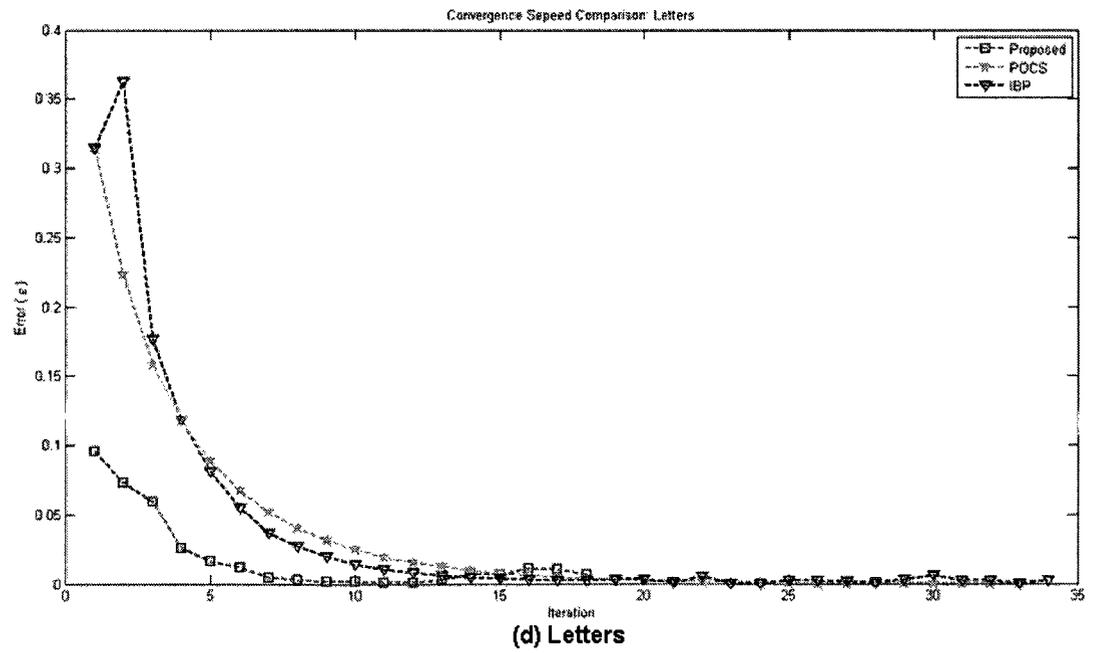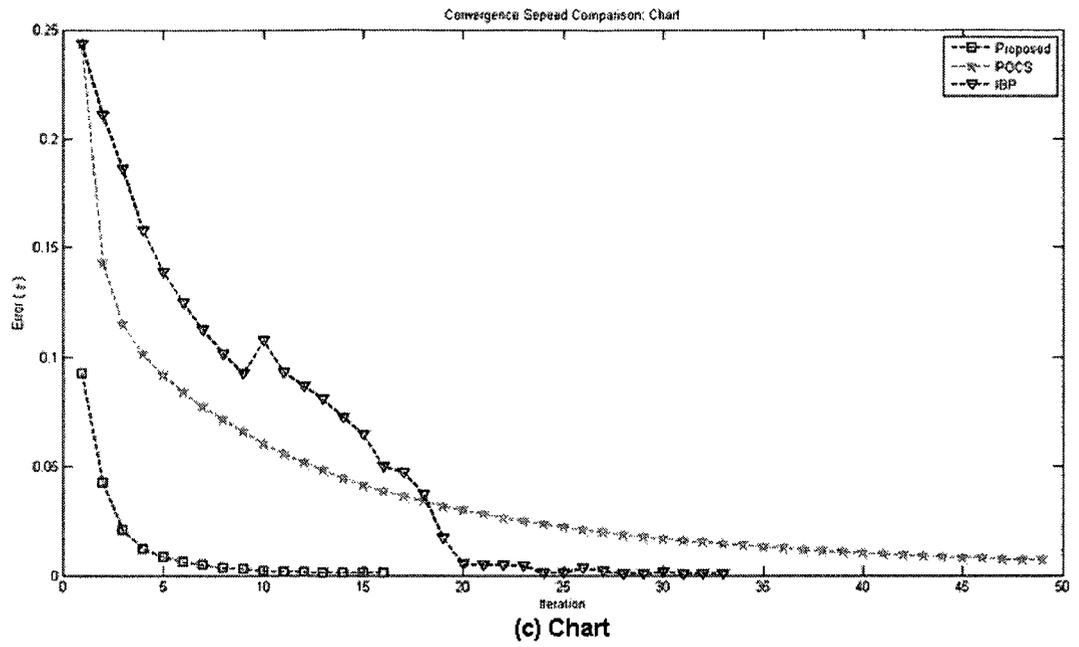
Figure 5.10: SR algorithm convergence speed comparison (2)

100

Table 3: Execution time comparison of different iterative SR approaches (Unit:Seconds)

	IBP	conv.POCS	Proposed
Average time for 1 iteration	2.76	1.73	5.31
Lena	80.04	84.77	106.2
Baboon	77.28	84.77	132.75
Chart	91.08	84.77	84.96
Letters	93.84	57.09	95.58

every iteration is relatively longer than the IBP and conventional POCS approaches. It is true that convergence speed is an important index for rating a processing speed. However, the execution time for each iteration of each different iterative SR algorithm also needs to be taken into account to judge the judge speed. The reason is that the actual total processing time is the number of iterations multiplied by the average executing time of one single iteration. In Table 3, it calculates the total processing time of a complete SR reconstruction of different approaches.

From Table 3, it can be seen that the proposed SR reconstruction algorithm takes longer total execution time when compared with other iterative approaches. However, in terms of the quality of the reconstructed HR image in Figure 5.6- Figure 5.9, a 60% more execution time to achieve a better quality might be acceptable and worthy.

5.4.2 Comparisons of Different Parameters Setup

In Section 5.3 and Sub-section 5.4.1 the subjective and objective experimental results of different RS reconstruction approaches have been presented. The parameters of the proposed SR algorithm used in therein are selected to produce most visually apparent HR images. In this sub-section, the simulation parameters of the proposed SR algorithm are set differently in order to show the effects of these parameters.

There are many parameters that control the quality of final HR image. In order to focus on the major process, the parameters to be modified are mainly the parameters of the

Table 4: The list of different settings of the simulation parameters

	Setting 1	Setting 2	Setting 3	Setting 4
Daubechies wavelet type	$db2$	$db3$	$db2$	$db2$
window size for variance	$d_2 = 5$	$d_2 = 5$	$d_2 = 7$	$d_2 = 5$
window size for F-test	$d_2 = 7$	$d_2 = 7$	$d_2 = 9$	$d_2 = 7$
level of confidence 1.	$\gamma_1 = 0.01$	$\gamma_1 = 0.01$	$\gamma_1 = 0.01$	$\gamma_1 = 0.05$
level of confidence 2.	$\gamma_2 = 0.05$	$\gamma_2 = 0.05$	$\gamma_2 = 0.05$	$\gamma_2 = 0.1$
predetermined attenuating factor	$\beta = 0.5$	$\beta = 0.5$	$\beta = 0.5$	$\beta = 0.5$
integer weighting factor	$\xi = 5$	$\xi = 5$	$\xi = 5$	$\xi = 5$

wavelet coefficient refinement process.

There are four different parameter settings used in this sub-section. The first setting is the original setting which was used in previous simulations. The second setting is the order of Daubechies wavelet function. The third setting are the window sizes d_1 and d_2. The fourth setting has both levels of confidence γ_1 and γ_2. The list of different settings of the simulation parameters for comparison is given in Table 4.

The HR images reconstructed by the proposed SR algorithm with four different parameter settings are shown in Figure 5.11. The images at the most left hand side column (a1), (b1), (c1) and (d1)are the reconstructed HR images with "Setting 1" which has the same parameters setting as the HR images shown in Figure 5.6(g)- Figure 5.9(g). The images at the second column (a2), (b2), (c2) and (d2) are the HR images reconstructed by replacing the order of Daubechies wavelet function with 'db3'. The images at the third column (a3), (b3), (c3) and (d3) are the HR images reconstructed by changing the size of variance estimation window d_1 and F-distribution test window d_2. The images at the fourth column (a4), (b4), (c4) and (d4) are the HR images reconstructed by changing the value of level of confidence parameters γ_1 and γ_2. From Figure 5.11it is clear that changing the wavelet function order from 'db2' to 'db3' does not affect the final result a lot. In contrast, changing the value of level of confidence γ_1 and γ_2 parameters makes the estimated HR images more blur.

The PSNR measurements of the HR images shown in Figure 5.11 are listed in Table

|			
(a1)	(a2)	(a3)	(a4)
(b1)	(b2)	(b3)	(b4)
(c1)	(c2)	(c3)	(c4)
(d1)	(d2)	(d3)	(d4)
Setting1	Setting2	Setting3	Setting4
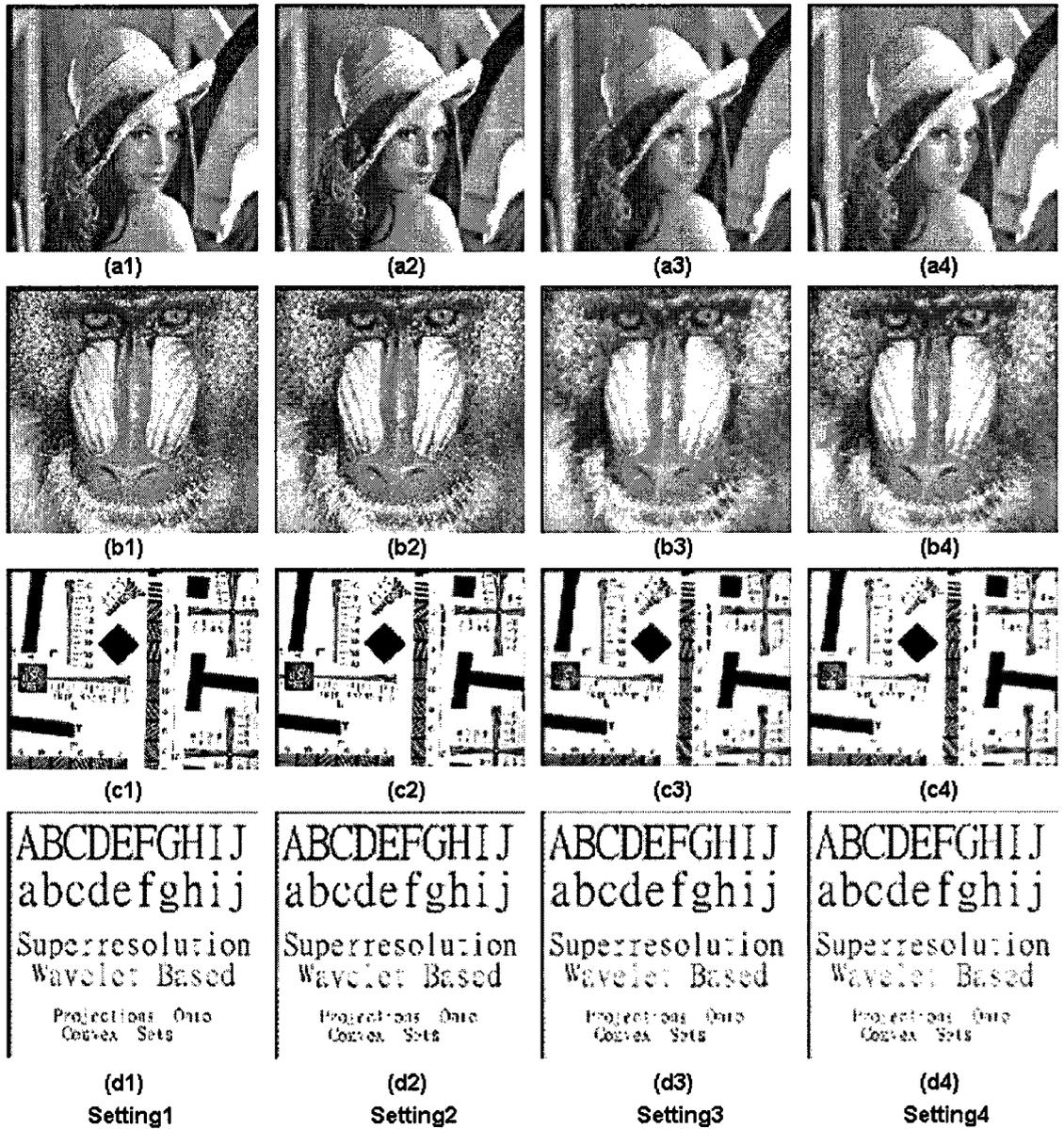
Figure 5.11: HR images reconstructed by four different parameter settings

Table 5: PSNR comparison of different parameter settings (Unit:dB)

	Setting 1	Setting 2	Setting 3	Setting 4
Lena	22.1960	22.1470	20.9985	18.5020
Baboon	17.5065	17.1655	16.5690	15.9525
Chart	13.3675	12.9430	12.1430	11.9950
Letters	12.8450	12.1100	12.2100	11.9590

5. This table provides an numerical comparison on the different parameter settings. It is seen that the PSNR value drops a little bit while changing of the order of wavelet function. However, changing the value of level of confidence parameters affects the PSNR significantly.

The comparison of convergence speed on different parameter settings is presented graphically in Figure 5.12. The convergence plots of the four reconstructed images, Lena, Baboon, Chart and Letters, with respect to the four different settings are shown in Figure 5.12(a), (b), (c) and (d), respectively. By observing the convergence plots in each figure, it can be summarized that (1) changing the order of wavelet function keeps the convergence speed unchange, (2) changing the size of the estimation windows could cause the convergence a little slower, and (3) changing the value of level of confidence causes the convergence much slower by at least two or more iterations.

5.5 Conclusion

Extensive simulation studies on the proposed wavelet-based SR approach in comparison with some existing methods have been conducted. The experimental setups and the comparison criteria are provided in Section 5.1 of this chapter. The intermediate result of two processes of the proposed SR algorithm, the "Projections onto wavelet constraint" and the "wavelet coefficient refinement" process, are firstly presented. A subjective comparison shows that the HR images reconstructed by the proposed algorithm have better appearance and quality. The numerical PSNR comparison reveals that a higher PSNR measurement

(a) Lena

(b)Baboon
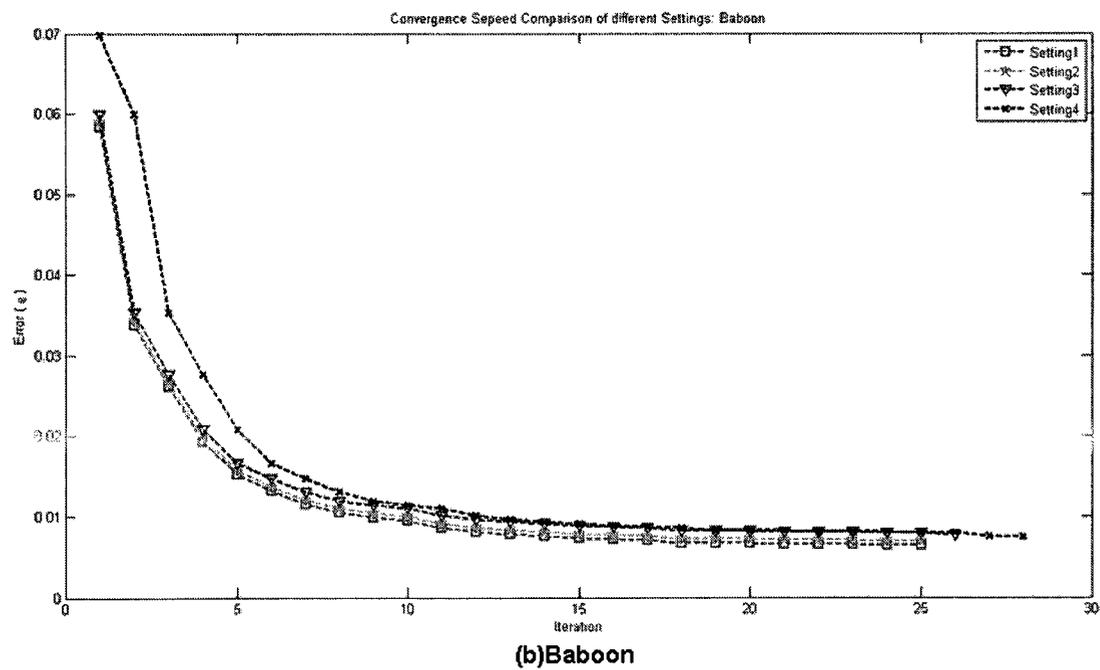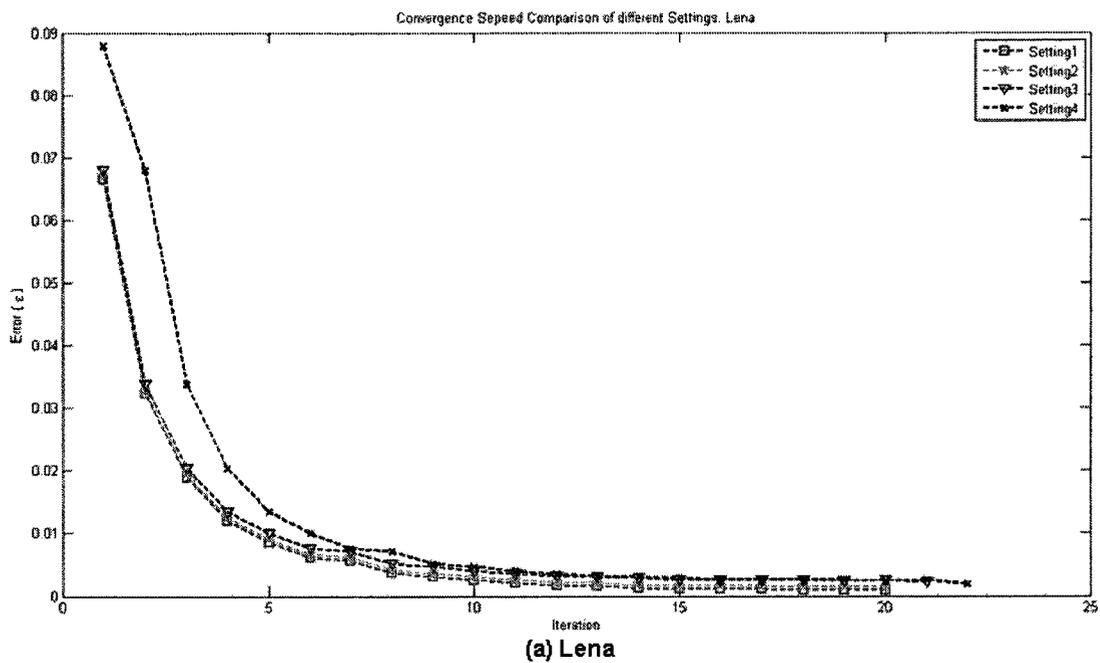
Figure 5.12: Convergence speed plots of different parameter settings for four images (1)
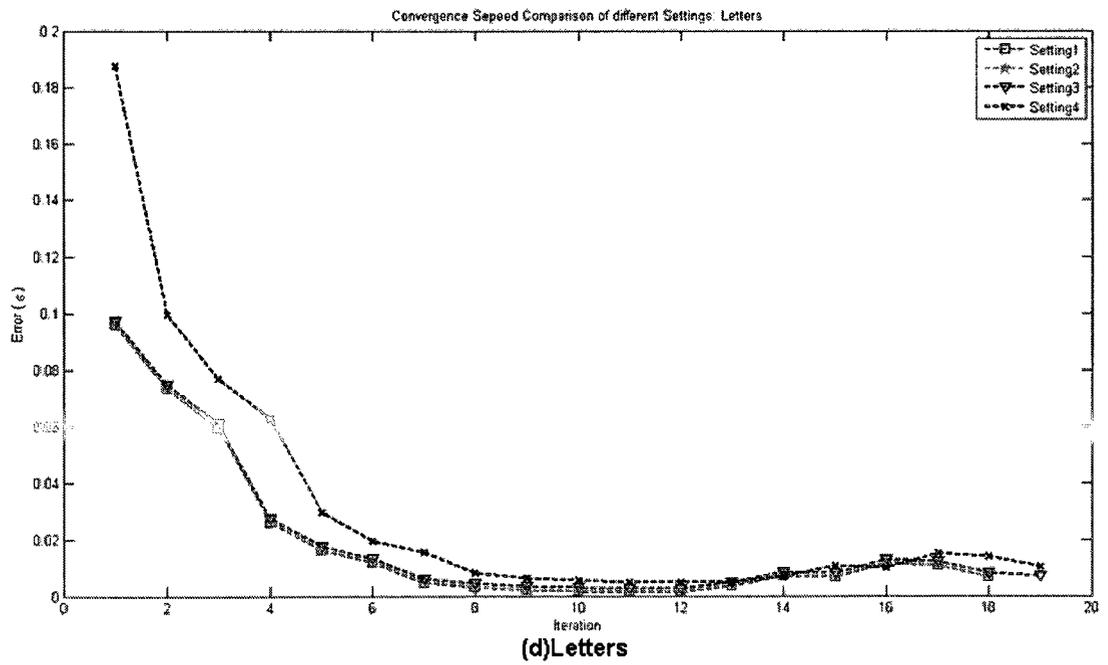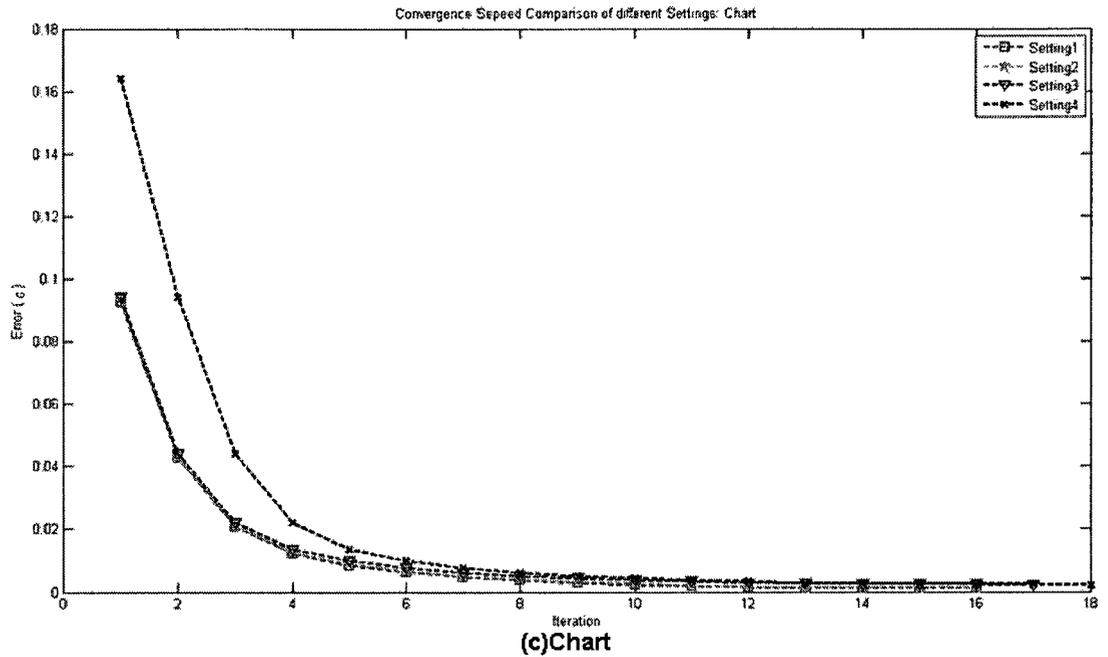
Figure 5.12: Convergence speed plots of different parameter settings for four images (2)

does not truly suggest a good SR algorithm because the PSNR can not show whether the reconstructed HR image has better resolution or not. After the PSNR comparison, the processing time of the proposed method is considered and compared with other methods. By considering the total execution time and subjective HR image results, we have found that when the speed and the quality can not be perfectly optimized, it is worthy to have a better HR image but fairly long computational time. In the end of this chapter, the comprehensive comparisons of the proposed SR method with different parameter setups is given to show the effects of changing parameters.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, several SR reconstruction approaches such as the bilinear interpolation, frequency domain approach, IBP and the conventional POCS method have been investigated and evaluated. The thesis mainly focuses on the improvement of the conventional POCS method by introducing wavelet constraint projection and the wavelet coefficients refinement process. The proposed SR reconstruction is designed algorithm in Chapter 4 and the simulation work is presented in Chapter 5, which constitute the major contributions of the thesis. The simulation result has demonstrated that the proposed algorithm produces better images as perceived by the human eye. The overall content of the thesis can be summarized as follows:

- Super-resolution refers to a set of image enhancement algorithms that reconstruct high resolution image from a sequence of degraded, low resolution frames. The forward observation model of the SR reconstruction problems described in Equation (1.1) serves as a foundation of SR reconstruction research.

- The general steps of an SR reconstruction approach shown in Figure 2.1 divide the

108

complicated SR reconstruction process into four simpler procedures. First, the motion registration step aligns the acquired consecutive LR frames. Second, the degrading function estimation step estimates the kernel of the blurring function of LR frames. Third, the SR reconstruction step fuses acquired frames into one HR image and last, the de-blurring step removes the blurring effects from the estimated HR image.

- The POCS is a method that is used to recover incomplete signals by performing the convex set projection operation. The conventional POCS SR reconstruction approach uses data consistency constraint C_D and amplitude constraint C_A as the convex sets for the projection operation. Based on the POCS method, the more the convex sets have been used, the better the SR image quality obtained. Therefore, in this thesis, one more constraint ,namely, the wavelet domain constraint C_W described in Equation (4.15) is introduced.

- In order to further remove the noise and the wrongly estimated pixels, the thesis introduces the stochastic wavelet coefficient refinement process. In the refinement process, the wavelet coefficient sub-images are modified according to the distribution of their variances. The variance estimation scheme applies 1-D orthogonal windows with respect to the detailed features to each wavelet coefficient sub-image. The coefficient modification scheme employs the F-distribution test to judge whether a wavelet coefficient should be preserved or modified.

- The complete procedure of the proposed SR image reconstruction algorithm shown in Figure 4.13 explains the process flow and the domain space of the input data.

- It is shown that the HR images reconstructed by the proposed SR reconstruction method does not have superior PSNR values as compared to other existing methods. However, the HR images from the proposed SR reconstruction method have a better

visual quality and higher resolution.

6.2 Future Work

Among many image and video signal processing studies, super-resolution is a spot-lighted research area. The following points out some possible research directions and applications where the proposed SR reconstruction algorithm could be employed to improve high definition visual systems.

- The SR observation model described in this thesis considers a simple frame to frame motion model. To make a broad use of the proposed SR algorithm, a more complicated motion model such as allowing multiple local motions between LR frames, or considering a various zooming ratios of consecutive frames, could be further discussed as new research directions.

- In the proposed SR reconstruction method, most adjustable parameters are in the wavelet coefficients refinement process. To achieve the best HR image quality, parameters may need to be independently fine tuned for different types of video scenes. An investigation on parameter adjustment for various types of LR image sequences would make the proposed SR reconstruction method more useful.

- The two extended processes in the proposed SR reconstruction method mainly use wavelet techniques. Recently, there are some new data transform techniques such as curvelet and bandlet which inherit the analysis capabilities of wavelet transform and have their own advantages in image processing research area. Therefore, the curvelet and bandlet could be valuable techniques to replace the wavelet used in the proposed SR method..

- The proposed SR reconstruction algorithm mainly considers reconstructing HR images from the luminance portion of the LR frames. Additionally, this algorithm can be further applied to color SR image reconstruction by considering chrominance components and the color CCD's Bayer pattern [47] . The Bayer pattern is a digital color image acquisition pattern. To compensate the effect of the Bayer pattern, color super-resolution research is usually combinesd with the demosaicing techniques [48] [49]. The application of the proposed algorithm may bring certain advantages to adapt to demosaicing techniques.

- There are many image and video compression and coding algorithms that use wavelet analysis and related techniques. The proposed wavelet-based POCS SR reconstruction algorithm provides the opportunity to combine other wavelet related techniques, since the proposed algorithm operates the frame data in the wavelet domain. Therefore, adopting wavelet based compression algorithms or transmission coding techniques to the proposed SR algorithm would lead to further development in the field of SR.

- To reduce the processing time of an SR algorithm, one possible solution is to build a hardware chip to perform the data processing. The proposed SR algorithm is a wavelet-based iterative algorithm; some steps of the algorithm such as wavelet transformation and coefficient variance calculation can be easily replaced by hardware circuits. Therefore, converting the proposed algorithm into a hardware chipset and estimating the real-time SR reconstruction performance can be a very good extension of this work.

The proposed SR reconstruction algorithm and the simulation work presented in this thesis can serve as a reference to further researches.

111

References

[1] M. Irani and S. Peleg, "Motion analysis for image enhancement resolution, occlusion, and transparency," *J. Visual Commun. Image Represent.*, vol. 4, pp. 324–335, Dec. 1993.

[2] N. X. Nguyen, *Numerical algorithms for image superresolution*. PhD thesis, Graduate Program in scientific computation and computational mathematics, Stanford University, July 2000.

[3] A. Tekalp, M. Ozkan, and M. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, vol. 3, pp. 169–172, 1992.

[4] H. Yang, J. Gao, and Z. Wu, "Blur identification and image super-resolution reconstruction using an approach similar to variable projection," *IEEE Signal Processing Letters*, vol. 15, pp. 289 –292, Feb. 2008.

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

[6] D. Slepian and H. Pollak, "Prolate spheroidal wave functions, fourier analysis and uncertainty-i," *BSTJ*, pp. 40:43–62, Jan. 1961.

[7] A. Papoulis, "A new algorithm in spectral analysis and bandlimited extrapolation.," *IEEE Trans. Cir. Sys., CAS-22(9)*, vol. 22, pp. 735–742, Sep. 1975.

[8] T. S. Huang and R. Y. Tsai, "Multi-frame image restoration and registration," *Advances in Computer Vision and Image Processing, JAI Press Inc., Greenwich, CT*, vol. 1, no. 2, pp. 317–339, 1984.

[9] M. Irani and S. Peleg, "Super resolution from image sequences," *10th International Conference on Pattern Recognition 1990. Proceedings.*, vol. 2, pp. 115 –120, jun. 1990.

[10] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Proc.*, May 1991.

[11] D. Youla and H. Webb, "Image restoration by the method of convex projections:part 1- theory," *IEEE Trans. on Medical Image*, vol. MI-1, pp. 81–94, Oct. 1982.

[12] M. Sezan and H. Stark, "Image restoration by the method of convex projections:part 2- applications and numerical result," *IEEE Trans. on Medical Image*, vol. MI-1, pp. 95–101, Oct. 1982.

[13] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *Optical Society of America*, vol. 6, pp. 1715–1726, Nov. 1989.

[14] M. Elad and A. Feuer, "Restoration of single super-resolution image from several blurred, noisy and downsampled measured images," *IEEE Trans. Image Processing 6*, pp. 1646–1658, Dec. 1997.

[15] N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient image superresolution algorithm," *IEEE Trans. Image Processing 10*, pp. 573–583, April 2001.

[16] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust super-resolution," *International Conference on Image Processing ICIP 2003. Proceedings*, vol. 3, pp. 291–294, Sept. 2003.

[17] S. Borman and R. L. Stevenson, "Super-resolution from image sequences - a review," *in Proceedings of the 1998 Midwest Symposium on Circuits and Systems*, vol. 5, April 1998.

[18] C. Papathanassiou and M. Petrou, "Super resolution: an overview," *2005 IEEE International IGARSS 2005 Proceedings*, vol. 8, pp. 5655–5658, July 2005.

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

[20] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," *Eur. Conf. Computer Vision*, pp. 237–252, 1992.

[21] Y. Wang, J. Ostermann, and Y.-Q. Zhang, *Video Processing and Communications*. Prentice Hall, 2002.

[22] W. M. van Eekeren, K. Schutte, and J. van Vliet, "Multi-frame super-resolution reconstruction of small moving objects," *IEEE Transactions on Image Processing*, vol. PP, no. 99, pp. 1–12, 2010.

[23] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transaction on Communication*, vol. 31, no. 31, pp. 532–540, 1983.

[24] X. Wei, F. Zhang, H. Chen, and Q. Qin, "Blind super-resolution image reconstruction based on pocs model," *ICMTMA 2009. International Conference on Measuring Technology and Mechatronics Automation*, vol. 1, pp. 437–440, Apr. 2009.

[25] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Wiley, New York, 1949.

[26] M. Almeida and L. Almeida, "Blind deblurring of foreground-background images," *2009 16th IEEE International Conference on Image Processing (ICIP)*, pp. 1301–1304, Nov. 2009.

[27] C. Lindley, *Practical Image Processing in C*. Big Apple Tuttle-Mori Agency Inc. New York, April 1994.

[28] G. Ramponi, "Warped distance for space-variant linear image interpolation," *IEEE Trans. on Image Processing*, vol. 8, pp. 629–639, May 1999.

[29] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Signal Processing*, vol. 29, pp. 1153–1160, Dce. 1981.

[30] H. Stark, *Image Recovery: Theory and Application*. Academic Press, Orlando, Florida, 1987.

[31] L. Bregman, "Finding the common point of convex sets by the method of successive projections," *Dokl. Akud. Nauk. SSSR*, vol. 162, no. 31, pp. 487–490, 1965.

[32] L. Gubin, B. Polyak, and E. Raik, "The method of projections for finding the common point if convex sets," *USSR Comput. Math. Math. Phys. (Engl. Transl.)*, vol. 7, no. 61, pp. 1–24, 1967.

[33] A. Levi and H. Stark, "Signal restoration from phase by projections onto convex sets," *Optical Society of America*, vol. 73, pp. 810–822, June 1983.

[34] P. A. Jansson, *Deconvolution of Images and Sepctra*, pp. 478–499. E. I. du Pont de Nemours and Company (Inc.) Experimental Station Wilmington, Delware, second ed., 1997.

[35] A. Katsaggelos, *Digital Image Restoration*, vol. 23. Heidelberg, Germany: Springer-Verlag. Springer., 1991.

[36] H. Yu, X. Ma, H. Huang, and C. Qi, "Face image super-resolution through pocs and residue compensation," *VIE 2008 5th International Conference on Visual Information Engineering*, pp. 494–497, Jul. 2008.

[37] V. Patanavijit and S. Jitapunkul, "A lorentzian bayesian approach for robust iterative multiframe super-resolution reconstruction with lorentzian-tikhonov regularization," *2006. ISCIT '06. International Symposium on Communications and Information Technologies*, pp. 1044–1049, Sep. 2006.

[38] B. Wei and W. Hui, "Pocs-embedded map method for image super-resolution restoration," *ICIEA 2009 4th IEEE Conference on Industrial Electronics and Applications*, pp. 3791–3794, May 2009.

[39] R. Schultz and R. Stevenson, "A bayesian approach to image expansion for improved definition," *IEEE Trans. Image Processing*, vol. 3, pp. 233–242, May 1994.

[40] Q. Razlighi, N. Kehtarnavaz, and A. Nosratinia, "Computation of image spatial entropy using quadrilateral markov random field," *IEEE Transactions on Image Processing*, vol. 18, pp. 2629–2639, Dec. 2009.

[41] M. Tanaka and M. Okutomi, "Locally adaptive learning for translation-variant mrf image priors," *CVPR 2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Jun. 2008.

[42] H. I. Koo and N. I. Cho, "Prior model for the mrf modeling of multi-channel images," *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International*, vol. 1, pp. 713–716, Apr. 2007.

[43] T. Chen and D. Metaxas, "Integration of gibbs prior models and deformable models for 3d medical image segmentation," *2002 Proceedings 16th International Conference on Pattern Recognition*, vol. 1, pp. 719–722, Dec. 2002.

[44] T. Hebert and R. Leahy, "Statistic-based map image-reconstruction from poisson data using gibbs priors," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2290–2303, Sep. 1992.

[45] G. Lei and H. Zhiming, "A projection on convex sets super-resolution algorithm using wavelet transform," *2008. ICSP 2008. 9th International Conference on Signal Processing,*, pp. 1039–1041, Oct. 2008.

[46] Y. Hawwar and A. Reza, "Spatially adaptive multiplicative noise image denoising technique," *IEEE Transactions on Image Processing*, vol. 11, pp. 1397–1404, Dec. 2002.

[47] R. Lukac, K. Plataniotis, and D. Hatzinakos, "Color image zooming on the bayer pattern," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, pp. 1475 – 1492, nov. 2005.

[48] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *Image Processing, IEEE Transactions on*, vol. 15, pp. 141–159, Jan. 2006.

[49] T. Y. Jung, S. Yang, and J. Jeong, "Multi-directional demosaicing for digital still cameras," *2009 WRI World Congress on Computer Science and Information Engineering*, vol. 7, pp. 374–378, Mar. 2009.