AN EXAMINATION OF SUPER RESOLUTION METHODS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

YILCA BARIŞ SERT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

APRIL 2006

Approval of Graduate School of Natural and Applied Science.

_____

Prof. Dr. Canan ÖZGEN

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____

Prof. Dr. İsmet ERKMEN

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____   _____

Assoc. Prof. Dr. Gözde BOZDAĞI AKAR   Asst. Prof. Dr. Çağatay CANDAN

Co-Supervisor   Supervisor

Examining Committee Members:

Assoc. Prof. Dr. Aydın ALATAN   (METU, EEE) ——————————

Asst. Prof. Dr. Çağatay CANDAN   (METU, EEE) ——————————

Assoc. Prof. Dr. Gözde BOZDAĞI AKAR (METU, EEE) ——————————

Asst. Prof. Dr. Ali Özgür YILMAZ   (METU, EEE) ——————————

Umur AKINCI (M.S.)   (ASELSAN)   ——————————

ii

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, surname   :   Yılca Barış SERT

Signature   :

ABSTRACT

AN EXAMINATION OF SUPER RESOLUTION METHODS

Sert, Yılca Barış

M.S., Department of Electric and Electronics Engineering

Supervisor: Asst Prof. Dr. Çağatay CANDAN

Co-Supervisor: Assoc. Prof. Dr. Gözde BOZDAĞI AKAR

April 2006, 114 pages

The resolution of the image is one of the main measures of image quality. Higher resolution is desired and often required in most of the applications, because higher resolution means more details in the image. The use of better image sensors and optics is an expensive and also limiting way of increasing pixel density within the image. The use of image processing methods, to obtain a high resolution image from low resolution images is a cheap and effective solution. This kind of image enhancement is called super resolution image reconstruction.

This thesis focuses on the definition, implementation and analysis on well-known techniques of super resolution. The comparison and analysis are the main concerns to understand the improvements of the super resolution methods over single frame interpolation techniques. In addition, the comparison also gives us an insight to the practical uses of super resolution methods. As a result of the analysis, the critical examination of the techniques and their performance evaluation are achieved.

Keywords: super resolution, image enhancement, image reconstruction

ÖZ

SÜPER ÇÖZÜNÜRLÜK METODLARI ÜZERİNE BİR İNCELEME

Sert, Yılca Barış

Yüksek Lisans, Elektrik-Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Yard.Doç. Dr. Çağatay CANDAN

Ortak Tez Yöneticisi: Doç. Dr. Gözde BOZDAĞI AKAR

Nisan 2006, 114 sayfa

Çözünürlük, imge kalitesi için ana ölçütlerden biridir. Yüksek çözünürlük, daha çok ayrıntı demek olduğundan çoğu uygulamada istenmekte hatta gerekmektedir. Daha iyi imge algılayıcıları ve daha kaliteli optik teçhizatın kullanılması ise pahalı ve sınırlayıcı bir çözüm olarak karşımıza çıkmaktadır. Ucuz ve etkili bir çözüm olması açısından düşük çözünürlükteki imgelerden yüksek çözünürlükte imge elde edilmesi için görüntü işleme yöntemlerinin kullanılması önemlidir. Bu tür iyileştirme, süper çözünürlükte imge yapılandırılması olarak adlandırılmaktadır.

Bu tez, iyi bilinen süper çözünürlük tekniklerinin tanım, uygulama ve değerlendirilmeleri üzerinde yoğunlaşmıştır. Süper çözünürlük yöntemlerinin karşılaştırılması ve analizi, bu yöntemlerin interpolasyona dayalı tek imgeli iyileştirme yöntemleri karşısındaki gelişmelerini anlamak açısından ön plandadır. Buna ek olarak, her yöntemin analizi ile kullanımsal yönden bir anlaşılırlık oluşmuştur. Bu karşılaştırmanın sonucunda teknikler üzerinde eleştirel bir inceleme yapılmış ve bir başarım incelemesi gerçekleştirilmiştir.

Anahtar Kelimeler: süper çözünürlük, imge iyileştirme, imge yapılandırılması

To My Beloved Wife and My Family

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Çağatay CANDAN and co-supervisor Gözde BOZDAĞI AKAR for their supervision, guidance and help throughout this study.

Special thanks to ASELSAN Inc. MGEO division for providing me every kind of convenience for the past four years to complete this study.

Of course, I greatly appreciate my colleagues in KSTM and ETM for their valuable friendship, help and support.

Especially, I am deeply grateful to my LOVE and my beloved wife Gülşen for her endless love and being there for me when I need her the most.

I would like to express my love and appreciation to my precious family; Cahit SERT my dear father, Yıldız SERT my dear mother and Gül Berrak SERT my little sister, who have been my great supporters in every aspect of my life.

I would like to extend my appreciation to my unique friends for their fellowship and encouragement throughout my life.

Thank you so much to all of YOU and me[*]!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SR   Super Resolution

IBP   Iterative Back Projection

POCS   Projection Onto Convex Sets

RANSAC   Random Sample Consensus based Image Registration Method

QM   Quality Metrics

LR   Low Resolution

HR   High Resolution

FT   Fourier Transformation

DFT   Discrete Fourier Transformation

Keren   Image Registration Method developed by Keren et al.[4]

Marcel   Image Registration Method developed by Marcel et al.[7]

Vandewalle   Image Registration Method developed by Vandewalle et al.[1]

MSE   Mean Square Error

PSNR   Peak Signal to Noise Ratio

SSIM   Structural SIMilarity Index

GUI   Graphics User Interface

CHAPTER 1

INTRODUCTION

1.1. INTRODUCTION TO SUPER RESOLUTION

Digital imaging is taking a great part in our life day by day and constantly we require better image quality, higher resolution and more functionality. In the scope of the high-resolution requirements, the imaging chips and optical components necessary to capture very high-resolution images become very expensive. On the other hand, the scientific research to build up better components is almost reached a limiting level, which encourages us to consider a cheaper and promising solution to the resolution problem.

The wide range of capabilities through signal processing, specifically image processing solves this problem in a cheap but efficient way. The use of a series of low resolution frames captured by a moderate digital camera or a video recorder, to build up a high resolution image is a very interesting and useful way, which is called super resolution image enhancement. Thereby, this approach cost less and the existing low resolution imaging systems can be utilized. The basic idea behind Super-Resolution (SR) is the fusion of a sequence of low-resolution noisy blurred images to produce a higher resolution image.

In super resolution the low resolved images represents different views at the same scene. The key idea is strongly related to the fact that every low-resolution image contains different information on the same scene and the fusion of this information pieces, makes it possible to extract the subpixel information on the low-resolution

1

image. The subpixel information means that new pixels are present among our existing pixel values that lead us to a higher resolved image.

If the low-resolution images are shifted by integer values, then each image contains the same information and we would finally have a bunch of shifted versions of the same image not the same scene. This means that every image can be obtained from the other one, but we need more to achieve the goal of information synthesis. If only the low-resolution images have subpixel shifts, extra information of the scene is at hand. The new information within the low-resolution images can be exploited to get a higher resolution copy of the scene.

As the image-capturing environment is not ideal, many distortions are also present in the low-resolution images. We may have blurred, noisy, aliased low resolution captures of the scene. Although the main concern of the super resolution methods is to obtain higher resolution images from the low-resolution image sequences, it also covers techniques of image restoration and image enhancement techniques [9].

1.2. FIRST FORMULATION

Tsai and Huang were the first to consider the problem of obtaining a high-quality image from several lower quality and translationally displaced images in 1984 [5]. Their data set consisted of terrestrial photographs taken by Landsat satellites. They modeled the photographs as aliased, translationally displaced versions of a constant scene. Their approach consisted in formulating a set of equations in the frequency domain, by using the shift property of the Fourier transform. Optical blur or noise was not considered. Tekalp, Ozkan and Sezan [11] extended Tsai-Huang formulation by including the point spread function of the imaging system and observation noise.

Super-resolution techniques have found many other applications since the first formulation of the problem. Some of these applications are [12]:
- Satellite imaging
- Astronomical imaging

- Video enhancement and restoration

- Video standards conversion

- Confocal Microscopy

- Digital mosaicing

- Aperture displacement cameras

- Medical computed tomographic imaging

- Diffraction tomography

- Video freeze frame and hard copy

- Restoration of MPEG-coded video streams

1.3. MAIN STUDIES ON SUPER-RESOLUTION

The Super resolution algorithms can be categorized into two groups as Frequency Domain Methods and Spatial Domain Methods:

Frequency-Domain Superresolution Restoration Methods:
- Restoration via Alias Removal [5, 19]

- Recursive Least Squares Methods [20, 21, 22]

- Recursive Total Least Squares Methods [12]

- Multichannel Sampling Theorem Methods [23, 24]

Spatial-Domain Superresolution Restoration Methods:
- Interpolation of Nonuniformly-Spaced Samples [25, 26, 27, 28, 29]

- Algebraic Filtered Back-Projection Methods [30]

- Iterative Back-Projection Methods [31, 32, 33, 34]

- Stochastic Methods [35-48]

- Set Theoretic Methods [11, 51, 52, 53]

- Hybrid Methods [54, 55]

- Optimal and Adaptive Filtering Methods [56, 57]

A detailed exposition of major super resolution methods is given in Chapter 3..

1.4. SCOPE OF THE THESIS

The scope of the thesis is on the implementation, comparison, of some of the well-known super resolution techniques. This study examines the SR schemes as the cascade of two steps image registration and image fusion. The two steps will be covered in detail. Our goal is to understand and distinguish the advantages and disadvantages of major super resolution methods.

A MATLAB GUI has been implemented to test the super resolution techniques. By the help of the implementation, methods have been critically examined and some additions have been made to improve the visual quality. Besides, the comparison and analysis of the super resolution methods the robustness of the methods under noisy conditions has also been examined.

1.5. ORGANIZATION OF THE THESIS

The thesis contains five chapters. A brief introduction to super resolution is given in Chapter 1.

In Chapter 2, the super resolution algorithms are surveyed. Some image registration methods that will be useful in this study are also examined. Based on this examination the later steps of the method are shaped. The successful image registration methods are selected to be used in image fusion step of super resolution. In addition to these, we introduce the quality metrics used throughout the thesis in this chapter.

In Chapter 3, the vital part of super resolution algorithms, that is image fusion, is surveyed. The advantages and disadvantages are briefly discussed for all of the methods. The single and multi frame methods are illustrated out by some exemplary examples. The analysis and comparison of super resolution algorithms are given. The methods are compared using both artificially generated and captured image and video sequences.

In Chapter 4, the discussions on the different approaches of super resolution are given.

In Chapter 5, the conclusions on the experimental results and the related future work are presented.

CHAPTER 2

SUPER RESOLUTION METHODOLOGY

2.1. INTRODUCTION

In this chapter, we will discuss the methodology of super-resolution. We start with the mathematical description of super-resolution concept to expand our understanding of the problem. The next step will be the capturing stage of real world images. The problems and difficulties of image acquisition are discussed. Super resolution techniques are qualitatively introduced to solve some practical high quality image acquisition problems.

In this chapter, the image registration methods used throughout this thesis is described. A comparison of these methods under different motion types is given. Finally, the objective image quality metrics used for the comparison of super-resolved images are introduced.

2.2. THE FORMAL DEFINITION

The super-resolution application suggests a method for reconstructing a high quality image from a sequence of lower -resolution images. The problem can be is defined as a construction of a Multi Input Single Output (MISO) system for resolution increment. The MISO system has an input of multiple frames, which can be taken by a video camera or still image camera. The output is a single image with higher resolution than the original frames. The MISO problem can be extrapolated to a Multi Input Multi Output (MIMO) problem such as super resolving a video sequence, in which the LR frames are put into consecutive windows of high-

resolution frames of a HR video. During this study, we will discuss the MISO problem rather than the MIMO counterpart.

Formally, the super-resolution image reconstruction can be represented as follows Let f denotes the time-varying virtual image of the scene in the image plane coordinate system [2.1]. Given a sequence g of P low-resolution, typically noisy and undersampled images, acquired by imaging of the scene $f(x_1, x_2, t)$ at times $t_1 \leq t_2 \leq ... \leq t_p \leq ... \leq t_{P'}$ [2.2].

$$f(x_1, x_2, t), x_1, x_2, t \in \Re, \qquad\qquad\qquad\qquad [\textbf{2.1}]$$

$$g[m_1, m_2, p]; m_1 \in \{1, 2, ..., M_{1p}\}, m_2 \in \{1, 2, ..., M_{2p}\} \text{ and } p \in \{1, 2, ..., P\} \qquad [\textbf{2.2}]$$

The objective is to form S estimates $\hat{f}[n_1, n_2, s]$ of $f(x_1, x_2, \tau_s)$ on the discrete sampling grid at the arbitrary time instants $\tau_1 \leq \tau_2 \leq ... \leq \tau_s \leq ... \leq \tau_{PS}$ [2.3].

$$\hat{f}[n_1, n_2, s], \ s \in \{1, 2, ..., S\}, \ n_1 \in \{1, 2, ..., N_{1s}\} \text{ and } n_2 \in \{1, 2, ..., N_{2s}\} \qquad [\textbf{2.3}]$$

$$N_{1s} > M_{1p}, N_{2s} > M_{2p}, \forall p, s \text{ and } S > P. \qquad\qquad\qquad [\textbf{2.4}]$$

Superresolution refers to the restoration of a sequence of images $\hat{f}[n_1, n_2, s]$ that has information content beyond the spatial and/or temporal band limit of the imaging system [12].

The problem definition above summarizes the direct problem of imaging process. We need to reverse this task to obtain a high-resolution view of the real-world scene, which is an inverse problem with ill-posed properties. It is an inverse problem because SR process is aimed to invert the image-capturing task, which is the acquisition of images of real world by using a limited and non-linear imaging environment. SR approach is also an ill-posed problem because the number of low-resolution images is limited and since we cannot see every point on the scene through these images; this is a direct reason of information loss.

2.3. IMAGE ACQUISITION MODEL

The SR image enhancement is an ill posed inverse problem. The solution to the problem is not unique. We need to understand the imaging process, before attempting to invert it. This inverting process requires a modeling of the relation between the high and the low-resolution images, at its first step.

The acquisition of an image has many details to consider. For example, optical distortions through the optics of the camera, aliasing effect inside the sensor, blurring caused by the unwanted camera shaking and scene motion, additional noise through every part of the pipeline plus the undersampling of the camera make the captured images suffer from spatial resolution loss.

We will refer y_k, where $k=1...p$ as the p low-resolution images and x as the real world high-resolution observation that we try to reach as close as possible at the end of the process. During the observation of the scene, assume that x remains constant. By this way, all of the p observations are of the same scene. All of the differences between low-resolution images are due to varying imaging conditions of the camera. In addition, the unknown noise is always present on all of the LR images. As a result, we will have p different observations of x. This model of observation can be represented as:

$$y_k = D\, B_k\, M_k\ x + n_k \quad for \quad k=1,2...,p \tag{2.5}$$

Where M_k is a transformation matrix, which transforms *x* in vertical and horizontal shifts and scale variances as well as rotational motions in all 3D coordinate axes. B_k is the blur matrix that can be a result of optical disorder, fast motion, point spread function (PSF) of the sensor etc. *D* is the subsampling matrix that is the cause of the loss in spatial resolution. In addition, n_k represents the noise, which is present at all, times. (Figure 2.1)
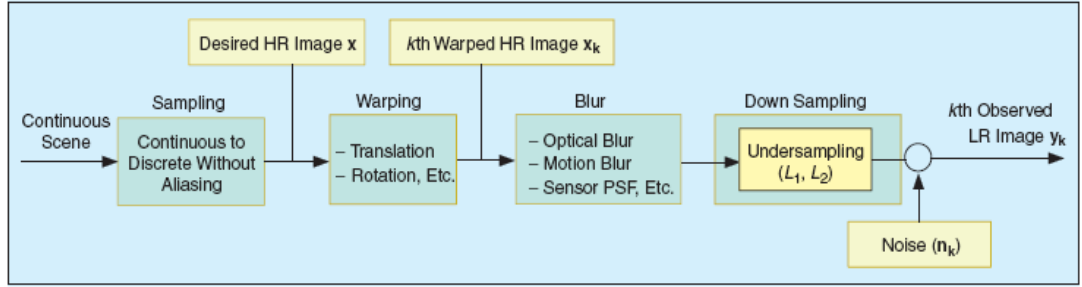
Figure 2.1 The observation model relating LR images to HR counterparts [16]

Alternatively, the observation model can be simplified to sum up all the effects in a single operator to make it easier to visualize the concept. This is possible if these models are unified in a simple matrix-vector form since the LR pixels are defined as a weighted sum of the related HR pixels with additive noise. As a result, equation [2.6] can be expressed as follows.

$$y_k = W_k\, x + n_k \quad for \quad k=1,2...,p \qquad [\textbf{2.6}]$$

Where *Wk* is the effects of the blurring, subsampling and transformations takes place on the original high-resolution pixels of *x*. Again, *nk* is the additive noise coming from the environment.

2.4. SUPER-RESOLUTION APPROACH

Image restoration is a well-defined concept of visually improving the quality of a single image. It focuses on the cancellation of the effects, which take place during the image acquisition process. For instance, deblurring operations can overcome blurring caused by an optical system, relative motion between the imaging environment and the scene and the PSF of the sensor. As well as deblurring, the denoising methods are used to cancel or at least minimize the effects of unwanted noise as much possible. However, neither of these image restoration methods is able to increase the spatial resolution of the images.

For increasing the size of the image, many interpolation techniques are extensively researched and there are a number of well-defined interpolation methods; but in fact, the information loss is unrecoverable and there is no way to find out the lost pixel values. One can have some estimates of the lost values through a distribution function on the image but single frame interpolation techniques are not enough to recover the lost high frequency terms that are lost during the downsampling operation. As the aim is to recover the details of the original scene successfully, there is the need of acquiring and fusing different information of the same scene. Nevertheless, without the application of the image restoration and interpolation methods, the methods of superresolution are broadly understood to mean bandwidth extrapolation beyond the diffraction limit of the optical system [12]. It is possible to say that SR image fusion can be considered as a second-generation technique of image restoration.
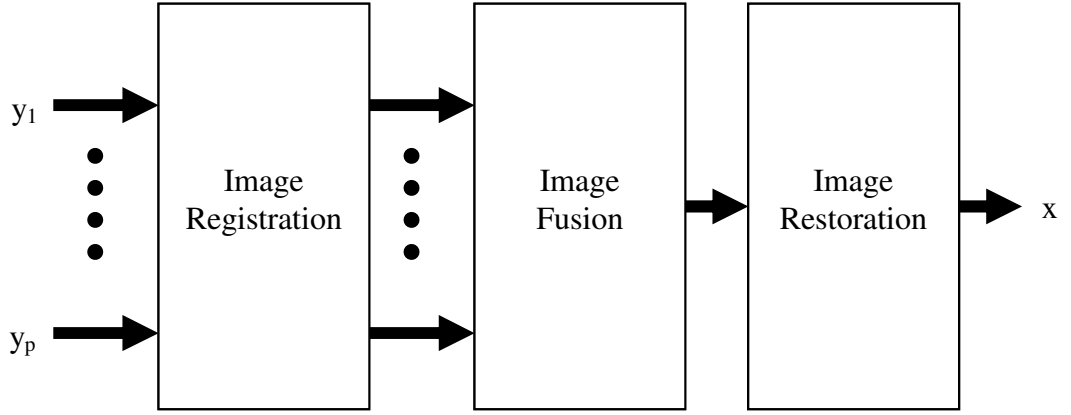
Figure 2.2. Scheme for super resolution

Super-resolution methods have two main parts to recover the lost terms of an image. First, the images at hand should be aligned to register every pixel value to the position of reference and then these information bits about the original pixel value are combined to recover the lost parts of the image. Formally, most of the super resolution methods consist of three basic steps to obtain the high-resolution image. As in Figure 2.2, these steps are registration of low-resolution images to a reference grid, fusion of the LR images to a HR image and the restoration of images in which deblurring and denoising methods are used; this is a conceptual classification only, as sometimes some steps are performed simultaneously.

In this thesis, we aim to examine the image registration techniques and major image fusion techniques in detail, so we will examine the results of the registration-fusion scheme prior to the application of a suitable image restoration technique.

Lastly, super-resolution methods critically depend on the accuracy of image registration. The only way to improve the resolution is the correct utilization of subpixel shifts between images. Without subpixel accuracy, we only have shifted copies of the same image, which gives no extra information to recover the lost information. For the experimentation purposes, some artificial images are generated by sub-pixel shifts. In addition we have used a video camera to capture a series of pictures, giving us a sufficiently rich sets of sub-pixel shifts between images The

camera should not be static during the image acquisition or at least the shots of the scene should be taken from different locations. If these conditions are satisfied, at the end of the image registration step, we will have sub-pixel shifted images of the same scene as in Figure 2.3 to be used in fusion algorithms.
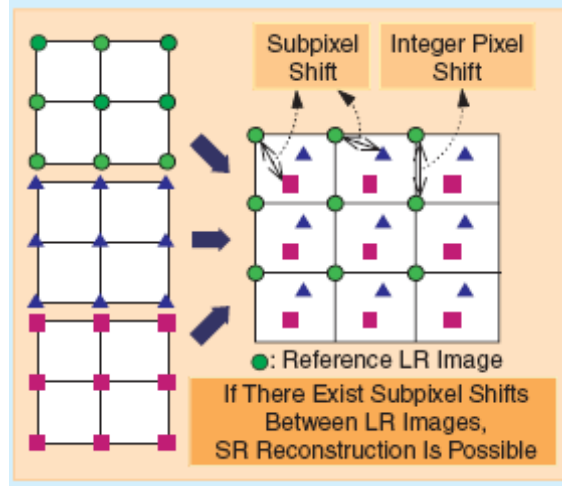
Figure 2.3. Subpixel shifts are vital [16]

During the study, we need to undo the effects of the spatial transformations occurring on the LR images, either to compensate the motion between frames, or to align some features. For this purpose, we discuss the planar motion of camera or imaging plane in the next section.

2.5. PLANAR TRANSFORMATIONS

Spatial transformation is the process of transforming an image into another image in the spatial domain by using a mapping function. Before the application of image registration, images captured by different sensors from different viewpoints at different instants are distorted with respect to each other. Image registration is concerned with the alignment of the image over the same grid. The images to be registered have a mapping function to the reference image. Spatial transformations are applied to the images using these mapping functions to align the images to a reference.

Throughout the image registration process, spatial transformations are extensively used. First, if there is a search for the distortion of the image, possible transformations are applied to the image and compared to the reference, by minimizing the spatial error iteratively, the distortion is found. In addition, after finding the distortion, to compensate its effects we apply the mapping by spatial transformations. This mapping is the transformation matrix and its matrix is called the homography matrix. The term homography is used for the planar transformation matrices of the images, which is in fact our main point of interest.

Homography matrix is a two-way guide for both the reference image and the input images. By applying the homography to the input image, the input image is transformed into the reference image space and by applying the homography inversely to the reference image, reference image is aligned to the input image space (Figure 2.4).
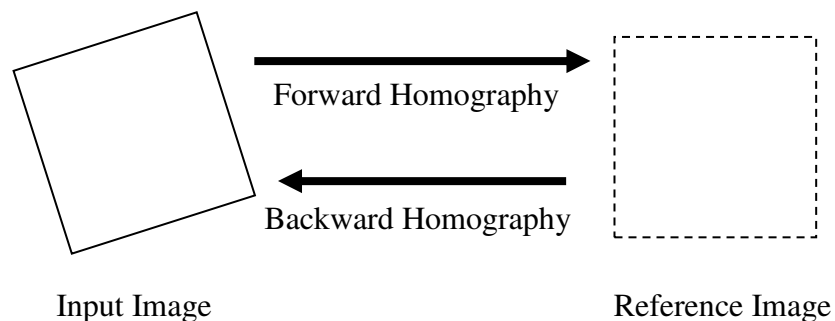
Figure 2.4 The forward and backward homography

Homography matrices include the transformations in eight degrees of freedom at most to represent spatial transformation in 2D space.

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \qquad \begin{bmatrix} Sx' \\ Sy' \\ S \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$
\begin{aligned}
x' &= \frac{xh_{11} + yh_{12} + h_{13}}{S} \\
y' &= \frac{xh_{21} + yh_{22} + h_{23}}{S} \\
S &= xh_{31} + yh_{32} + 1
\end{aligned}
\qquad [\textbf{2.7}]
$$

where x' and y' represents the transformed coordinates and x, y are the original coordinates of the pixels.
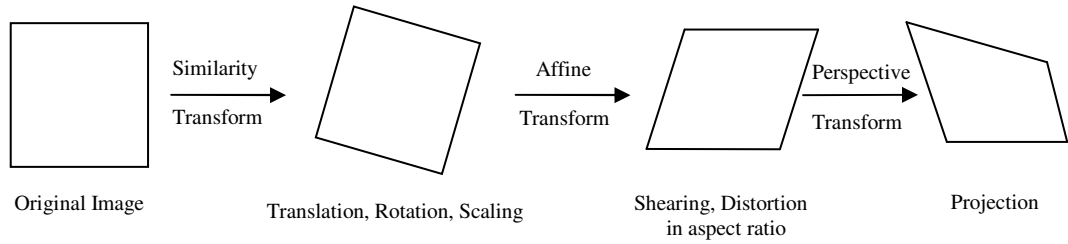
Figure 2.5 2D Image Transformations

For four degrees of freedom in the homography, the transformation is called similarity transform that contains the rotation (θ), translation (dx, dy) and scaling (S).

$$\begin{bmatrix} Sx' \\ Sy' \\ S \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & dx \\ \sin(\theta) & \cos(\theta) & dy \\ 0 & 0 & S \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \begin{aligned} x' &= \frac{x\cos(\theta) + y\sin(\theta) + dx}{S} \\ y' &= \frac{y\cos(\theta) + x\sin(\theta) + dy}{S} \end{aligned} \qquad [\textbf{2.8}]$$

14

With six degrees of freedom in the homography, the transformation is affine transform, which is still linear and preserves straight lines in the image with shearing angle (ø) and aspect ratio (A) (No similarity transform in the formulation below).

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & \tan(\phi) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \begin{aligned} x' &= Ax + y\tan(\phi) \\ y' &= y \end{aligned} \qquad [\textbf{2.9}]$$

and for a full homography, the transformation is perspective transform in which the flat scene is deformed (No similarity or affine transform in the formulation below).

$$\begin{bmatrix} Zx' \\ Zy' \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ C & D & 1 \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \begin{aligned} x' &= \frac{x}{Cx + Dy + 1} \\ y' &= \frac{y}{Cx + Dy + 1} \end{aligned} \qquad [\textbf{2.10}]$$

The captured or created frames that we are considering here are mostly the still scenes with no projective distortion and with minimal local motion in it. Therefore, we will discuss mostly the planar global similarity transforms though this study.

2.6. IMAGE REGISTRATION

Image registration is the method of aligning multiple images on the same grid. Registering frames of a video or images from a sequence is mainly about solving the problem of geometric relation with the reference image and finding the right way to put them on the same geometrical grid. It is the key step in all image analysis tasks in which the desired information is related to some motion in the picture or the camera. Image registration is extremely important in super-resolution

scheme since the artifacts caused by an incorrectly aligned image are more disturbing than the blurring effect caused by interpolation of only one image.

Image Registration Algorithms considered in this thesis classified in two main groups as follows:

1. Spatial Domain Image Registration Techniques
   - Random Sample Consensus (RANSAC) Algorithm discussed by Capel and Zisserman [3] and by Fischler and Bolles [9] that focuses on feature matching
   - Taylor series expansions method discussed by Keren et al [4].

2. Frequency Domain Image Registration Techniques
   - Low frequency image matching method by Vandewalle et al.[1,2].
   - A log-polar based phase correlation method discussed by Marcel et al. [7] [1].
   - A noise-robust Cartesian coordinate frequency domain technique discussed by Luchesse et al.[8] which describes the rotation in a different manner than the other frequency domain methods.[1].

We have implemented all of the mentioned methods. All these methods are implemented and used. The aim is to generate a result that finds out a fast, robust and competent method to register images precisely. At the end of the image registration procedure, a homography matrix will be at hand for each input image that aligns them into the reference image. This homography matrix will be used in subsequent steps of super resolution. Following the registration discussion, we give the image registration algorithms used in this thesis. Then we compare the registration methods by experiments on artificially generated images. During the generation of these synthetic images, we implement different the transformation parameters (rotation, translation and scaling) so that we can understand shortcomings of different methods.

2.6.1. RANDOM SAMPLE CONSENSUS ALGORITHM

Spatial domain image algorithms rely on the fact that images to be aligned have some common points to pair. Some of the methods of spatial domain use direct relations between images such as intensity values of a block of pixels and seeking these blocks on both images to find correspondences. On the other hand, some algorithms are based on some interest points to match between images.

As discussed by Capel and Zisserman [3] and by Fischler and Bolles [9] using RANSAC methodology, one can find correspondences by automatic detection and analyze these features among the images (Figure 2.6). Typically, in each image several hundred "interest points" are automatically detected with sub-pixel accuracy using an algorithm such as the Harris feature detector[6]. Putative correspondences are identified by comparing the image neighborhoods around the features (Figure 2.7).

A robust search algorithm such as RANSAC extracts a consistent homography of these correspondences. Finally, these correspondences are optimized and a very accurate estimate of the homography is found (Figure 2.6).

Detect features of interest using Harris corner detector.	→	Compute a set of putative correspondences between images.	→	Using RANSAC, to estimate homography.
After the results are stable Homography is finalized.	←	Use the optimized homography to search further interest points.	←	Optimize the homography matrix by including all the inliers.
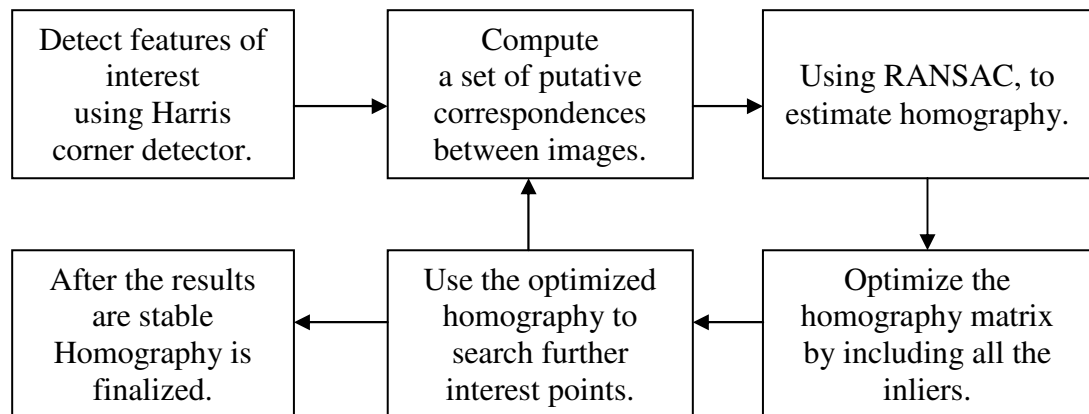
Figure 2.6 Pipeline of the RANSAC algorithm
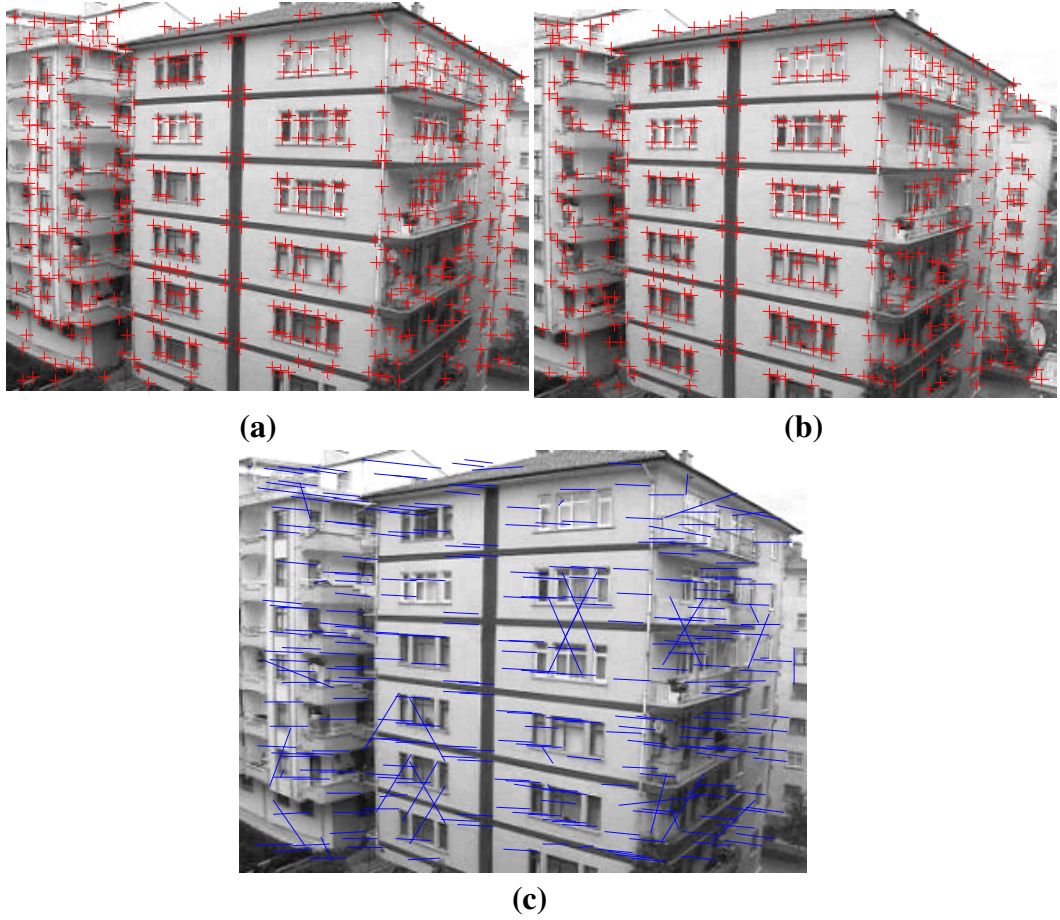
Figure 2.7 The working principle of RANSAC algorithm: (a) corners in the 1st frame, (b) corners in the 2nd frame, (c) match the corresponding interest points

2.6.2. KEREN ALGORITHM

Keren et al algorithm is a very efficient and straightforward image registration method. It simply uses the Taylor series expansion of the spatial transformation. For the two images f_1 and f_2, there exists a horizontal shift "a" and vertical shift "b" and the rotation angle around the origin θ:

$$g(x, y) = f(x.\cos(\theta) - y.\sin(\theta) + a, \; y.\cos(\theta) - x.\sin(\theta) + b) \qquad [\mathbf{2.11}]$$

As we expand $\sin(\theta)$ and $\cos(\theta)$ to the first two terms in their Taylor series, we will get:

$$g(x,y) \approx f(x + a - y.\theta - x.\theta^2/2,\ y + b - x.\theta - y.\theta^2/2) \qquad \textbf{[2.12]}$$

Expanding f to the first term of its own Taylor series gives the first order equation:

$$g(x,y) \approx f(x,y) + (a - y.\theta - x.\theta^2/2).\frac{\partial f}{\partial x} + (a + x.\theta - y.\theta^2/2).\frac{\partial f}{\partial y} \qquad \textbf{[2.13]}$$

The error function is then:

$$E(a,b,\theta) = \sum \left[f(x,y) + (a - y.\theta - x.\theta^2/2)\frac{\partial f}{\partial x} + (a + x.\theta - y.\theta^2/2)\frac{\partial f}{\partial y} - g(x,y) \right]^2 \qquad \textbf{[2.14]}$$

where the summation is in the overlapping part of the images f and g.

If we look for the minimum of E by computing its derivatives by a,b and θ and comparing them to zero, then after neglecting the non-linear terms and some small coefficients we get the following system of linear equations, where the summation is over the overlapping area:

To estimate a,b and θ precisely we need to apply the iterative process of updating g with the accumulated values of rotational and translational parameters where the reference frame is f and always the same.

$$\left[\sum\left(\frac{\partial f}{\partial x}\right)^2\right]a + \left[\sum\left(\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\right)\right]b + \left[\sum R\left(\frac{\partial f}{\partial x}\right)\right]\theta = \sum\frac{\partial f}{\partial x}(f-g)$$

$$\left[\sum\left(\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\right)\right]a + \left[\sum\left(\frac{\partial f}{\partial y}\right)^2\right]b + \left[\sum R\left(\frac{\partial f}{\partial y}\right)\right]\theta = \sum\frac{\partial f}{\partial y}(f-g) \qquad [2.15]$$

$$\left[\sum R\left(\frac{\partial f}{\partial x}\right)\right]a + \left[\sum R\left(\frac{\partial f}{\partial y}\right)\right]b + \left[\sum R^2\right]\theta = \sum R(f-g)$$

$$R = x\frac{\partial f}{\partial y} + y\frac{\partial f}{\partial x} \qquad [2.16]$$

2.6.3. VANDEWALLE ALGORITHM

Frequency Domain methods of image registration are mainly based on three principles:

- Shifting property of the Fourier transform (FT)
- Aliasing relationship between continuous FT of HR image and the DFT of LR images
- Band limited HR images

Vandewalle et al. looks from a different perspective to the problem. The algorithm prefers to use not only the whole frequency spectrum of the image but the low frequency region of the spectra, where the signal to noise ratio is highest and aliasing is minimal. The four low-resolution images are necessarily undersampled. Otherwise, our algorithm is not able to reconstruct a better image as it uses exactly this undersampled information [2].

The motion estimation is done in two steps. First, the rotation is recovered and then the translations between images are found. This process is accomplished due to some properties of Fourier transform. These properties are as follows:

The Translation Property: Shifts in spatial domain cause a linear shift in the phase component. That is, the magnitude components of Fourier transformation do not affected by linear shifts in spatial domain.

$$f(x_1 + a, x_2 + b) \xleftrightarrow{FT} F(k_1, k_2) \exp[-j(ak_1 + bk_2)] \tag{2.17}$$

Rotation Property: Rotating the image through an angle θ in the spatial domain causes the Fourier representation to be rotated through the same angle.

$$f(x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta)$$
$$\xleftrightarrow{FT} F(k_1 \cos\theta - k_2 \sin\theta, k_1 \sin\theta + k_2 \cos\theta) \tag{2.18}$$

According to the pipeline, after getting the Fourier Transform of the image f(x) we have, F(u), and when F(u) is transformed into polar coordinates we will have F(r; θ) at hand, the frequency content h is computed as a function of the angle by integrating over radial lines:

$$h(\alpha) = \int_{\alpha - \Delta\alpha/2}^{\alpha + \Delta\alpha/2} \int_0^{\infty} |F(r,\theta)| .dr.d\theta \tag{2.19}$$

In practice, |F(r; θ)| is a discrete signal. Therefore, we compute the discrete function h(α) as the average of the values on the rectangular grid that have an angle:

$$\alpha - \frac{\Delta\alpha}{2} < \theta < \alpha - \frac{\Delta\alpha}{2} \tag{2.20}$$

21

As we want to compute the rotation angle with a precision of 0.1 degrees, h(α) is computed every 0.1 degrees. To get a similar number of signal values, |F(r; θ)| at every angle, the average is only evaluated on a circular disc of values for which r < ρ (where ρ is the image radius, or half the image size). Finally, as the values for low frequencies are very large compared to the other values and are very coarsely sampled as a function of the angle, we discard the values for which r < ε. ρ, with ε = 0:1. Thus, h(α) is computed as the average of the frequency values on a discrete grid with

$$\alpha - \frac{\Delta \alpha}{2} < \theta < \alpha - \frac{\Delta \alpha}{2} \quad \text{and} \quad \varepsilon \rho < r < \rho \qquad \text{[2.21]}$$

This results in a function h(α) for both |F1(u)| and |F2(u)|. The exact rotation angle can then be computed as the value for which their correlation reaches a maximum [2]. Just as we recover rotation and cancel its effect by rotating the image in the reverse direction, we will find the vertical and horizontal shifts of the images. This is in practice rather simple and by using the translation property of Fourier Transform. It is well known that the shift parameters Δx can thus be computed as the slope of the phase difference $\angle (F_2 (u) / F_1 (u))$.

After applying these terms to the images, we will have the rotation angle and vertical, horizontal shifts at hand. To use these values in our later work we will have to transform them into the homography matrix, which is in fact quite simple.

2.6.4. LUCHESSE ALGORITHM

Lucchese and Cortelazzo [8] developed a rotation estimation algorithm based on the property that the magnitude of the Fourier transform of an image and the mirrored version of the magnitude of the Fourier transform of a rotated image has a pair of orthogonal zero-crossing lines. The angle that these lines make with the axes is

equal to half the rotation angle between the two images. The horizontal and vertical shifts are estimated afterwards using a standard phase correlation method.

2.6.5. MARCEL ALGORITHM

Most of the frequency domain registration methods are based on the fact that two shifted images differ in frequency domain by a phase shift only, which can be found from their correlation. Using a log-polar transform of the magnitude of the frequency spectra, image rotation and scale can be converted into horizontal and vertical shifts. These can therefore also be estimated using a phase correlation method. Reddy and Chatterji [19] and Marcel et al. [7] describe such planar motion estimation algorithms.

2.6.6. COMPARISON OF REGISTRATION METHODS

All of the image registration methods mentioned in the previous section is the results of some prior studies. All of them have their own impregnability and frailties. Throughout this section, we will compare the results of some experiments on the image registration methods. First, the test methodology will be examined, and then the test results will be considered.

2.6.6.1. TEST METHODOLOGY

The best way of surveying a number of methods is to carry out some experiments on them. In this study, we use synthetic images with known transformation parameters. As the motion parameters are fixed throughout the image, we expect the image registration methods estimate these parameters as close as possible to the real values. Two images are used to generate synthetic LR images during this study. These are the famous "Lena" image (Figure 2.8(a)) and a test pattern (Figure 2.8(b)) for resolution assessment [2].

(a) **(b)**
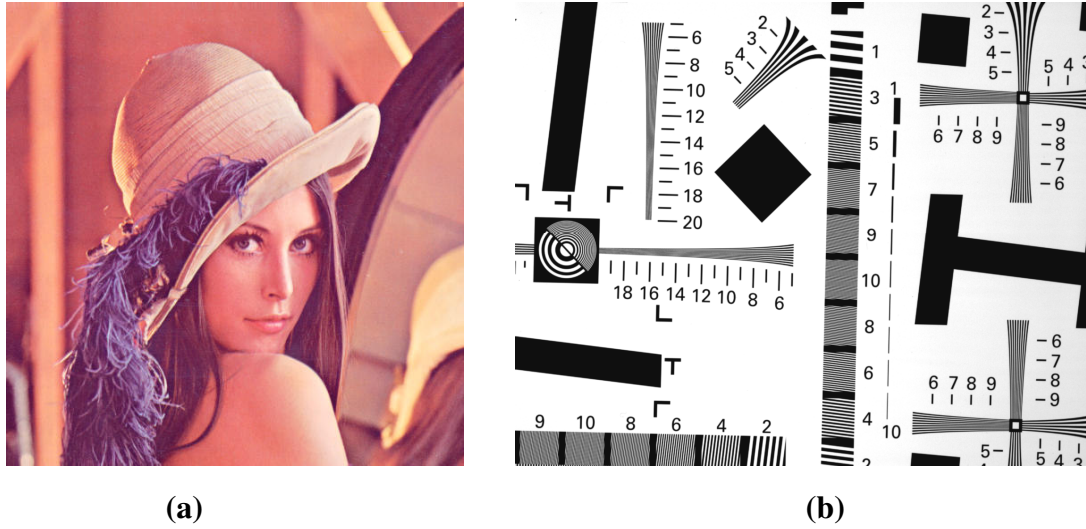
Figure 2.8 Lena (a) and Reschart (b) are used in examples of methods

The following transformations are applied to the images (Figure 2.9):

- Pure Rotation
- Pure Pixel
- Sub-pixel Translation
- Transrotations
- Transrotations with scaling
- Transrotations with additive noise (20dB)

Every condition is evaluated for the maximum number of the methods explained in the previous sections. Scaling transformation is only assessed for RANSAC algorithm since other methods can only deal with rotation and translation (transrotation) by design. Noisy transrotation examination is applied to multi parameter variance case (only to transrotations), but not to pure rotation or pure translation.

Figure 2.9 (a) pure rotation, (b) pure translation (c) pure translation with subpixel shifts, (d) transrotation, (e) scaling with transrotation (f) transrotations with noise

2.6.6.2. TESTS OF IMAGE REGISTRATION

At the end of the tests, the ultimate goal is to identify the best image registration algorithm among these five algorithms explained in this thesis. This identification will lead us to the use of the "best" method in our image reconstruction algorithms in the super-resolution stage. Throughout the tests, all algorithms except RANSAC are deterministic. RANSAC may find close but different motion estimates since the algorithm is probabilistic, so the first trial for RANSAC is saved for the analysis. As a result, it is advantageous to run the RANSAC algorithm with noisy data and average out the effects of noise on motion estimation.

2.6.6.2.1. PURE ROTATION

For the pure rotational transformation test, the test patterns are created by rotating the HR image at 0°, 2°, 5° and 10° about the middle of the image and downsampling it (i.e. Figure 2.9(a)). The figures shows the rotational errors for all methods, where the real rotation angles are shown in the horizontal axis of the figures.
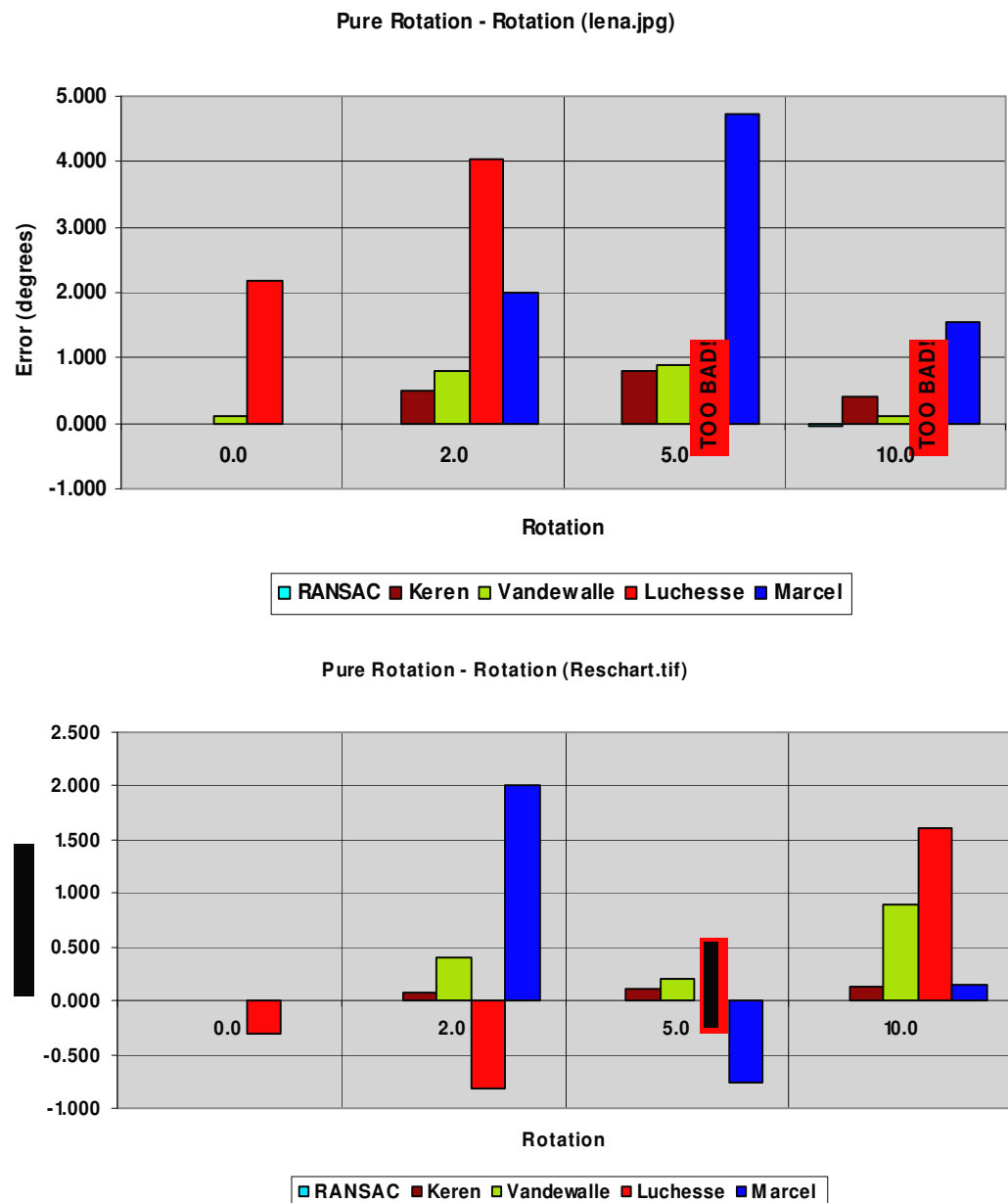
Figure 2.10 Pure rotational motion is tested for both images

Most of the methods, besides Luchesse et al, are able to find a close estimate. As Marcel et al has a greater error value than the remaining three methods (RANSAC, Keren and Vandewalle) which perform satisfactorily. As a result of the test, RANSAC based algorithm has estimated the pure rotation perfectly (Figure 2.10).

2.6.6.2.2. TRANSLATION WITH NO SUBPIXEL SHIFTS

For the pure translational motion, LR images are created by shifting the downsampled images in vertical and horizontal directions by 0, 5 and 20 pixels (i.e. Figure 2.9(b)). The distance between the estimated and actual places of the pixels represents the error value of the results:

$$error = \sqrt{\left[(x-x')^2 + (y-y')^2\right]} \qquad [\mathbf{2.22}]$$

where x, y are the actual values of vertical and horizontal shifts and x', y' are the shift values that are estimated by the corresponding image registration method (Figure 2.11).

Pure Translation - Distance (lena.jpg)

Pure Translation - Distance (reschart.tif)
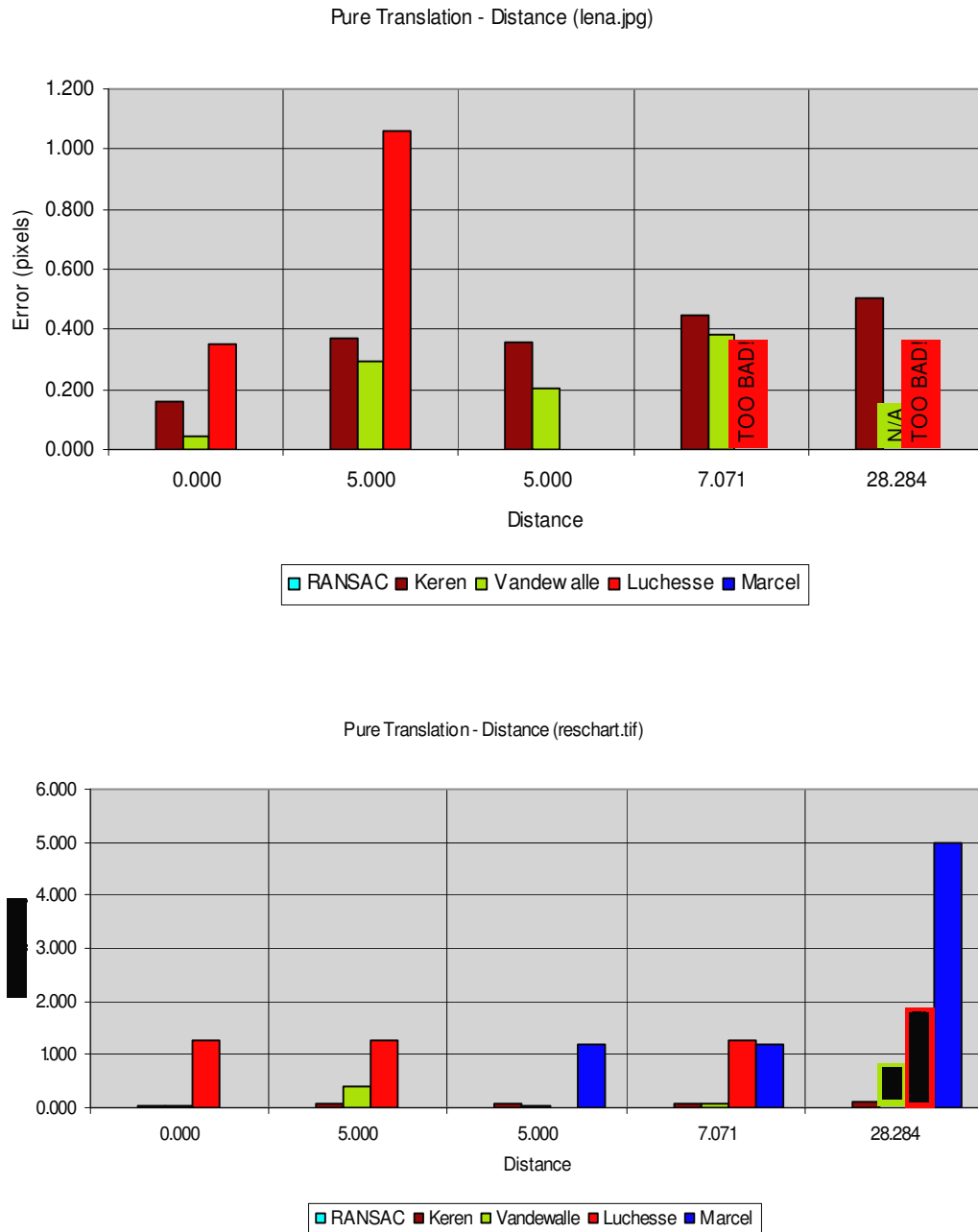
Figure 2.11 Pure translational motion is tested for both images

For all of the methods, the algorithms are able to find a close solution. Lucchese et al. is an unstable method for the Lena image. As Marcel et al has a greater error value in reschart image, it performs impressively in Lena image in which the high and low frequency components are both present; the remaining three methods (RANSAC, Keren and Vandewalle) perform very well in this test.
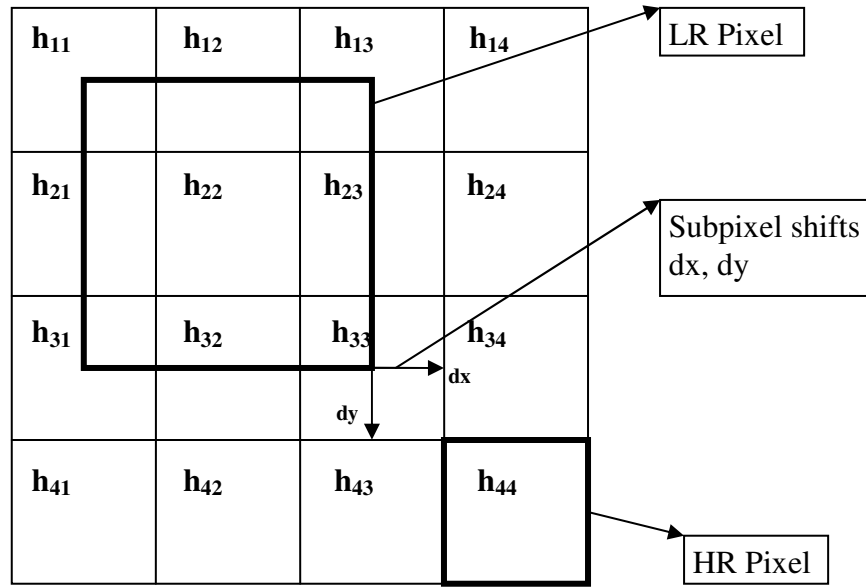
2.6.6.2.3. TRANSLATION WITH SUBPIXEL SHIFTS

 The digital image acquisition is the sampling of the energy reflected by the real world scenes with the imaging sensors. The light coming into the sensor creates an intensity value on every cell of the sensor. When the sensor and the scene moves with respect to each other, the intensity value on every cell changes. If the magnitude of the motion is exactly an integer value then the image taken will shift in integer pixel values as in the previous test. However, if the motion occurs in a decimal level, the imaging cells will share the total energy reflected by the scene in a different dispersion. As a result, the intensity values of the actual LR image and the sub-pixel shifted version of it has a different distribution of the intensities. After direct downsampling of the HR image the LR pixel values are $L_{11} = (h_{11} + h_{12} + h_{21} + h_{22})/4$. If a sub-pixel shift is applied, the LR pixel values become (Figure 2.12):

$$L_{11} = \frac{[h_{11}d_xd_y + h_{12}d_y + h_{13}(1-d_x)d_y] + [h_{21}d_x + h_{22} + h_{23}(1-d_x)] + [h_{31}d_x.(1-d_y) + h_{32}(1-d_y) + h_{33}(1-d_x).(1-d_y)]}{4}$$ [2.23]

The rest of the LR pixels L_{12}, L_{13}… L_{21}, L_{22}… are calculated as the shifted versions of L_{11}.

Figure 2.12 Sub-pixel translation, test image forming

The figure above shows the case with the downsampling level two for both horizontal and the vertical axes.

During this test, 0.3, 0.6 and 0.9 pixel shifts are applied to the images in both horizontal and vertical axes. According to the sub-pixel registration test, with the exception of Luchesse and Marcel algorithms, the remaining methods perform quite impressively to find the sub-pixel shifts. Especially, Vandewalle Algorithm submits outstanding results for sub-pixel level translations (Figure 2.13).

Pure Translation - Subpixel Distance (reschart.tif)

Pure Translation - Subpixel Distance (reschart.tif)
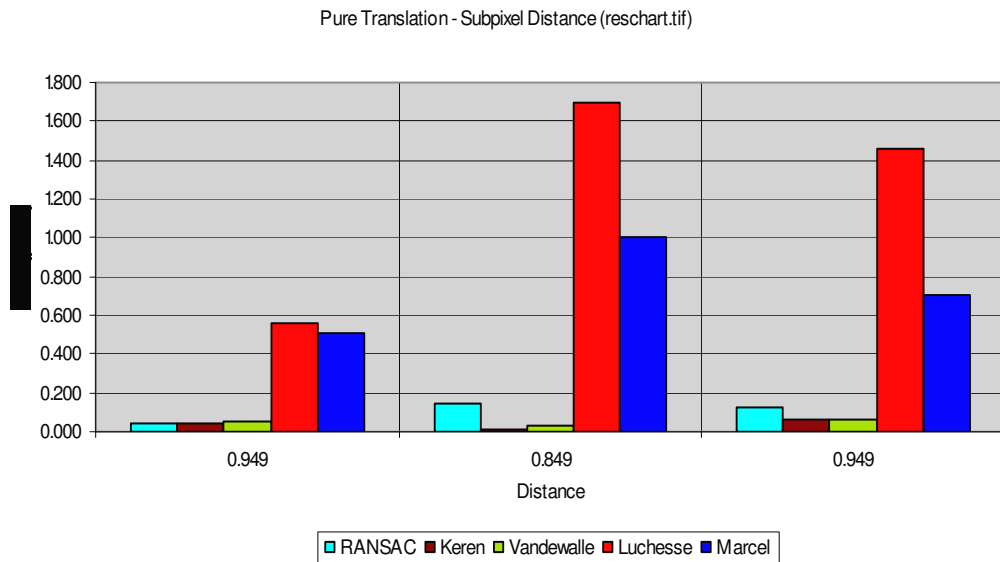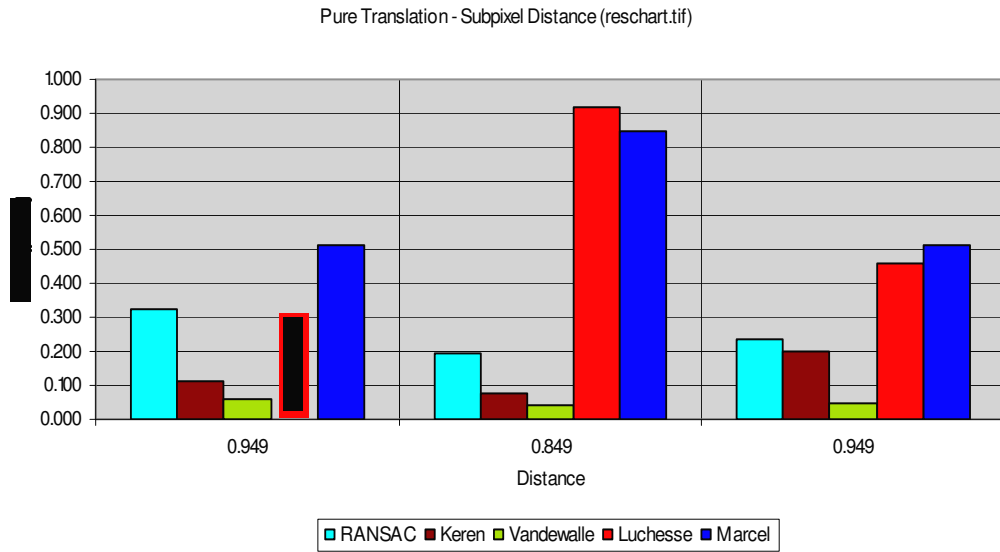
Figure 2.13 Pure sub-pixel translational motion is tested for both images

2.6.6.2.4. TRANSROTATIONS

The combination of a series of rotational and translational parameters is called as the transrotational experiments on registration algorithms. Both the rotation and the translational distance errors are compared here. The translational parameters include both integer and decimal level shifts at once. Rotational translations are the cases where our hand-held camera is taking a series of pictures or a video. Thus, the

31

results of this test are important for our evaluation of which algorithm to use for super-resolution.

Transrotation Distance (lena.jpg)

Distance

RANSAC ■ Keren ■ Vandewalle □ Luchesse ■ Marcel

Transrotation Distance (reschart.tif)

Distance

RANSAC ■ Keren ■ Vandewalle ■ Luchesse ■ Marcel
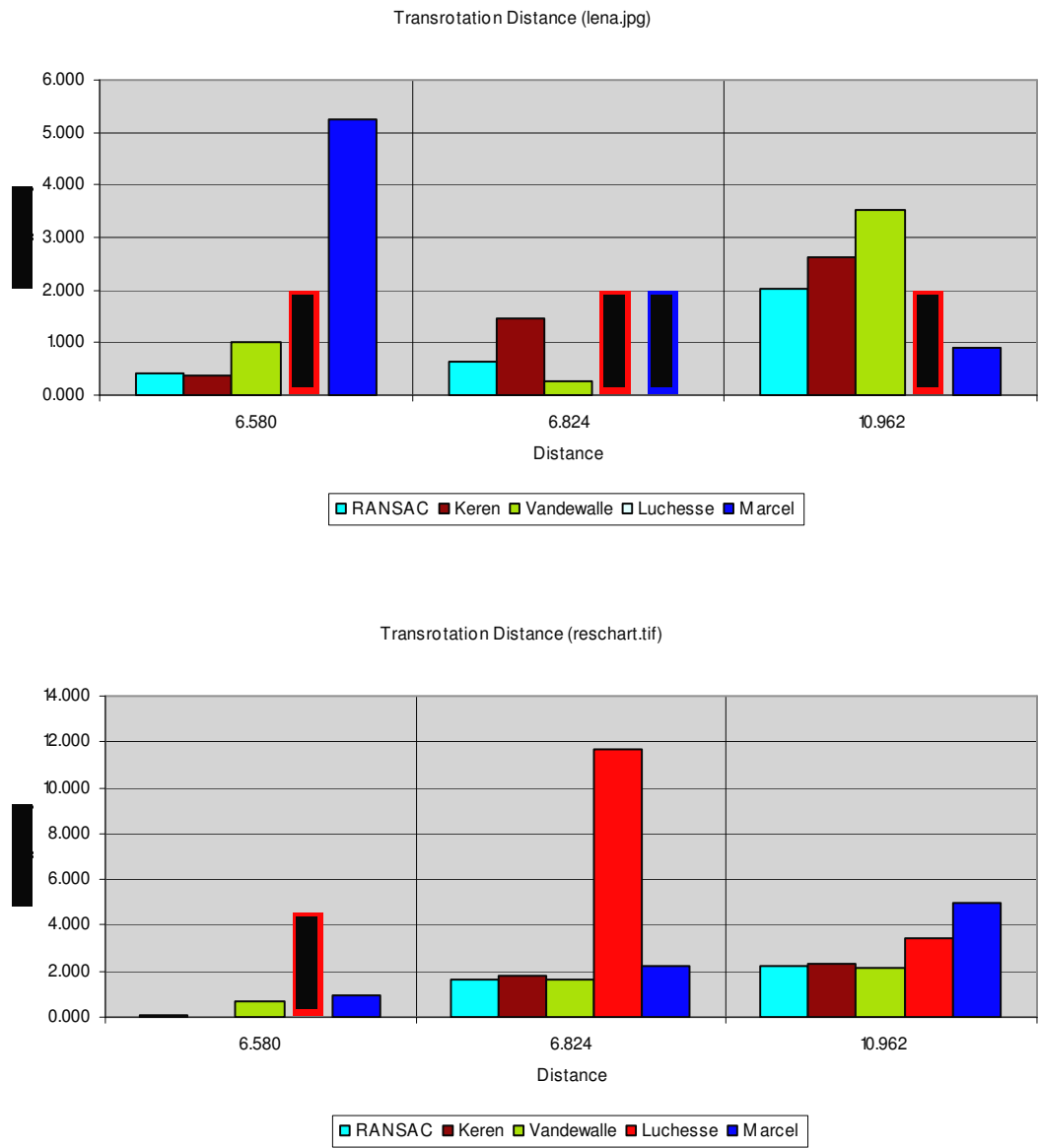
Figure 2.14 Translation Errors in transrotational motion

First, the distance error increases as the translation parameter increases. Once again, the Luchesse performs really badly in most of the cases (especially for the Lena image). As the Marcel's algorithm is not very competitive with respect to the other three methods which perform quite well (Figure 2.14).

Transrotation - Rotation (lena.jpg)

Transrotation - Rotation (Reschart.tif)
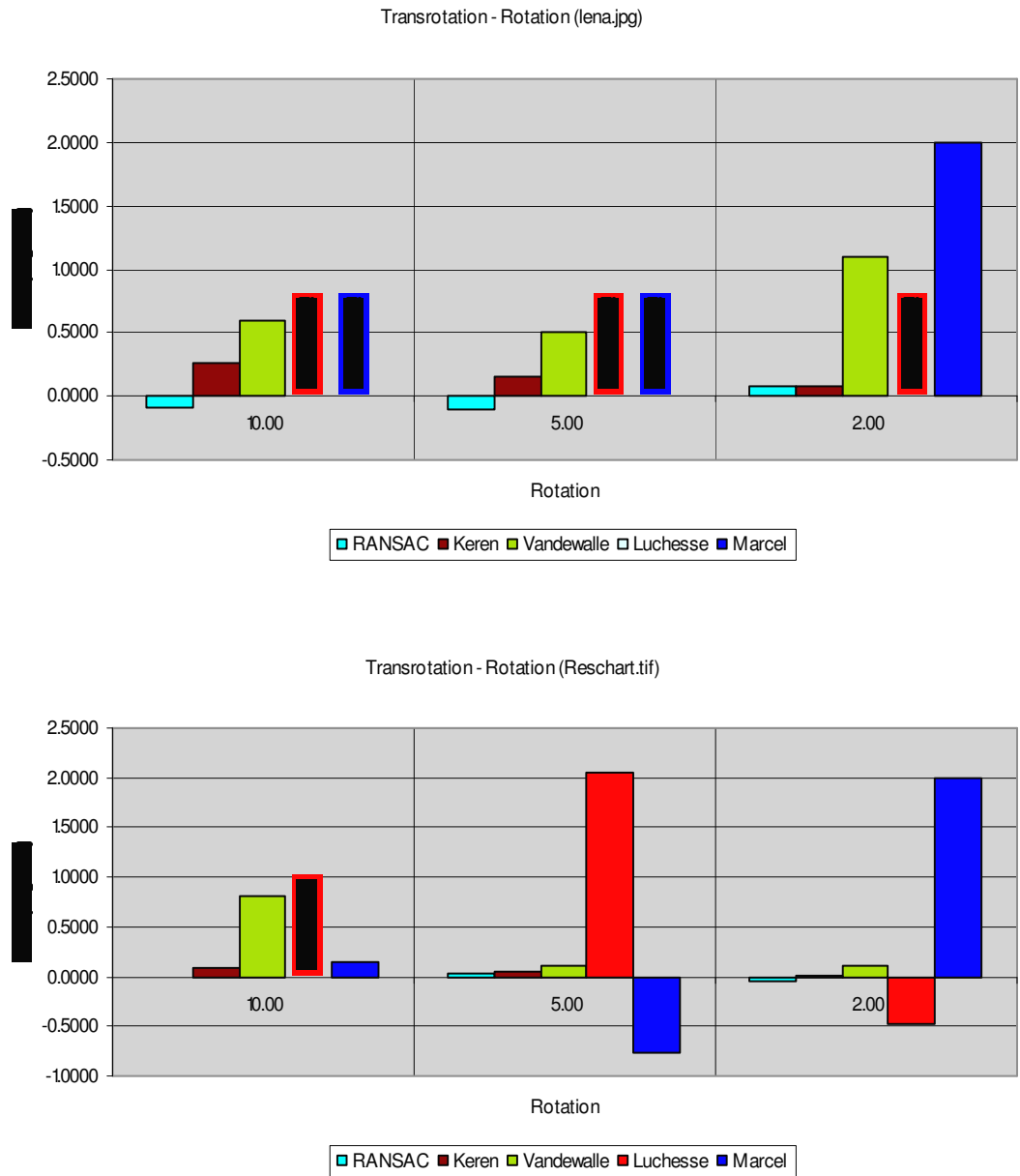
Figure 2.15 Rotation Errors in transrotational motion

33

Additionally, the rotation estimation error has a parallel nature with the distance error behavior of the methods. Again, the RANSAC, Keren and Vandewalle algorithms prove themselves very competent in the area of image registration as the transrotational effects are applied or present on the image series at hand.
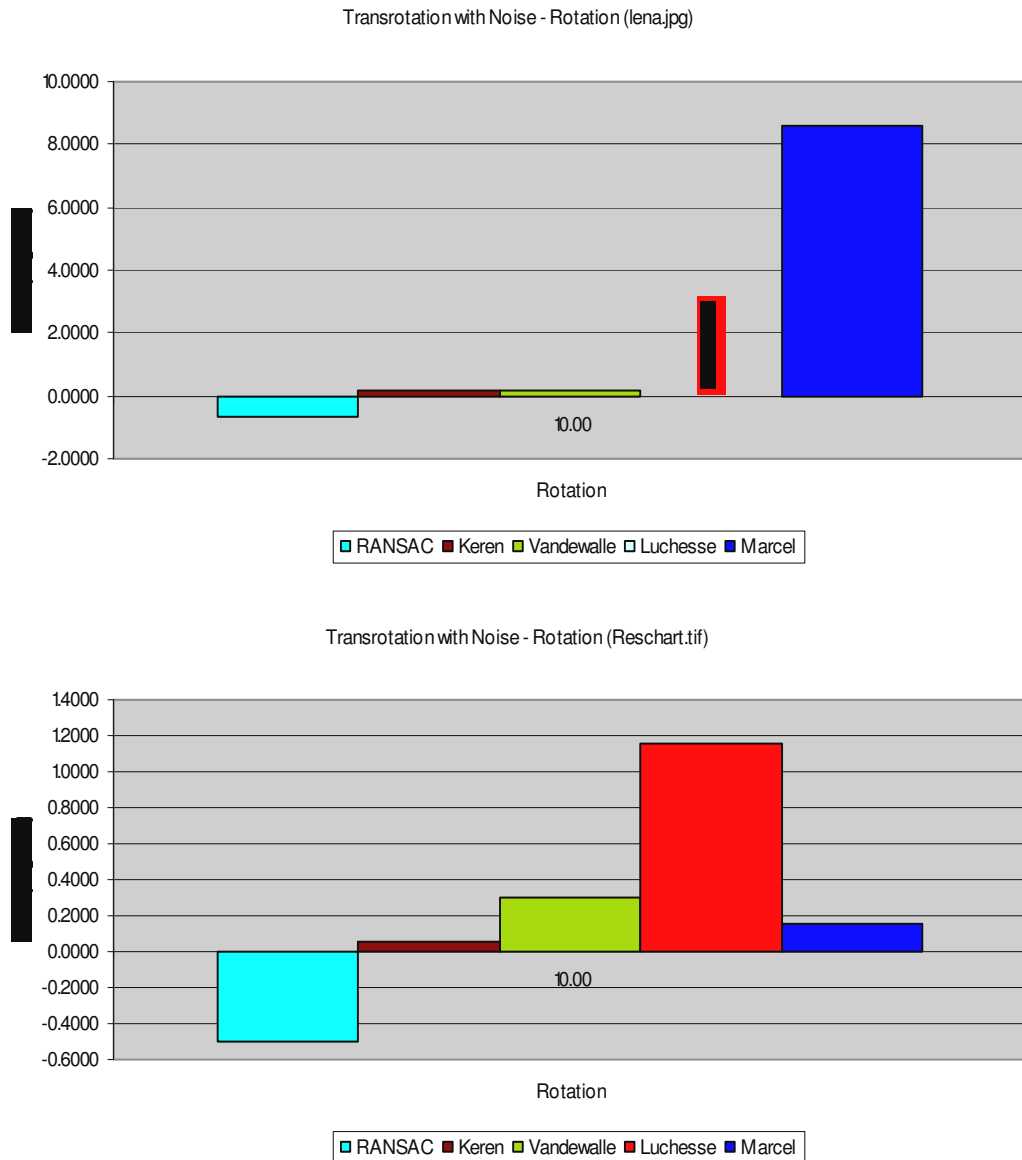
2.6.6.2.5. EFFECTS OF NOISE IN REGISTRATION

The image acquisition is subject to many unwanted effects as it was stated in the previous chapters. The experiments should objectively treat the methods. Therefore, we should also use noise added test patterns of artificial images, to correctly examine the registration methods. The way of adding noise to the test images is applying random noise to decimate the quality of image to a targeted level. The noise is added to the image in a simple way (2.24, 2.25) .

$$Noise = \sqrt{\frac{\sum (LRpixels)^2}{\sum (RandomNoise)^2}} \cdot 10^{\left(-\frac{NoiseLevelIndB}{20}\right)} \cdot RandomNoise \qquad [\textbf{2.24}]$$
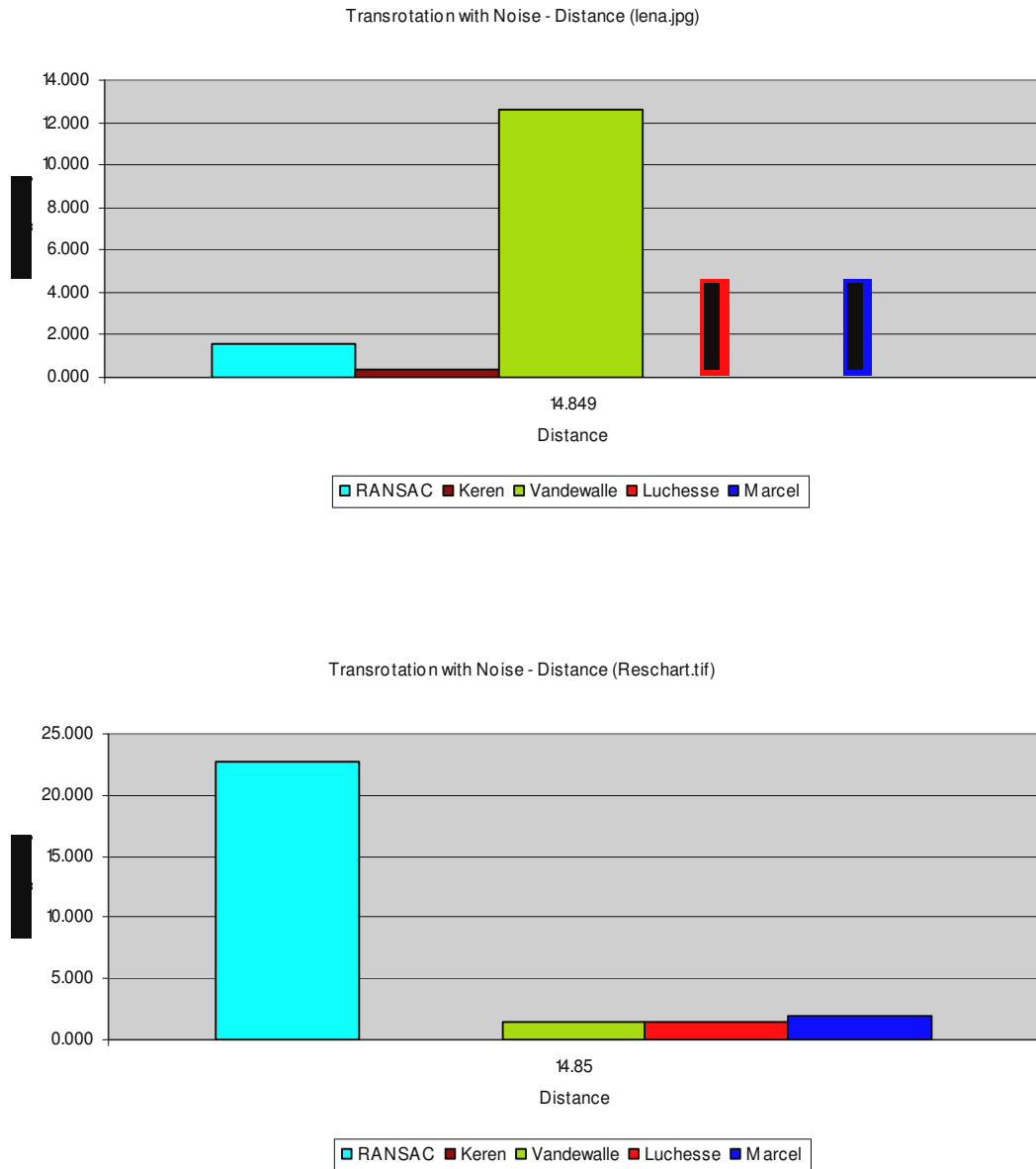
$$Noisy\ Image = Zero\ Noise\ LR\ Image + Noise \qquad [\textbf{2.25}]$$

RandomNoise is a pattern with random numbers between 0.0 and 1.0 at the size of the low-resolution image grid in which the noise is going to be applied. NoiseLevelIndB is the desired noise level of the image at the end (20dB, 10dB etc.). LRPixels are the intensity values for the LR image.

Transrotation with Noise - Rotation (lena.jpg)

RANSAC Keren Vandewalle Luchesse Marcel

Transrotation with Noise - Rotation (Reschart.tif)

RANSAC Keren Vandewalle Luchesse Marcel

Figure 2.16 Rotation Errors in transrotational motion with noise

The results for the rotational errors are in an acceptable level. The values estimated by the RANSAC, Keren and Vandewalle methods are close to the real values. However, for the "Lena image" Luchesse and Marcel methods perform quiet badly. The two methods give good results in the reschart image (Figure 2.16).

Transrotation with Noise - Distance (lena.jpg)

14.849

Distance

RANSAC ■ Keren ■ Vandewalle ■ Luchesse ■ Marcel

Transrotation with Noise - Distance (Reschart.tif)

14.85

Distance

RANSAC ■ Keren ■ Vandewalle ■ Luchesse ■ Marcel

Figure 2.17 Distance Errors in transrotational motion with noise

The distance errors are very different for the two test images at hand. The only method remains unaffected is the spatial domain based algorithm Keren. The other methods end with random results changing from run to run. For the "Lena image", the frequency-based methods end up with incorrect results, but the RANSAC method performs better. In the contrary, the frequency domain methods (Vandewalle, Marcel, and Luchesse) perform better, but the RANSAC method
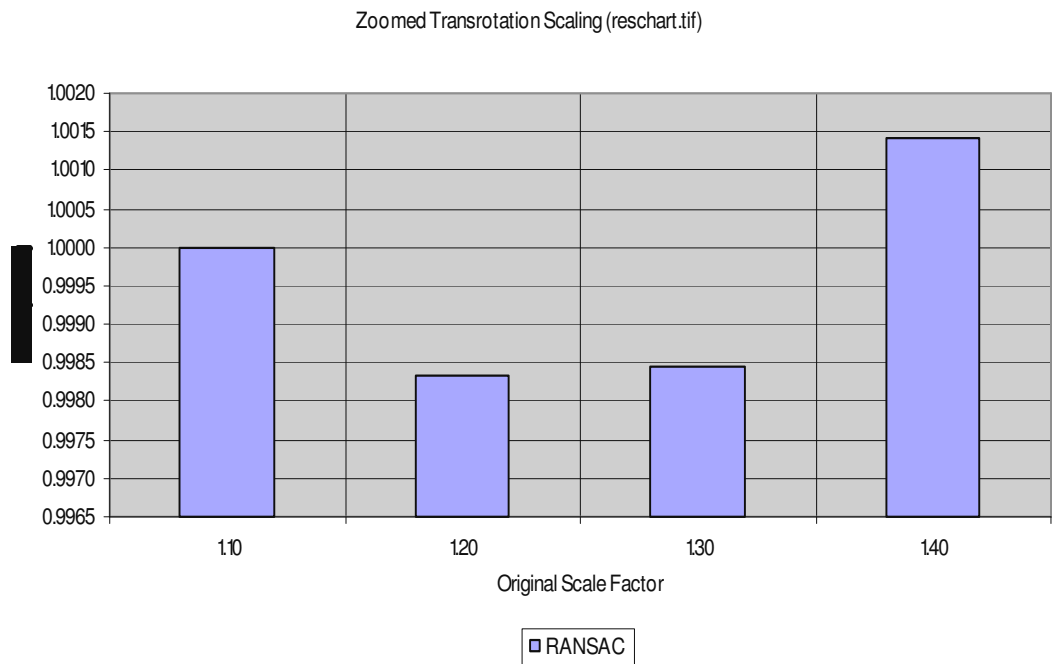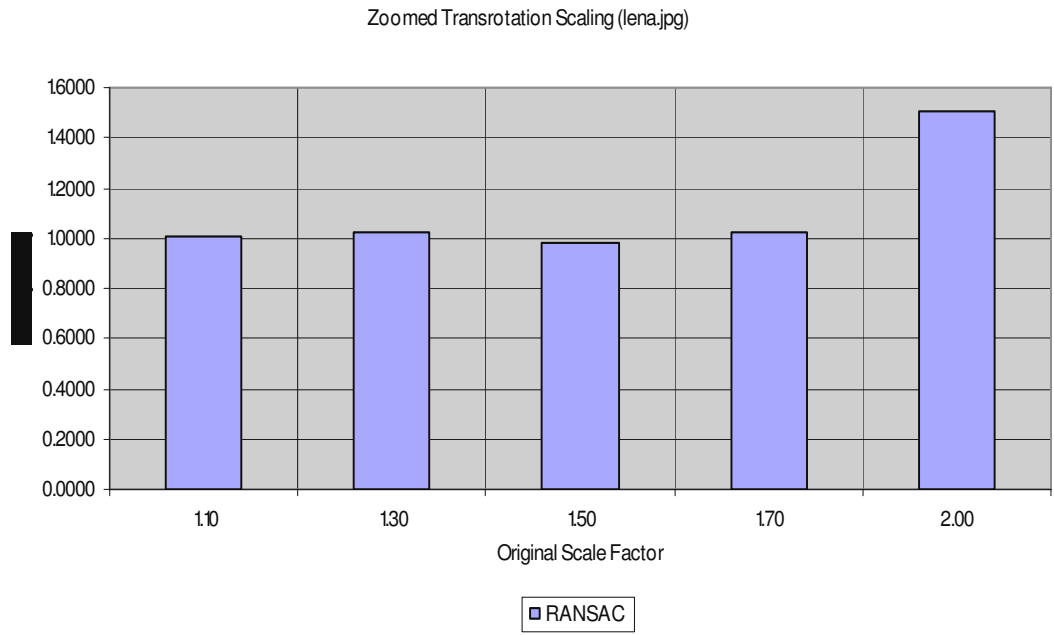
36

totally fails with the reschart image. The reason of the failure of the RANSAC is its failure in the failed detection of the interest points. The reschart image is unpopulated besides the numbers and letters on it and the noise components blocks the detection of image features (Figure 2.17).

2.6.6.2.6. ZOOM (SCALING)

Zoom detection of is beyond the capabilities of the all algorithms except RANSAC. It is because that those algorithms can only cope with planar translation and rotation but not the scaling. Therefore, we will examine the RANSAC algorithm only for scaling estimation, where a maximum of the transrotation parameters applied in the previous tests are also present.

For the reschart image the features present in the image almost vanishes in double scaling so the scaling levels for reschart image is 1.1x, 1.2x, 1.3x, 1.4x where the scaling for Lena image are 1.1x, 1.3x, 1.5x, 1.7x and 2x. The figures below show the rate of the error with respect to the scale level. (Perfect estimation means "1" in the figure 2.18)

The RANSAC algorithm can estimate scale factor successfully. The other registration methods can handle the scaling transformation either, but algorithms need some further study. Beyond some level of scaling, the number of interest points that are used for registering the images, decreases dramatically. Therefore, the registration errors increase proportionally with the increasing level of scaling. The rotational and distance errors of the scaled image are as follows.

Zoomed Transrotation Scaling (lena.jpg)

Original Scale Factor

☐ RANSAC

Zoomed Transrotation Scaling (reschart.tif)

Original Scale Factor

☐ RANSAC

Figure 2.18 Scale Factor Errors in transrotational motion for RANSAC.

Table 2.1 Rotational and Translational Errors for the scaled reschart image

Scaling	Rotation (10 degrees)	Horizontal Shift Error (10.5 pixels)	Vertical Shift Error (10.5 pixels)
1.1x	0,9360	-0,7090	-0,5220
1.2x	1,7000	-0,7840	-1,1810
1.3x	2,3800	-1,0330	-1,1930
1.4x	2,8610	0,2840	-0,4770
1.5x	100.0000	-82.2090	164.4300
1.1x with 20dB noise	-0.4190	-2.0750	-7.8540

Table 2.2 Rotational and Translational Errors for the scaled Lena image

Scaling	Rotation (10 degrees)	Horizontal Shift Error (10.5 pixels)	Vertical Shift Error (10.5 pixels)
1.1x	1,0740	-0,7650	-0,5310
1.3x	2,7940	-1,1130	0,2160
1.5x	3,4960	-0,6590	-2,5620
1.7x	4,2980	-1,2180	1,0400
2.0x	-13,0010	42,1410	-26,4020
1.1x with 20dB noise	0.8700	-2.4830	-3.7090

As it appears in the previous figures of tables on the scaling, the rotational and translational errors are strictly related to the scaling factor estimation. (Table 2.1; 1.5 x case & Table 2.2; 2.0 x cases). Even if the RANSAC scaling estimation gathers incorrect results from some level of zooming, the algorithm copes with most of the cases quite well. In addition, to see the noise's effect on the scaling concept; a 20dB noise is applied, and the scale factor is 1.1x simultaneously. The results of the

trial are affected severely from the noise. For both images, the degradation of performance with respect to noiseless case is obvious.

2.6.7. DISCUSSIONS ON REGISTRATION METHODS

The image registration techniques discussed here are only a part of the available techniques. The spatial domain and frequency domain methods both have some drawbacks. In the light of the tests we have completed, we can identify which techniques works and which do not. If pure rotational motions are present, on the scenes, we are investigating the use of Marcel and Luchesse seems to be problematic while the other three methods (RANSAC, Keren & Vandewalle) can be used. If a translational motion dominant scene is at hand all of the five methods shall work quiet well. However if the aim of the study is the estimating the sub-pixel motions as well as the other translational shifts, the use of RANSAC, Vandewalle and Keren is recommended.

The combination of rotational and translation motion is possible during image acquisition. The transrotational motion tests give us clue for such situations. Based on the findings of the transrotational transformation tests, Luchesse and Marcel methods cannot cope with this type of motion. However, RANSAC, Keren and Vandewalle methods survive for both test images in this experiment.

In addition to motion, the acquisition device suffers from the ambient noise in the environment. A noise robust method successful with the transrotational transformations can be used reliably in the super resolution studies. Under noisy conditions, Marcel and Luchesse methods are not reliable; the other methods can cope with noise levels implemented in the experiments.

Finally, we have examined the scaling transformation. Scaling occurs when the motion of the camera is away from the imaging plane. Only the RANSAC method has the capacity to identify the scaling parameter therefore the test is only limited to one method. The scaling results show that until some point of scaling RANSAC can identify the scale factor as well as the transrotational parameters successfully.

Because of the fact that after some level of zooming, the interest points get out of the viewpoint; the estimates become completely wrong.

Throughout the tests, the only algorithm that succeeded is the Keren. Nonetheless, The Vandewalle and RANSAC are both useful in most of the cases. Even more, since the RANSAC has the ability of detecting scaling, it may be preferable to use it.

One of the important points is the execution time of the algorithms. The Table 2.3 indicates the execution time of algorithms. This test is completed for one image only with the same transformation parameters.

Table 2.3 Execution times of the methods

	RANSAC	Keren	Vandewalle	Luchesse	Marcel
Elapsed Time (lena)	0.666 s	3.057 s	88.351 s	3.995 s	5.011 s

The results of the execution time test show that the RANSAC or Keren algorithm gives us the results fast. Finally, the use of Keren or RANSAC in our further discussions and tests will be appropriate.

2.7. IMAGE FUSION

Multi frame Image reconstruction image fusion is the integration of the sorted, aligned images into one common high-resolution grid. The information bits inside the images are fused to form a complete picture of the whole data set. Since every single pixel on the newly formed high-resolution image is a combination of the corresponding low-resolution pixels, the misalignment of these LR images will result in false convergences in data fusion step, which are obviously very disturbing (like ghosting). As a result, only the results of the best image registration method available should be used in the image reconstruction step. If the results of the current image registration algorithms are not precise enough, the known parameters

41

of the synthetic images will be used. In this chapter, the operational stages of the super-resolution techniques are examined. The description of image fusion methods, their comparison and experimentation results are reported in the next chapter.

2.8. QUALITY METRICS

There are two classes of objective quality or distortion assessment approaches. The first are mathematically defined measures such as the widely used mean square error (MSE) and peak signal to noise ratio (PSNR) which is a derivation of MSE. The formulations for these are:

$$MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}[\text{Im}(x, y) - \text{Im}'(x, y)]^2}{M.N}$$

[2.26]

$$PSNR = 20 * \log\left[\frac{255}{\sqrt{MSE}}\right]$$

where M, N stands for the size of the image in both horizontal and vertical axes, Im is the original HR image and Im' is the reconstructed HR image that is to be examined. MSE stands for error between two images, PSNR stands for error variance against the maximum possible image variance.

The second class of measurement methods considers human visual system (HVS) characteristics in an attempt to incorporate perceptual quality measures. Unfortunately, these complex metrics do not show any clear advantage over algebraic metrics such as MSE and PSNR under strict testing conditions and different image distortion environments.

The main function of the human visual system is to extract structural information from the viewing field, and the human visual system is highly adapted for this purpose. Therefore, a measurement of structural information loss can provide a good approximation to perceived image distortion [18]. Wang et al. regard the structural information in an image as those attributes that reflect the structure of objects in the scene, independent of the average luminance and contrast [17]. Structural Similarity (SSIM) index is an improved version of the method that Wang et al. [13] proposed before as a mathematically defined universal image quality index. The quality measurement approach does not depend on the images being tested, the viewing conditions or the individual observers. To find the quality index (Eqn. 2.21), first, the original ($x = \{x_i \mid i = 1,2,...,N\}$) and the test ($y = \{y_i \mid i = 1,2,...,N\}$) images are subjected to a 8×8 sliding window and for each position of the window, the formula below is calculated, where bars over letters designate average and σ stands for the variance of the pixel values within the window.

The sliding window calculations results in a quality map of the image where the dynamic range of the map is [-1, 1]. The best value 1 is achieved if and only if $y_i = x_i$ for all i. The overall quality index value is the average of the quality map. The quality index can be stated as:

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}.\bar{y}}{\bar{x}^2 + \bar{y}^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \qquad [2.27]$$

The first component is the correlation coefficient between x and y, which measures the degree of linear correlation between x and y, and its dynamic range is [-1, 1]. Even is x and y are linearly related, there still be relative distortions between them, which is evaluated in the second and third components. The second component, with a value range of [0, 1], measures how close the mean luminance is between x

43

and y. It equals one if and only if $\bar{x} = \bar{y}$. σ_x and σ_y can be viewed as estimate of the contrast of x and y, so the third component measures how similar the contrasts of the images are. Its range lies between 0 and 1, where the best value 1 is achieved if and only if $\sigma_x = \sigma_y$ [18].

CHAPTER 3

SUPER RESOLUTION METHODS

3.1 INTRODUCTION

The resolution improvement is the process of magnifying the image into a larger size. In the process of resolution enhancement, we have a number of pixels at hand. We create an empty grid of the targeted high-resolution image, depending on these pixels. According to the available low-resolution pixel intensities, we fill the entire image pixel by pixel. At the end, we have a higher resolution image, which has pixel values based on the available pixels. The filling process is the key point here. Depending on how we fill the empty grid, the resolution enhancement is classified.

For the single frame resolution enhancement, we have a single image and the available pixels on this image are placed on the high-resolution grid. While transferring the LR pixels to the HR grid, empty pixels between the existing pixels are left. After then, these empty pixels are filled up by a function of choice. However, suppose that we have a number of images to utilize, to create a high-resolution image. The images have different functions to show the same scene. As the functions are different, the intensities of corresponding pixels include different information for each point on the scene. Therefore, we fill whole high-resolution grid using a particular method of combining the distinct information of the available images. The general name of that particular method is the multi-frame resolution enhancement.

The main difference of multi frame and single frame methods is the number of images that they use. Throughout this chapter, we will discuss the two classes of the resolution enhancement techniques, systematically. First, the single frame methods are reviewed and then the multi-frame based techniques are discussed. As we understand the methods, the tests will take place, the discussions on the tests and comparison of the methods will be the subject of the next chapter. During the tests, we use synthetically transformed images to generate a high-resolution image (Figure 3.1).

Figure 3.1 The Transformed lena images with scaling (a) downsampled, no transformation (b) 3.32˚, (2.13, 3.92) pix, 1.02 x(c) 2.67˚, (4.67, 6.71) pix, 1.10 x (d) 6.76˚, (9.87, 1.39) pix, 1.08 x (e) 4.88˚, (6.59, 8.83) pix, 1 x

3.2. SINGLE FRAME RESOLUTION ENHANCEMENT

Single frame resolution enhancement is equivalent to image interpolation. Interpolation of an image is enlarging its size by estimating the unknown pixels of larger image from their neighboring pixels. We examine some interpolation techniques next. These results of these techniques will be used at benchmarks for super resolution methods.
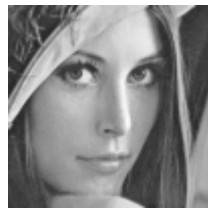
3.2.1. NEAREST NEIGHBOUR INTERPOLATION

Nearest Neighbor Interpolation is the most straightforward solution to the interpolation problem. As it is obvious from the name of the method, the subject of this technique is finding the nearest pixel value to the missing image value at a location then, assigning that nearest pixel values to the missing image value as in Eqn.3.1. In this method, the areas of pixels are enlarged by the targeted level of interpolation so the image quality is negatively affected. Mostly the image is in a shape of mosaics that disturbs the visual quality (Figure 3.2).
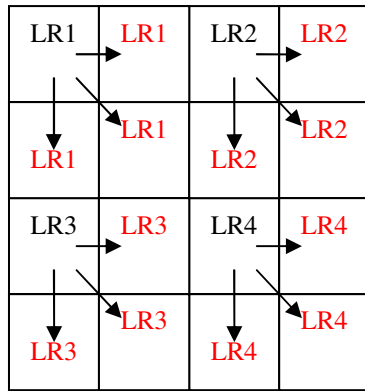
$$HR(2x-1,2y-1) = LR(x, y)$$
$$HR(2x,2y-1) = LR(x, y)$$
$$HR(2x-1,2y) = LR(x, y)$$
$$HR(2x,2y) = LR(x, y)$$

$$\forall x, y \in LR \ [\mathbf{3.1}]$$

(A)

(B)

(C)

Figure 3.2 The nearest neighbour interpolation scheme, (A) fill empty cells in the grid, (B) the original size lena image, (C) nearest neighbour interpolated lena image
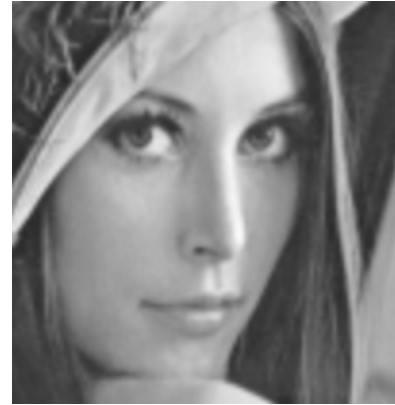
3.2.2. BILINEAR INTERPOLATION

Bilinear Interpolation is a simple but efficient way of enlarging images. The main concern of this method is to fit a bilinear surface through existing data points. The resultant image will be a smoother than the nearest neighbor interpolation. In this method, the available pixels are placed into a HR grid, leaving zeros between them. In both vertical and horizontal directions, empty pixels are filled with the linear function values between the existing pixels. This process is done for the rest of the empty cells including the newly found pixel values into the process. By using this method, every empty pixel is filled with a value affected by the nearest four existing pixels depending on the distance to them (Figure 3.3).

LR1	LR1/2 + LR2/2	LR2	LR2/2 + LR3/2	LR3
LR1/2 + LR4/2	(LR1+ LR2+LR 4+LR5)/4	LR2/2 + LR5/2	(LR2+ LR3+LR 5+LR6)/4	LR3/2 + LR6/2
LR4	LR4/2 + LR5/2	LR5	LR5/2 + LR6/2	LR6
LR4/2 + LR7/2	(LR4+ LR5+LR 7+LR8)/4	LR5/2 + LR8/2	(LR5+ LR6+LR 8+LR9)/4	LR6/2 + LR9/2
LR7	LR7/2 + LR8/2	LR8	LR8/2 + LR9/2	LR9

(A)

$$HR(2x-1,2y-1)=LR(x,y)$$
$$HR(2x-1,2y)=[LR(x,y)+LR(x,y+1)]/2$$
$$HR(2x,2y-1)=[LR(x,y)+LR(x+1,y)]/2$$
$$HR(2x,2y)=[LR(x,y)+LR(x+1,y)+LR(x,y+1)+LR(x+1,y+1)]/2$$

$$\forall x, y \in LR \qquad\qquad [\mathbf{3.2}]$$

(B)

(C)

Figure 3.3 The bilinear interpolation scheme, (A) fill empty cells in the grid, (B) the original size lena image, (C) bilinear interpolated lena

3.2.3. BICUBIC INTERPOLATION

Bicubic Interpolation is an advanced version of the bilinear interpolation. Bicubic interpolation uses a 4 by 4 neighborhood to find the missing pixels in the high-resolution grid. Bicubic interpolation uses a polynomial passing through four pixels to make a decision. Therefore, bicubic interpolation creates enlarged images that are smoother and higher quality (Figure 3.4). The following equation [3.3] is the continuous time convolution kernel of the cubic interpolation. When the interpolation is separately applied to rows and columns of an image, we have the bicubic interpolation [15].

$$
h(t) = \begin{cases} \dfrac{3}{2}\,|t|^3 - \dfrac{5}{2}\,|t|^2 + 1 & 0 \leq |t| \leq 1 \\[2mm] -\dfrac{1}{2}\,|t|^3 + \dfrac{5}{2}\,|t|^2 - 4\,|s| + 2 & 1 \leq |t| \leq 2 \\[2mm] 0 & |t| > 2 \end{cases}
$$

[3.3],

(A)

(B)

Figure 3.4 The bicubic interpolation scheme, (A) the original size lena image, (B) bicubic interpolated lena image

Throughout the thesis, the bicubic interpolation will be used as a reference to the quality, because it has a proven degree of quality and ease of use.

3.3. MULTI-FRAME RESOLUTION ENHANCEMENT

As we have mentioned in the previous chapter, the super-resolution methodology consists of two main phases. One of these phases is the image registration phase, which was discussed in a detailed manner. The other phase is the image fusion. After the LR images are registered, they should be combined in a suitable way. The image fusion allows us to combine these details into one high-resolution image of the scene. The name multi-frame resolution enhancement is used for the image fusion step since the resultant image is a higher resolution image than the input images that are low resolution. The two phrases will be used interchangeably throughout this thesis.

For the registration, we have selected the Keren and the RANSAC algorithm. As noted before, the noise components affect the RANSAC algorithm negatively, we will use Keren based registration for noisy images. However, the noiseless images are registered with RANSAC better than the others are.

3.3.1. DIRECT ADDITION

After the input images are registered, the most basic method to fuse these images is to get mean or median of the images. These two methods are extremely basic and easy to implement. Inevitably, they have some drawbacks such as blurring, and degradation of details that are not present in every image. As an advantage, these methods can reduce the effects of misregistrations and noise successfully, because of the low pass filtering nature of the mean and median operations. As compared to the single-frame methods, these methods are quite discouraging in consideration of being multi-frame methods. However, the noise suppression capability and the opportunity of adding the image restoration methods to the algorithm, make these

methods considerable. In addition to these, the two variants of direct addition methods are single-run methods and have a very low computational complexity.

The direct addition methods have a few steps, which do not alter as the image filtering scheme changes. First, we have the registered images at hand. These images are upsampled by using bilinear or bicubic interpolation as the first step to the resolution enhancement. As we have the upsampled and aligned pictures of the scene, we can add them to form the final image. The type of addition determines the filtering type as median or mean. The mean filtering of upsampled-registered images takes the mean of all overlapping pixel values, whereas the median filtering takes the median value of each overlapping pixel set (Figure3.5).

$$HR(x, y) = \sum_{i=1}^{n} \frac{(\uparrow LR)(x, y, i)}{n} \qquad \forall x, y \in HR \qquad [3.5]$$

$$HR(x, y) = median((\uparrow LR)(x, y, n)) \qquad \forall x, y \in HR \qquad [3.6]$$

Where HR is the targeted high-resolution image, LR is the available low-resolution images, $\uparrow$ is the upsampling operator, x, y is the pixels of the high-resolution grid and n is the number of available LR images.
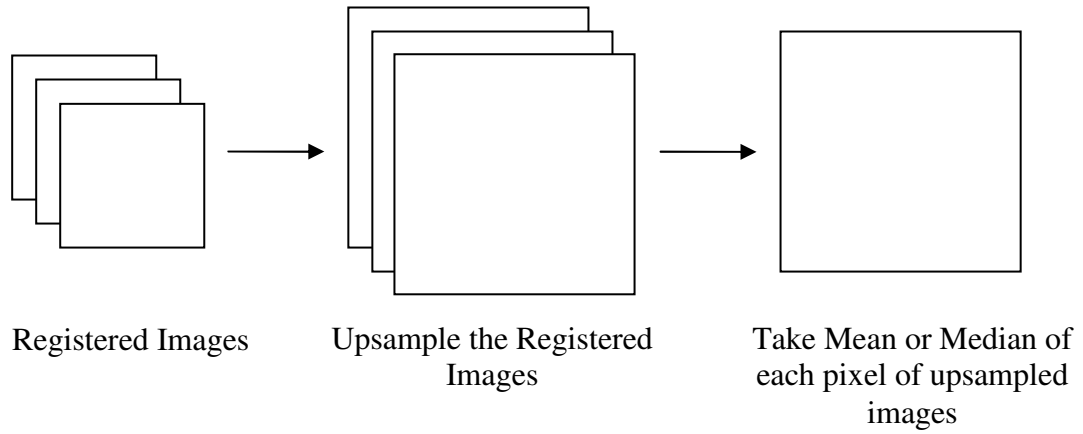
Figure 3.5 Pipeline of Direct Addition with Median Filtering Algorithm

The tests takes place for the direct addition with median filtering only, because the mean and median operation give similar results and most of the time median filtering operates more accurately. The fusion tests show that major draw back of the addition method is the blurring effect. This blurring reduces the SSIM index and PSNR values negatively. The scaling is a corruption that is hard to recover. Even the RANSAC cannot register perfectly. Since the RANSAC algorithm registers the scaled images with some error, Figure 3.6 (A) (C) has some artifacts along the edges. Although the blurring is a drawback; as noise increases on the images, the cause of this effect suppresses noise. Even if the information is unrecoverably corrupted in one image due to noise, the application of direct addition to the noisy LR images, recovers some information and the quality increases slightly than any of the single frame methods (Figure 3.6(B) (D)). Despite adding whole image set as in the previous section, seems to be appropriate; the effect of subpixel information is reduced or ignored while mean or median filtering.
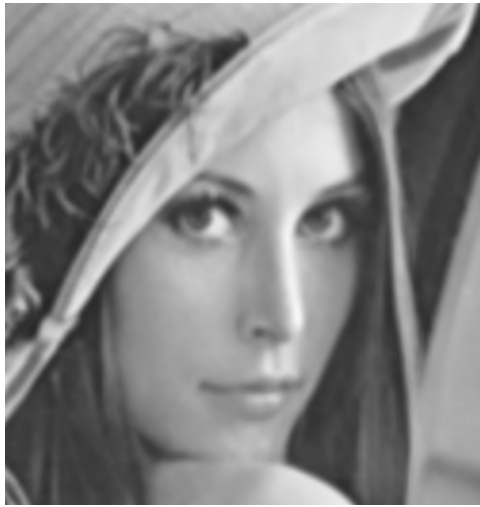
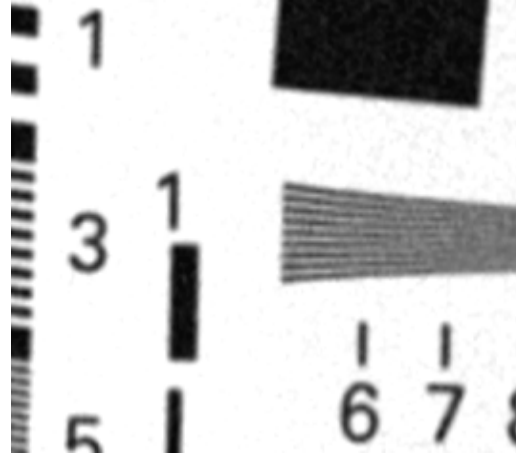|	
SSIM	0.851
PSNR	28.805
MSE	9.253

(A)

SSIM	0.803
PSNR	28.123
MSE	10.009

(B)
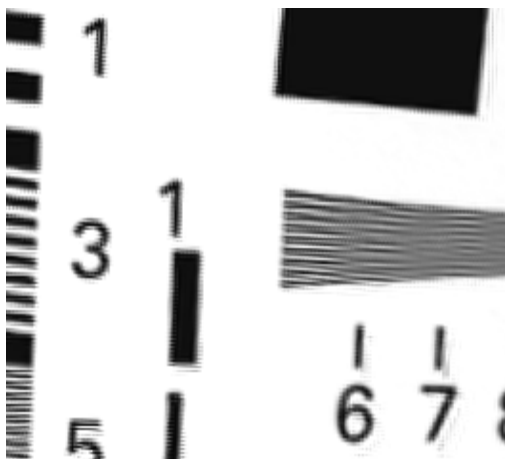
SSIM	0.781
PSNR	17.056
MSE	35.789

(C)

SSIM	0.721
PSNR	21.434
MSE	21.62

(D)

Figure 3.6 Direct Addition with Median Filtering on (A)(C) RANSAC registered images (rotation+translation+scaling applied); (B)(D) Keren registered images (rotation+translation+noise applied)

The direct addition methods suffer from their negligence of motion information. The motion information is discarded after image registration and disregarding the information fusion leads to noise removal or enhancement of signal quality but not to resolution improvement. Other methods such as non-uniform interpolation make use of information on sub-pixel motion to construct super resolution images.

3.3.2. NON-UNIFORM INTERPOLATION

The super resolution methods are based on the fact that, even if the images are of the same scene, the sub pixel shifts between them makes them distinct. All of the super-resolution algorithms, specifically the image fusion methods, intend to recover this subpixel information. In the direct addition method in the previous section, the effect of subpixel information is reduced or ignored while mean or median filtering. Therefore, the blurring occurs and the information hidden between the pixels cannot be recovered as desired. In the method of nonuniform interpolation, we consider every image, pixel by pixel to overcome the high frequency information suppression,

The non-uniform interpolation method aims to fuse all these information as affectively as possible. Since the images are shifted in integer pixels and fractional pixels together and when we register them in integer pixel level, the subpixel shifts left unregistered. As the pre-registered images put together, the pixel values of all images are scattered surrounding the reference frame pixels (Figure 3.7). If the registration is perfect, the scattered pixels show the exact values of the corresponding points in the scene. However, we have some registration and calculation errors present. Since the high-resolution grid does not hold the pixels of each scattered pixel at exact place, the best we can do is to calculate the effects of the scattered pixels on high-resolution grid's cells and assigning these values as the high-resolution pixels. According to this method, the high-resolution pixels are composed of the combination of the pixels surrounding them [3.7].

After the integer level translations and rotations, scaling are compensated:

$$HR(x,y) = \frac{1}{D}\sum_{i=1}^{n}\frac{LR(a,b,i)}{d} \qquad \text{iff} \begin{cases} x-\delta_x < a < x+\delta_x \\ y-\delta_y < b < y+\delta_y \end{cases} \qquad [\mathbf{3.7}]$$

$$d = \sqrt{|x-a|^2 + |y-b|^2}$$

Where D is the total distances for that individual HR pixel, x, y∈ HR and a, b∈ LR and a, b includes fractional sub-pixel information.
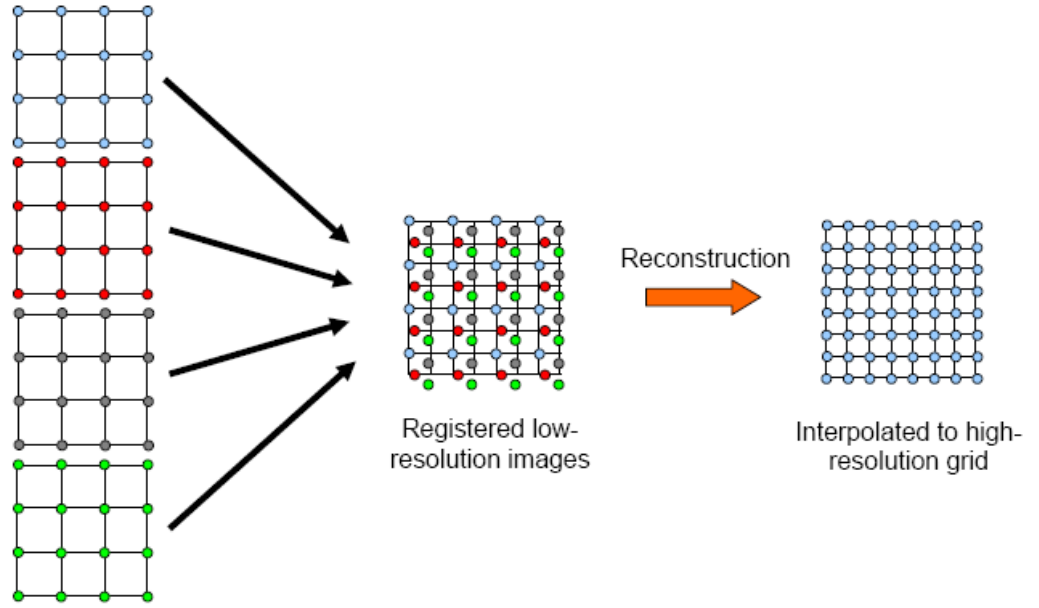
Figure 3.7 Four LR images are preregistered and aligned without compensating subpixel shifts. Their individual pixel values create the HR image on the HR grid.[16]

The procedure is very easy to understand but the implementation is a bit problematic since the pixels are computed one by one and the surrounding pixel values and distances of the LR pixels to the HR pixel are calculated for each pixel. (Figure 3.8)
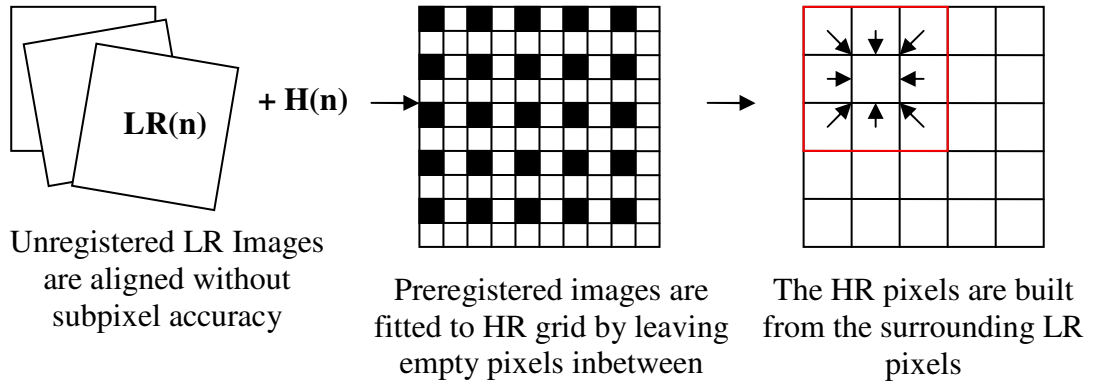
Figure 3.8 Pipeline of Nonuniform Interpolation Algorithm

During the tests, the nonuniform interpolation performs quite well (Figure 3.9). With some attention to the code efficiency, real-time applications are possible. However, the most important disadvantage of non-uniform interpolation is its sensitivity to noise and misregistrations. The algorithm ignores these errors and all the errors become artifacts at the end. If the blurring and noise characteristics of the LR images differ, the algorithm also fails to end up with a successful result. The algorithm almost overcome the noise present on the image and gives one of the best results among other methods. Nevertheless, the results of the scaled image fusion test are not very good. This is because the registration errors are present in that case. Non-uniform interpolation is a promising method, but the need of using image restoration methods for both the LR images and resultant HR image is undeniable. As the last word, we can say that the non-uniform interpolation is as close as we can get to the ideal of reconstruction of a continuous image form its non-uniformly sampled pixels gathered from multiple LR observations.

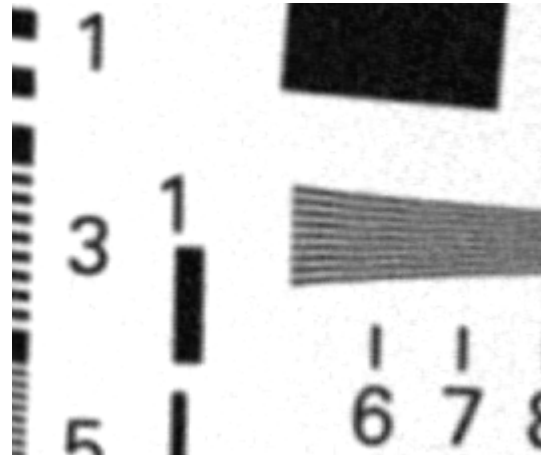| SSIM	0.877
PSNR	25.660
MSE	13.290

(A)

SSIM	0.847
PSNR	27.347
MSE	10.944

(B)
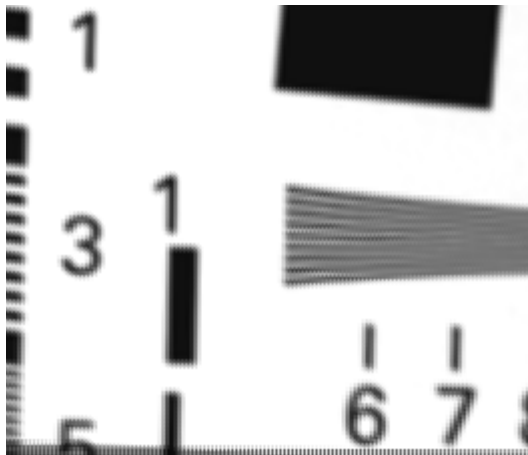
SSIM	0.832
PSNR	18.799
MSE	29.283

(C)

SSIM	0.804
PSNR	21.437
MSE	21.612

(D)

Figure 3.9 Nonuniform Interpolation on (A)(C) RANSAC registered images (rotation+translation+scaling applied); (B)(D) Keren registered images (rotation+translation+noise applied)

3.3.3. ITERATIVE BACKPROJECTION

The previous tests show that the performances of the single-run methods are not satisfactory. The expected performance should at least as the same as the single frame methods. One of the ideas to increase the performance is using an iterative methodology. Iterative methods use prior information of the previous results to get a better outcome. Many iterative multiframe superresolution restoration methods have in common a simple but powerful simulate-and-correct approach to restoration. The iterative backprojection method is one of them, which is very easy to understand in application. The IBP method provides a useful framework for solving the super resolution problem by providing a mechanism for constraining the super resolution restoration to conform to the observed data [12]. The algorithm needs a priori image to start the process. As the priori image, this method uses the mean of the registered images. By this way before beginning the iterations, we have a base image to build up our outcome. The basic motive in our algorithm is to generate a simulation image, generate LR images from the simulation image to compare them with the observed counterparts and using the error between them to generate a better quality simulation (Figure 3.10). As long as the ending rule is not satisfied, the iterations will continue. The ending rule may be a threshold value, the point where the quality stop increasing, or simply a predefined number of iterations.

The aim of the IBP method is to minimize the error between simulated LR images and the observed LR images iteratively. All corrections on the simulated HR image generated by the various LR pixels are combined by taking their weighted average, using the back projection kernel [14].

There is the ease of understanding of the scheme. However, the solution is not unique due to the ill-posed characteristic of the problem and choice of the priori image is vital, where the solution may not converge or the solution converges too slowly. In fact, many solutions that satisfy the constraints given by the observed low-resolution frames exist [12].

Therefore, the choice of the priori constraint should be carefully selected. As taking the mean of all registered images and then upsampling the result is a basic and easy to implement, in this study we use the mean operation as the priori constraint finding way [3.6]. Certainly, different methods of finding priori constraints may be used.

$$x^{n+1}[n_1,n_2] = x^n[n_1,n_2] + \sum_{k=1}^{N}(y_k[m_1,m_2] - y_k^n[m_1,m_2]) \times h^{BP}$$
[3.6]

Where n is the iteration number, x is the simulated HR image, N is the number of images, y_k is the observed LR images and y_k^n is the final simulation of LR images after n iterations and n1,n2 are the HR space and m1,m2 is the LR space. Finally, h^{BP} is the backprojection kernel where the LR image is mapped into HR grid, which is the mean and upsampled operation, in our study. However, h^{BP} may be utilized as an additional constraint, which represents the desired property of the solution [16].

Obviously, the RANSAC registered scaled images will contain some registration errors. The effects are clear from the previous tests. The IBP method successfully overcomes this kind of corruption (Figure 3.11 (A) (C)). In addition to this, the noise suppression capability of the IBP method is not very competent as the quality metrics are considered. This is because the priori image suffers from the noise and the noise present on all of the observed LR images exaggerates the error.
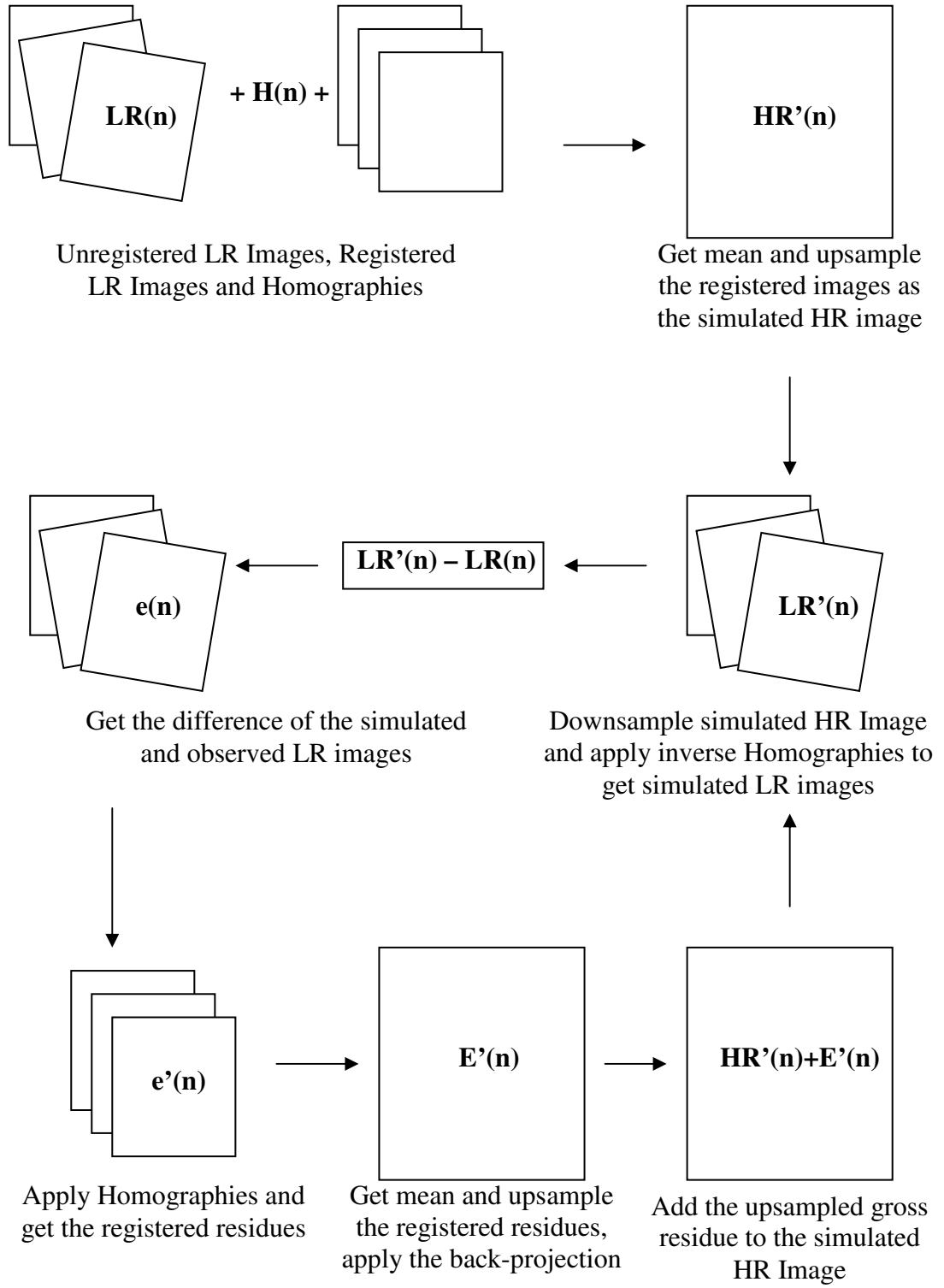
Figure 3.10 Pipeline of Iterative Backprojection Algorithm

SSIM	0.9
PSNR	30.531
MSE	7.586

(A)

SSIM	0.868
PSNR	28.804
MSE	9.254

(B)
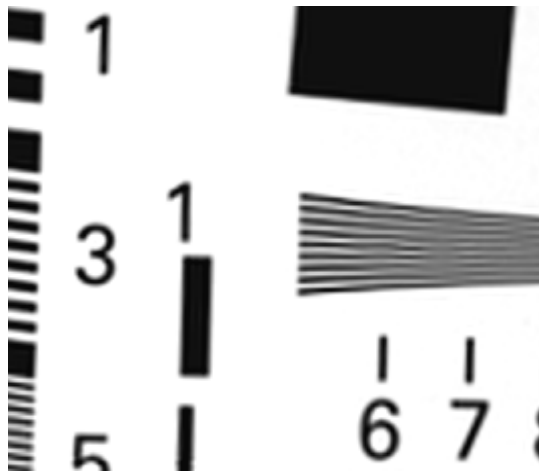
SSIM	0.925
PSNR	24.835
MSE	14.614

(C)
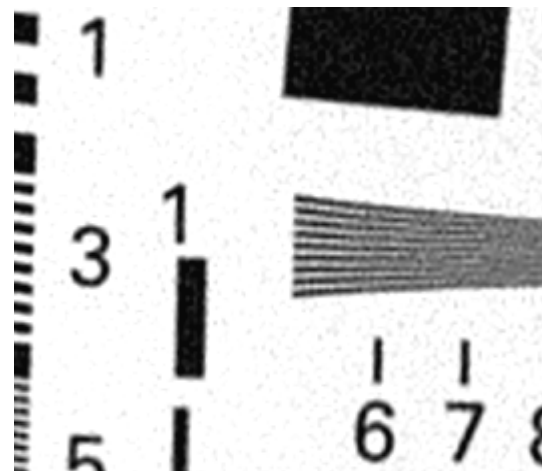
SSIM	0.669
PSNR	23.119
MSE	17.807

(D)

Figure 3.11 Iterative Backprojection (10 iterations) on, (A)(C) RANSAC registered images (rotation+translation+scaling applied); (B)(D) Keren registered images (rotation+translation+noise applied)

3.3.4. IBP WITH NONUNIFORM INTERPOLATION

The IBP with Nonuniform Interpolation method is a variant of the IBP. It also contains the non-uniform methodology. The main difference is the backprojection kernel used. Instead of using the upsampled mean of the registered input images as initial constraint in IBP, nonuniform interpolation is used in IBP with nonuniform interpolation. The main formulation is the same as the IBP [3.6] but the IBP changes. The method has both the nonuniform interpolation method's and IBP method's disadvantages. The noise suppression capability is not very promising also the registration errors affects the results negatively. Therefore, the use of additional methods is necessary. One of the powerful ideas in IBP is the error suppression by mean filtering. Since in IBP with non-uniform interpolation method does not have such steps, the addition of low-pass filters will be effective. The addition of filters in fact does not violate he rule of not using image restoration throughout the thesis, because we do not change the final image but the backprojection kernel IBP.

Throughout the IBP with nonuniform interpolation tests, the results show that without applying a proper back-projection kernel this method does not perform well (Figure 3.13). However, as soon as the kernel is fixed, the effect of misregistrations and errors are decimated on the initial estimate image and the simulated LR images (Table 3.1).

Table 3.1 The effect of back-projection kernel choice in IBP algorithms

Reschart.jpg	SSIM(scaling)	PSNR(scaling)	SSIM(noise	PSNR(noise)
Without filters	0.804	20.163	0.399	21.147
With filters	0.859	21.991	0.569	22.635
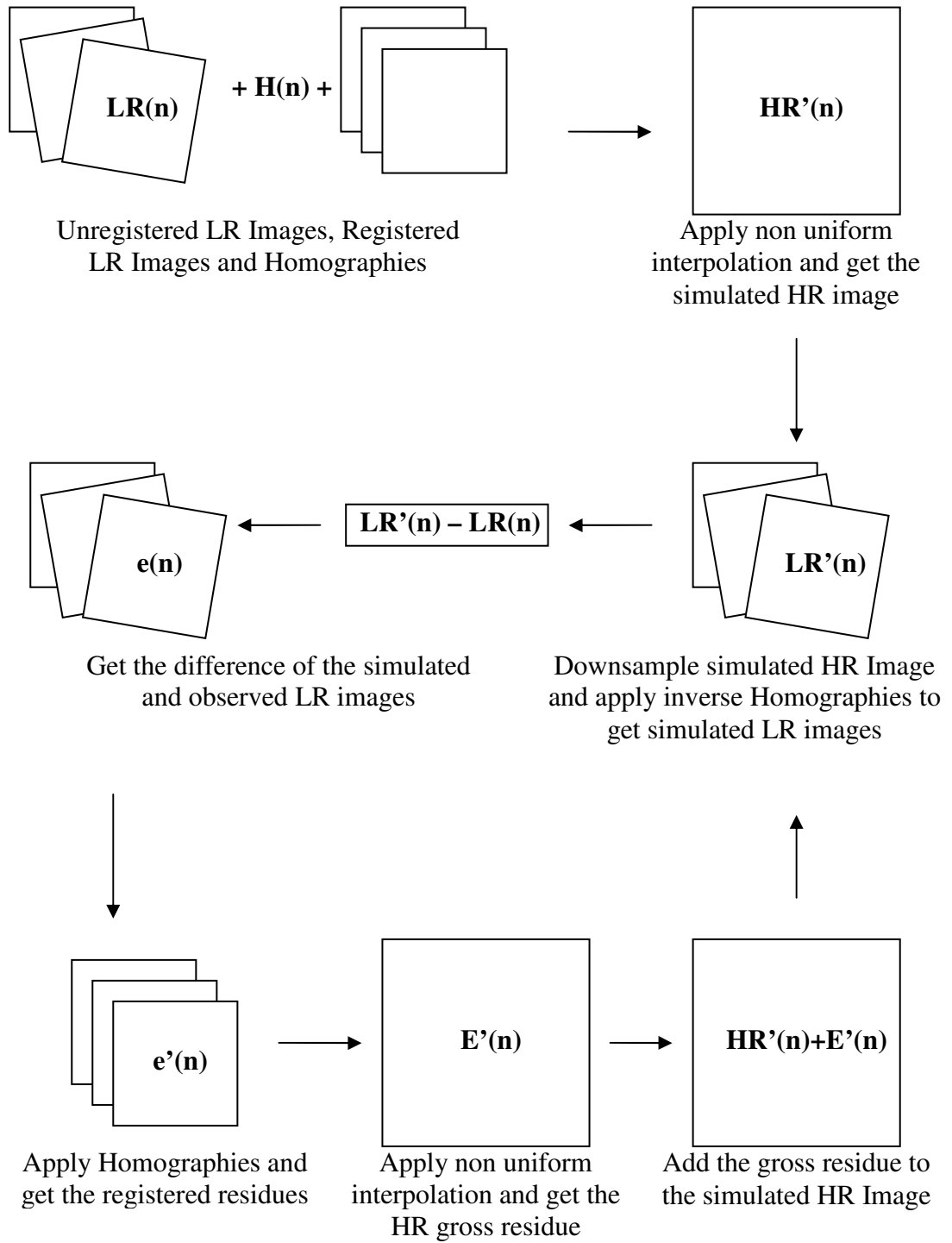
Figure 3.12 Pipeline of Iterative Backprojection with Nonuniform Interpolation Algorithm
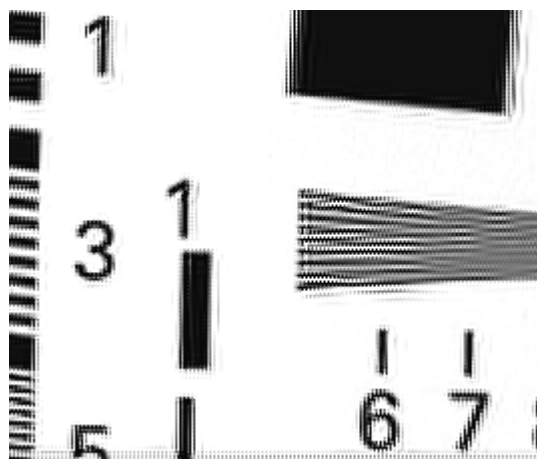
SSIM	0.811
PSNR	24.780
MSE	15.682

(A)

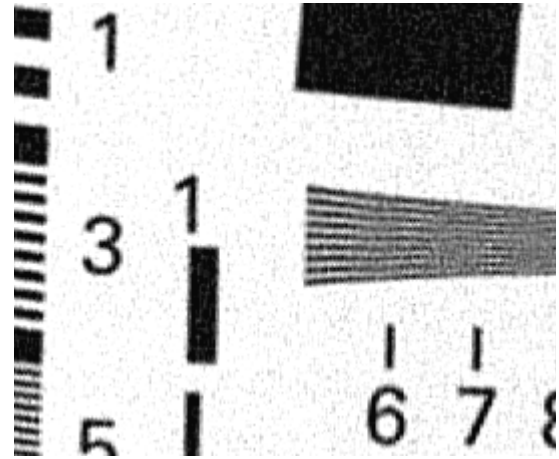|	
SSIM	0.599
PSNR	25.658
MSE	13.294

(B)

SSIM	0.804
PSNR	20.163
MSE	25.024

(C)

SSIM	0.399
PSNR	21.147
MSE	22.347

(D)

Figure 3.13 Iterative Backprojection with Nonuniform Interpolation on (A)(C) RANSAC registered images (rotation+translation+scaling applied); (B)(D) Keren registered images (rotation+translation+noise applied)

The pipeline of the IBP with nonuniform interpolation algorithm is shown in figure 3.13; the updated IBP methodology will be injected to the algorithm in the first estimate and for every difference operation of the simulated and observed LR images.

3.3.5. PROJECTION ONTO CONVEX SETS (POCS)

The iterative solutions to the image fusion problem aim to minimize the error as the process passes over the same procedures. The IBP algorithms above scan the whole image at once and correct the error to converge to any solution that satisfies the error minimalization, within the whole solution space. The set theoretic methods, by the most known variant Projection onto Convex Sets (POCS) method in this study, solve the restoration problem by defining sets of constraints which must be satisfied by candidate solutions [16]. The definition convex sets is very wide and special to the image processing, specifically our problem; image fusion, these sets are the observed data. POCS aims to converge to a solution, which does not violate the borders of the observed LR images (Figure 3.14). POCS method finds alternative solutions for every pixel value, which satisfies both convex sets, namely the observed LR images, simultaneously. Stark and Oskoui [10] did an early work on the subject. They use closeness and convexity of the constraint sets to ensure convergence of iteratively projecting the images onto the sets. Tekalp, Ozkan, and Sezan [11] propose a more robust POCS formulation.

The POCS method is a simple but powerful operation (Figure 3.15). The process begins with the determination of the priori constraint, which is the reference frame that is within the solution space. The reference frame interpolated to the HR copy and the first estimate is at hand. The iterative process starts here and passes over every frame and every pixel to locate the solution.
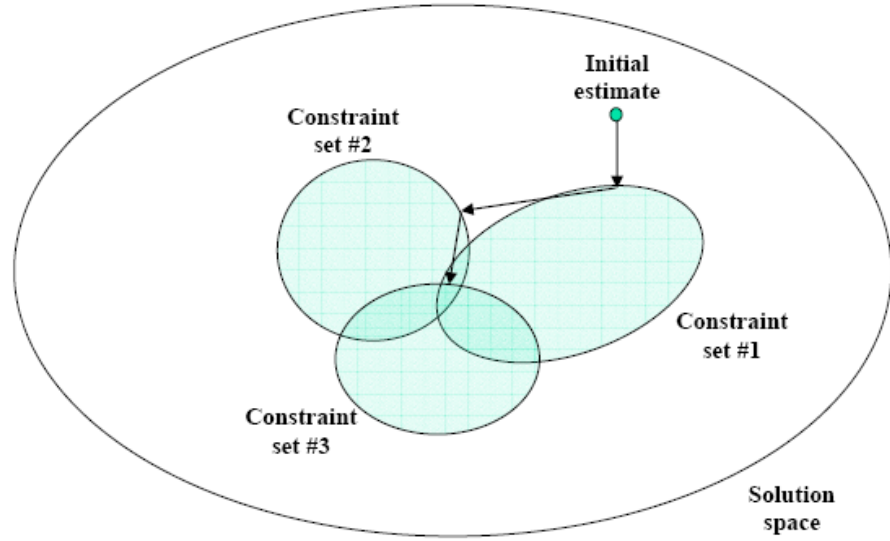
Figure 3.14 In the POCS technique the initial estimate is projected to the convex sets iteratively.

First, the images are registered, and we get the motion compensated coordinates of every pixel. These pixel values must be projected to the HR space and this is done by applying a Gaussian PSF for every pixel. For every image in our set, we have a solution then. Unifying these solutions into the intersection of these sets is the next step. Every pixel value has the prior information and the found value specific to that set. The next thing is to stretch the available estimate to satisfy the solution set we are working on. The pixel value is updated at the level of the threshold to match the projected value as close as possible without disturbing the continuity of the solution. After the prior estimate is stretched to every frame of the input set, first iteration is completed. The solutions are normalized to the intensity space [0,255] and the next simulated HR image is ready.

$$x^{n+1} = P_m P_{m-1} \cdots P_2 P_1 x^n \qquad\qquad [\textbf{3.8}]$$

where x^0 is an arbitrary starting point, and P_i is the projection operator which projects an arbitrary signal x onto the closed, convex sets, C_i. For each pixel within the LR images $y_k[m_1, m2]$

$$r^{(x)}[m_1,m_2] = y_k[m_1,m_2] - \sum_{n_1,n_2} x[n_1,n_2]W_k[m_1,m_2;n_1,n_2]$$ **[3.9]**

$$C_D^k[m_1,m_2] = \left\{ x[n_1,n_2] : \left| r^{(x)}[m_1,m_2] \right| \le \delta_k[m_1,m_2] \right\}$$ **[3.10]**

The projection of an arbitrary $x[n_1,n_2]$ onto $C_D^k[m_1,m_2]$ can be defined as:

$$x^{(n+1)}[n_1,n_2] = x^{(n)}[n_1,n_2] + \begin{cases} \dfrac{(r^{(x)}[m_1,m_2] - \delta_k[m_1,m_2]).W_k[m_1,m_2,n_1,n_2]}{\sum_{p,q} W^2_k[m_1,m_2,p,q]}, & r^{(x)}[m_1,m_2] > \delta_k[m_1,m_2] \\[4mm] 0, & \left| r^{(x)}[m_1,m_2] \right| \le \delta_k[m_1,m_2] \\[4mm] \dfrac{(r^{(x)}[m_1,m_2] + \delta_k[m_1,m_2]).W_k[m_1,m_2,n_1,n_2]}{\sum_{p,q} W^2_k[m_1,m_2,p,q]}, & r^{(x)}[m_1,m_2] < -\delta_k[m_1,m_2] \end{cases}$$ **[3.11]**

Additional constraints after [3.11] can be utilized to improve the results [10].

The drawbacks of the POCS method are mainly the nonuniqueness of the solution since the solution may converge to any member in the intersection set. The computational cost is high due to the pixelwise recovery of the values, rather than the scanning of the whole frame. Additionally, the method converges very slowly or even may not converge at all.

LR(n) + **H(n)** → **HR'(n)** →

Unregistered LR Images, and Homography Matrices

Upsample the reference LR image as the simulated HR image

For all LR images, apply the Homographies to find the motion compansated pixel coordinates

For every pixel in the image, find the difference between the simulated HR image pixels and gaussian PSF applied pixels.

To the found motion compansated coordinates apply the gaussian PSF for each pixel

+δ	0	-δ	•••
+δ	-δ	+δ	
-δ	+δ	0	

→ **Normalize HR(x,y) between 0 and 255** →

If diff > δ , subtract δ weight from that pixel,
If diff < δ , add δ weight to that pixel,
If diff = δ , leave as it is.

After this procedure applied to all pixels we have updated the simulated HR image

Figure 3.15 Pipeline of Projection onto Convex Sets Algorithm

SSIM	0.922
PSNR	18.553
MSE	30.122

(A)

SSIM	0.663
PSNR	21.224
MSE	22.148

(B)

SSIM	0.911
PSNR	17.056
MSE	35.789

(C)
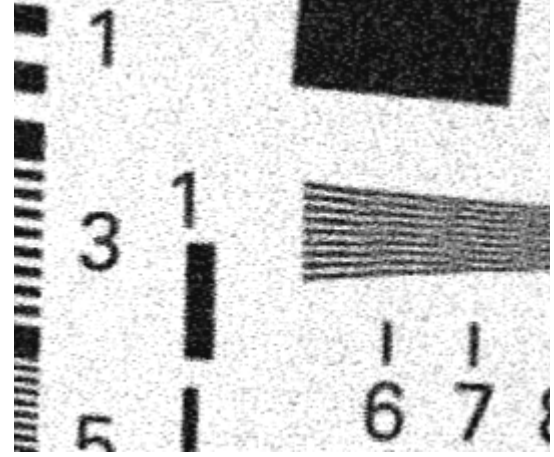
SSIM	0.679
PSNR	20.702
MSE	23.520

(D)

Figure 3.16 Projection onto Convex Sets on (A) (C) RANSAC registered images (rotation+translation+scaling applied); (B) (D) Keren registered images (rotation+translation+noise applied)

The tests on POCS algorithm show that the method is extremely sensitive to noise and registration errors. As appears in figure 3.16 (A) (C) the registration errors causes too many artifacts and these results in poor scores in both PSNR and SSIM index. Also as seen in figure 3.15 (B) (D) since the noise is not filtered a cumulative noise of all images appears in the final HR image, which also causes the degradation of quality. However, the good results in registration and with lower noise present on input images, the POCS is a promising method.

3.3.6. COMPARISON OF FUSION METHODS

Thus far, we have covered the definitions of image fusion algorithms, the flowcharts of them and their some sample applications. As we have done for the image registration algorithms in the previous chapter, the next step will be the comparison of these methods. Throughout the comparison, the methods will be tested for several images, several cases.

3.3.6.1. TEST METHODOLOGY

The tests are all done by using the Super Resolution GUI build specifically for this purpose. The use of image quality metrics that was discussed in the previous chapter is only possible by the presence of the original high-resolution image. By the absence of this reference frame, only visual assessment is possible. Therefore, we get an image as the reference image and then we tried to form some transformed synthetic low-resolution variants of this reference image. The results of the SR methods are compared to the reference frame and the success of the methods is analyzed based on a numerical background. Even for this case, visual assessment should be done, because of the fact that both SSIM and PSNR values depend on the image differences and in some cases, the indexes might differ from what we see visually.

During the tests, we use the image registration algorithms, Keren and RANSAC. The success of RANSAC algorithm in recovering every kind of spatial

70

transformation and Keren's durability in noisy conditions, make us to use them in various cases. The synthetic image tests are done for the two images as we use in image registration section. We use the bicubic interpolation as the control variable. The use of bicubic interpolation is because this kind of single-frame resolution enhancer gives very reliable and successful results and as we use multi-frame methods and use a lot of computational power, we should at least reach the performance of this method. Otherwise, we already have the working bicubic interpolation at hand, with less computational cost and more reliability.
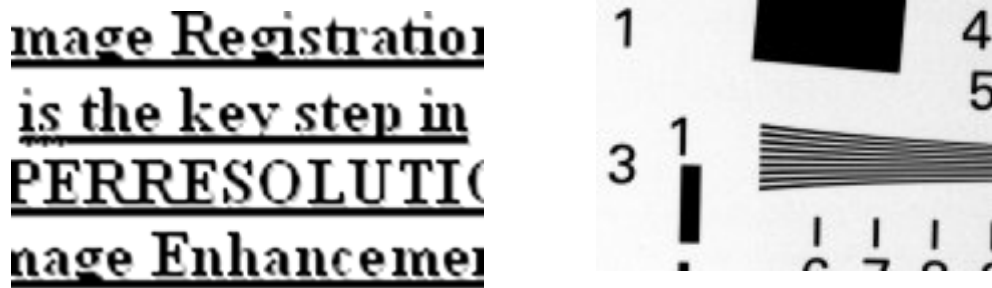
Figure 3.17 Images (Test1.bmp and Test2.bmp) that are used in the tests. with transformations : **(A)** 1.08 degrees (5.77, 7.24) pixels; **(B)** 0.51degrees (1.33, 1.89) pixels; **(C)** 1.22 degrees (5.59, 7.09) pixels; **(D)** 1.81 degrees (0.12,6.69) pixels; **(E)** 0.29 degrees (4.91, 2.41) pixels.

First, the images have no noise or scaling on them, but only the transrotational transformations. After registering the images with Keren algorithm, we fuse the images by using several algorithms one by one. Next, the scaling is applied to the images and RANSAC is used to recover this type of corruption. Once again, all methods are used to fuse RANSAC registered images. Finally, some noise is added and Keren is used to recover the transformations. Certainly, we repeat the fusion step once again. During these tests, we gather the SSIM index and PSNR values, as well as some visual patterns to compare the methods. As another step of performance evaluation, we registered the video frames by RANSAC algorithm and fuse the registered images by using whole set of methods. The visual patterns are

71

gathered for comparison, either. At the end of the tests, we collect some data to evaluate the effects of image number and iteration number to the success of the methods. During the tests, five LR synthetic images and seven video frames are used to evaluate the performance of the methods.

3.3.6.2. QUALITY METRIC BASED COMPARISON

The evaluation of some methods using a numerical constraint is relatively easy and preferred. The image quality is one of the though areas for assigning numerical values to the samples. The basic idea is to get the difference of the images and find the value as in PSNR, which may suffer from the fact that slight intensity differences or slight shifts may cause terrible outcomes, even if that is not the case. In addition more trustworthy methods based on HVS (human visual system) are also used, in which the assessment is structural and immune to slight variances. We use both for our comparison.

For the first test pattern, the the image of the writing have no noise or scaling and they are perfectly registered (Table 3.3). The results show that the POCS and IBP methods perform very well. POCS can even get more points than the bicubic interpolation. As taking the bicubic method as the reference to our test the POCS, IBP and Nonuniform Interpolation methods performs better, while IBP with nonuniform Interpolation and direct addition methods performs worse. Similar to this; the test2 image also gives similar results.

The registration errors increase slightly, when we register the images using the Keren algorithm instead of directly using known trans-rotational values. Misregistrations cause all methods to degrade in performance. Therefore, all of the methods looses SSIM and PSNR values similarly (~0.5 dB – 1 dB in PSNR and <0.04 pts in SSIM index), but nonuniform interpolation looses much more quality (~2dB in PSNR and >0.2pts in SSIM index). (Table 3.2)

Table 3.2 Comparison of methods based on the quality metrics

Method	Metric	Test1.bmp		Test2.bmp	
		Perfect Registered	**Keren Registered**	**Perfect Registered**	**Keren Registered**
Direct Addition with Median Filtering	SSIM	0.459	0.449	0.767	0.767
	PSNR	10.227	10.176	18.793	18.766
Nonuniform Interpolation	SSIM	0.685	0.458	0.822	0.748
	PSNR	12.007	10.522	20.708	18.547
Iterative Backprojection (IBP)	SSIM	0.725	0.695	0.798	0.788
	PSNR	12.392	12.118	20.733	20.616
IBP w Nonuniform. Interpolation	SSIM	0.483	0.433	0.716	0.675
	PSNR	10.374	10.063	19.293	19.064
Projection onto Convex Sets (POCS)	SSIM	**0.799**	**0.750**	**0.854**	**0.871**
	PSNR	**13.492**	**12.762**	**21.586**	**21.924**
Bicubic Interpolation	SSIM	0.671	0.671	0.819	0.819
	PSNR	11.993	11.994	20.504	20.508
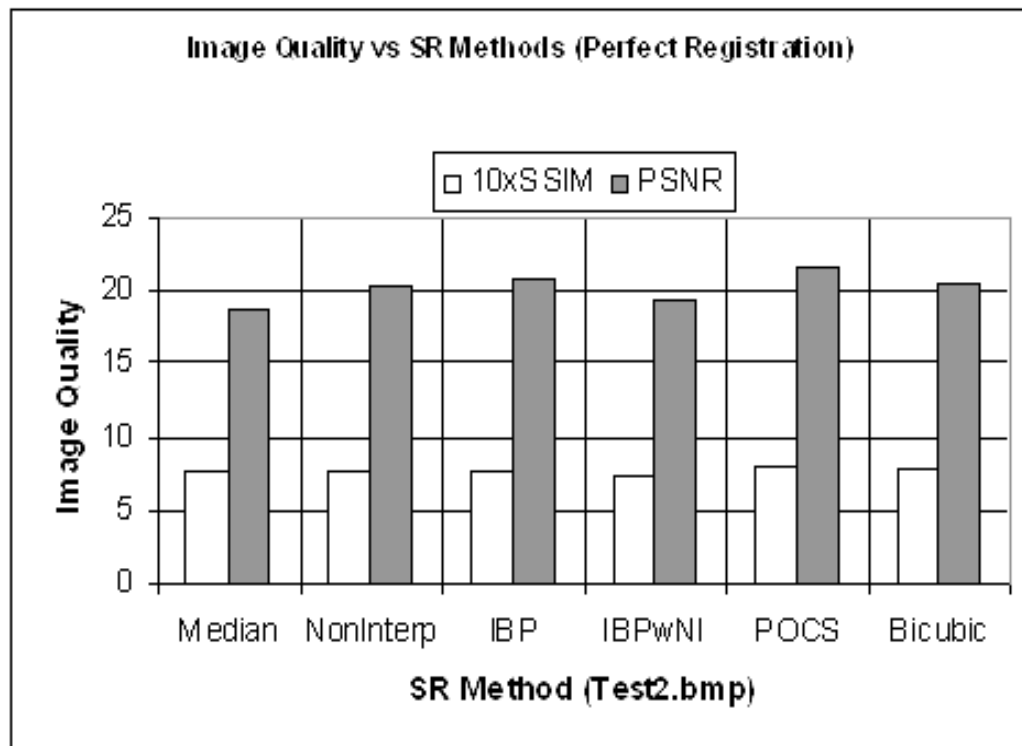
Figure 3.18 Test1 & Test2 images QMs

As a result, as long as the LR images preserves original outline, the single frame methods performs better since we downsample the HR image and then upsample it. As soon as the noise present in the LR images, the single frame methods need hard core image restoration, whereas the multi frame methods can cope with noise by their information fusion and noise suppressive capabilities. However if the registration is perfect and noise does not very effective, POCS performs very well as well as IBP and nonuniform interpolation (Table 3.2).
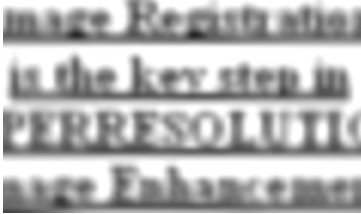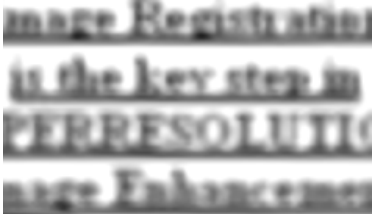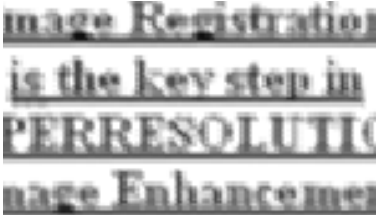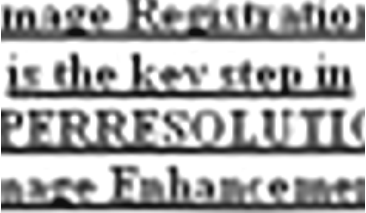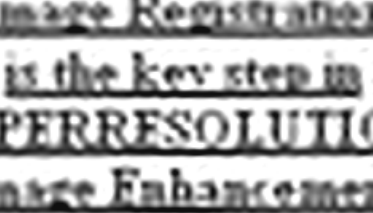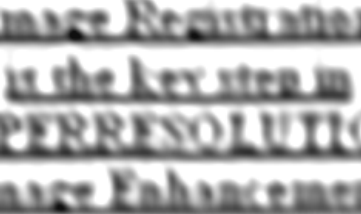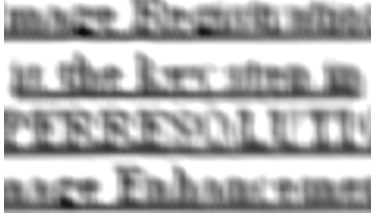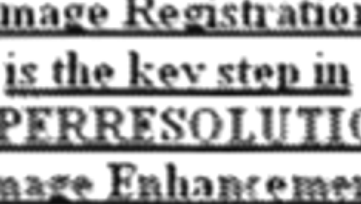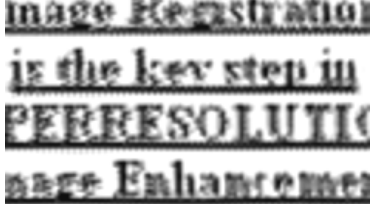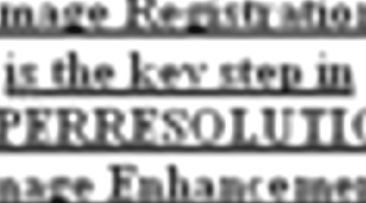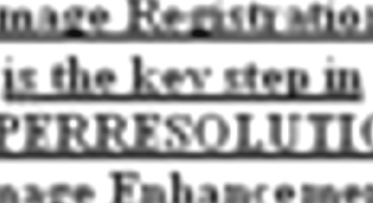
To have a better understanding on the methods and their individual performances, we need to give our attention to the visual assessment in the next chapter.
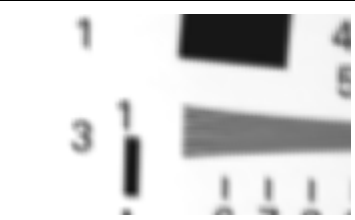
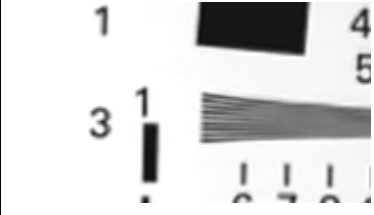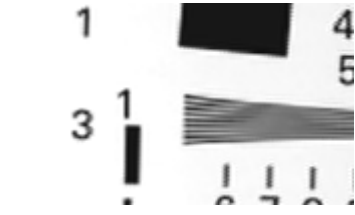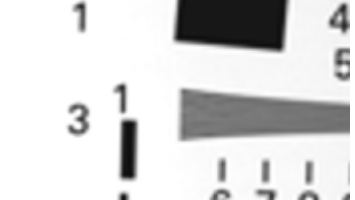## 3.3.6.3. VISUAL ASSESSMENT OF THE METHODS
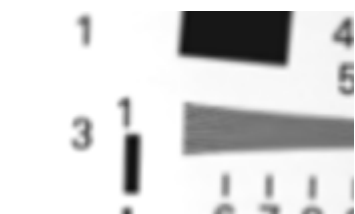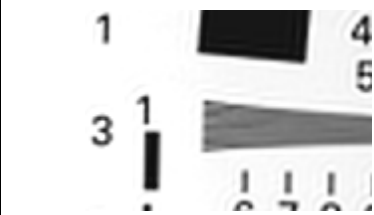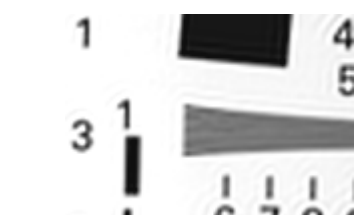
The evaluation of the SR methods is not only based on quality metrics. The use of quality metrics is possible if only we get a control image of the scene or the LR images are synthetically acquired, since we need a reference frame to compare the results of the algorithms. For this reason, if we have a video sequence or a sequence of images visual assessment is essential.

For the synthetic images, Table 3.3 and Table 3.4 contain some details of the super-resolved images. According to the Table 3.3 and 3.4 direct addition methods and IBP with nonuniform interpolation suffers from the blurring. On the other hand, the other three methods are subject to degradation of quality caused by noise. The IBP, nonuniform interpolation and POCS methods produce better results than others do.

Table 3.3 Comparison table for synthetic images for different methods Test1

Method	Perfect Registration	Keren Registration
Direct Addition with Median Filtering	image Registration is the key step in PERRESOLUTIO nage Enhancemen	image Registration is the key step in PERRESOLUTIO nage Enhancemen
Nonuniform Interpolation	image Registration is the key step in PERRESOLUTIO nage Enhancemen	image Registration is the key step in PERRESOLUTIO nage Enhancemen
Iterative Backprojection (IBP)	image Registration is the key step in PERRESOLUTIO nage Enhancemen	image Registration is the key step in PERRESOLUTIO nage Enhancemen
IBP with Nonuniform Interpolation	image Registration is the key step in PERRESOLUTIO nage Enhancemen	image Registration is the key step in PERRESOLUTIO nage Enhancemen
Projection onto Convex Sets (POCS)	image Registration is the key step in PERRESOLUTIO nage Enhancemen	image Registration is the key step in PERRESOLUTIO nage Enhancemen
Bicubic Interpolation	image Registration is the key step in PERRESOLUTIO nage Enhancemen	image Registration is the key step in PERRESOLUTIO nage Enhancemen

Table 3.4 Comparison table for synthetic images for different methods Test2

Method	Perfect Registration	Keren Registration
Direct Addition with Median Filtering		
Nonuniform Interpolation		
Iterative Backprojection (IBP)		
IBP with Nonuniform Interpolation		
Projection onto Convex Sets (POCS)		
Bicubic Interpolation		

In table 3.5 four patterns of three videos are subject to evaluation. First, one is the ventilation fan of a window, second is a structure's reflection on the window, and third is a detail of a building and the last one is the backside of a bus.

According to the patterns on table 3.5, the sharpness is only present in IBP and POCS results. Especially, IBP performs very well and gives clear images of the subjects. However POCS results are seem to be unclean, which is expected, due to the registration errors present on registered images. On the other hand, POCS results contain sharper edges of the subjects. Finally, the performance of the single frame, bicubic interpolation method is not very well with respect to IBP and POCS. The reason to this is the fact that in single frame methods in the presence of noise the contents of those particular pixels where noise is present looses information. As a result, the information loss also defects the neighboring pixels due to the interpolation structure told before. Therefore, the lost pixel both corrupts its belonging place and the pixels surrounding it. However, multi-frame methods have lots of information on particular places of the image. As long as the image registration is perfect, the lost pixels can be recovered by using the other frames in which the noise may not corrupted the data contents on that pixel.

Table 3.5 Comparison table for video frames for different methods

Method	Pattern 1	Pattern 2	Pattern3	Pattern4
Direct Addition with Median Filtering				
Nonuniform Interpolation				
Iterative Backprojection (IBP)				
IBP with Nonuniform Interpolation				
Projection onto Convex Sets (POCS)				
Bicubic Interpolation				

3.3.6.4. NOISE vs. IMAGE QUALITY

Noise on the image has many negative effects on the image quality. As for the image registration, the noise also attenuates the image reconstruction performance. In addition to the image characteristics, the random noise present on the image and the level of noise affects the performance of every method in a different manner. Table 3.6 shows the performance of the methods with respect to noise, for the test1.bmp image; POCS performs better than any other algorithm including IBP. However, as in the Figure 3.19 that the test2 image is used, the POCS algorithm performs very well for the first iteration and then the noise corrupts image and many artifacts appears, so the performance slightly decreases. But the IBP method has a better performance in case of test2.bmp image is used.

Table 3.6 Comparison table for noisy and zero noise image SR(test1)

		Median Filtering	Non-uni. Interp.	Iterative Backproj.	IBP with NUI	POCS	Bicubic Interp.
SSIM	0 noise	0.449	0.358	0.695	0.433	0.750	0.671
	20db Noise	0.451	0.399	0.677	0.486	0.724	0.645
PSNR	0 noise	10.174	9.522	12.114	10.664	12.762	11.992
	20db Noise	10.218	9.793	11.938	10.392	12.364	11.738

Nevertheless, the LPF affect of the IBP and high correction capability of POCS performs better than any other method for the noisy image sequences. Just like video frames that have no fixed behaviour of corruptive effects on them, the noise is random at the synthetic images. Therefore, the noise recovery accomplished and the results are similar for the video and the synthetic images. As a result, noise corrupts the results but still SR methods are more durable than the single frame methods, which looses all available information as the noise exists. Consequently, methods using error correction such as IBP and noise suppression such as direct addition with median filtering are able to construct successful HR estimates.
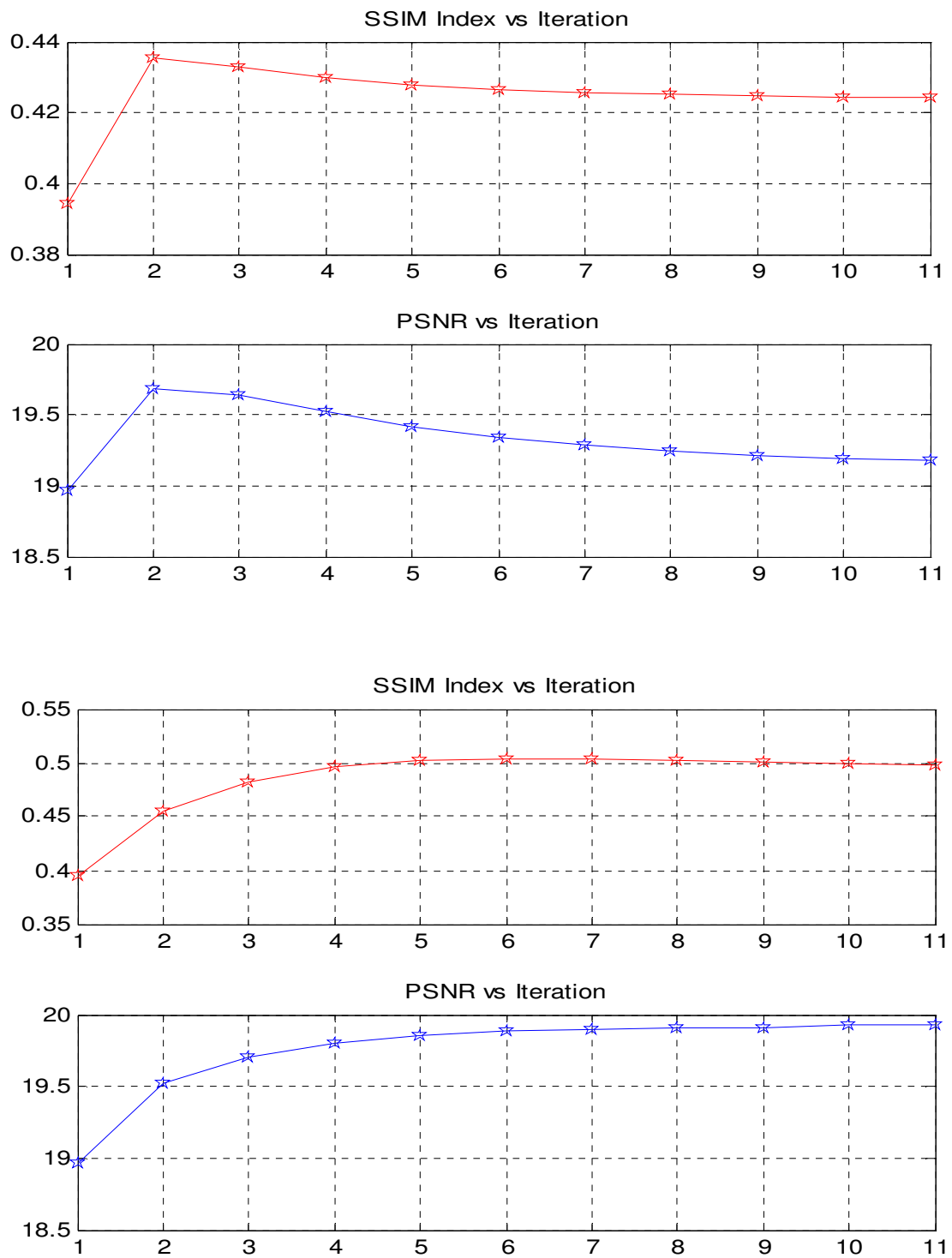
Figure 3.19 PSNR and SSIM values, as the noise present, first two graphs are for POCS and the other two are for IBP (test2.bmp)

3.3.6.5. IMAGE QUANTITY vs. IMAGE QUALITY

Every method uses the source frames in a different way. Therefore, the results of increasing the number of images may differ from method to method. The examinations are completed with reschart image, by increasing the synthetic image quantity.
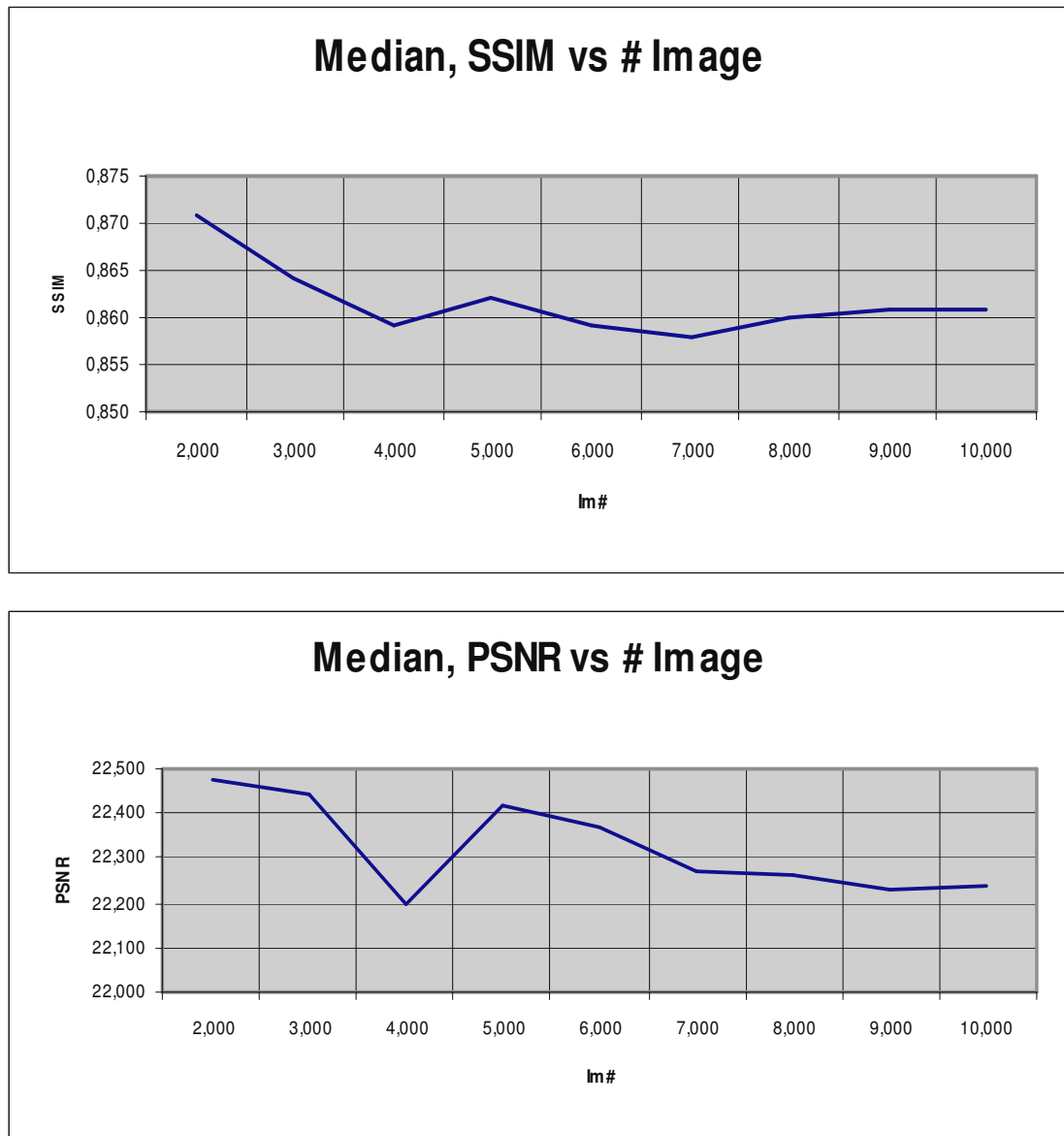
Figure 3.20 Effect of the image quantity on median filtered SR

Addition of the images by median filtering the data set is the most basic form of the SR methodology. As it is seen in Figure 3.20, the quality slightly decreases. This

may be caused by the fact that as we add more images, if the image does not contain new information, the previous images positive effects on the result decreases.
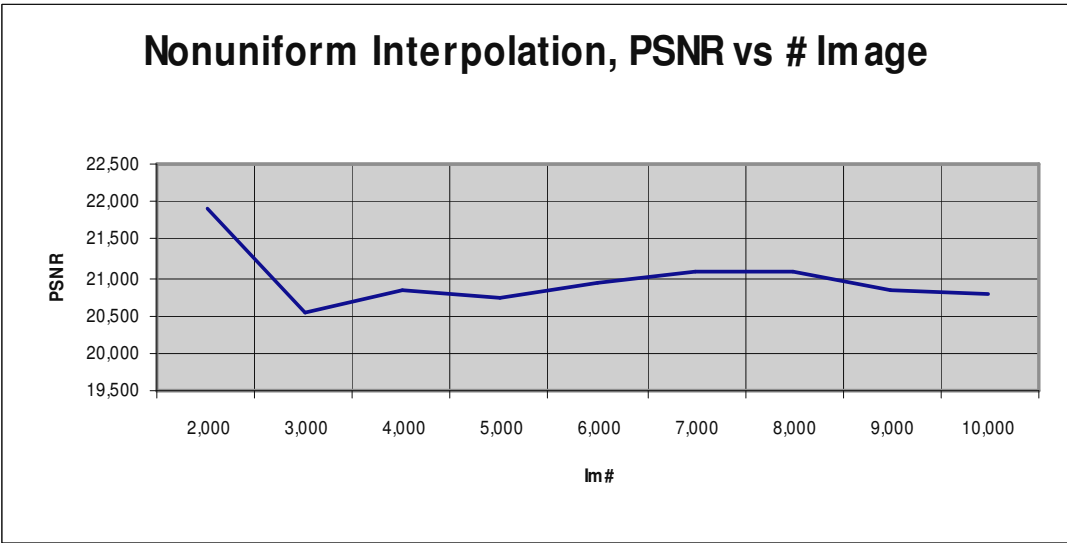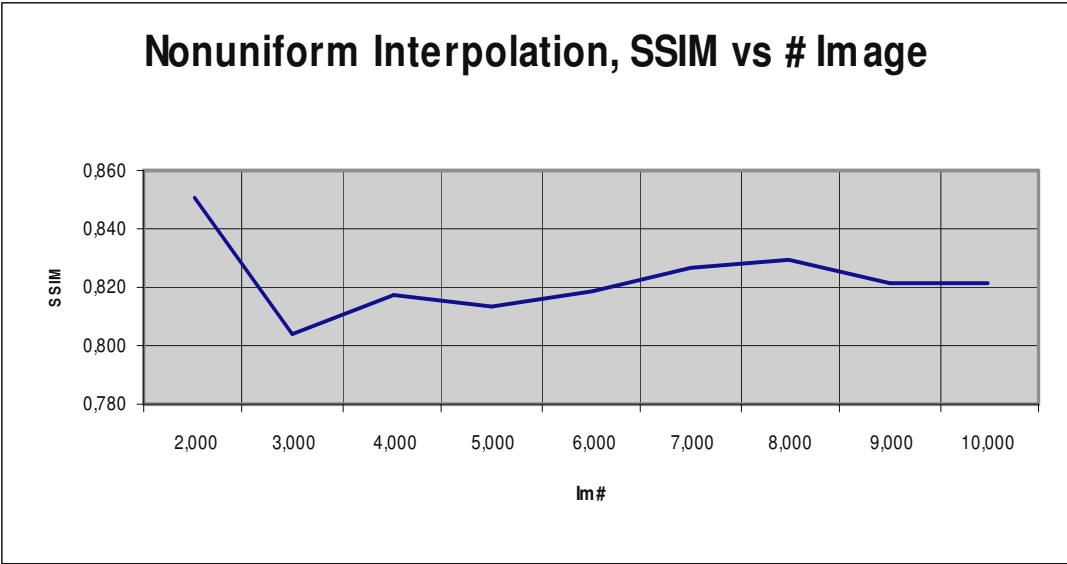
Figure 3.21 Effect of the image quantity on nonuniform interpolation

Nonuniform Interpolation Algorithm looses details in some extent but by addition of new high frequency terms, it gains some quality after the second image.

Iterative Backprojection, SSIM vs # Image

Iterative Backprojection, PSNR vs # Image
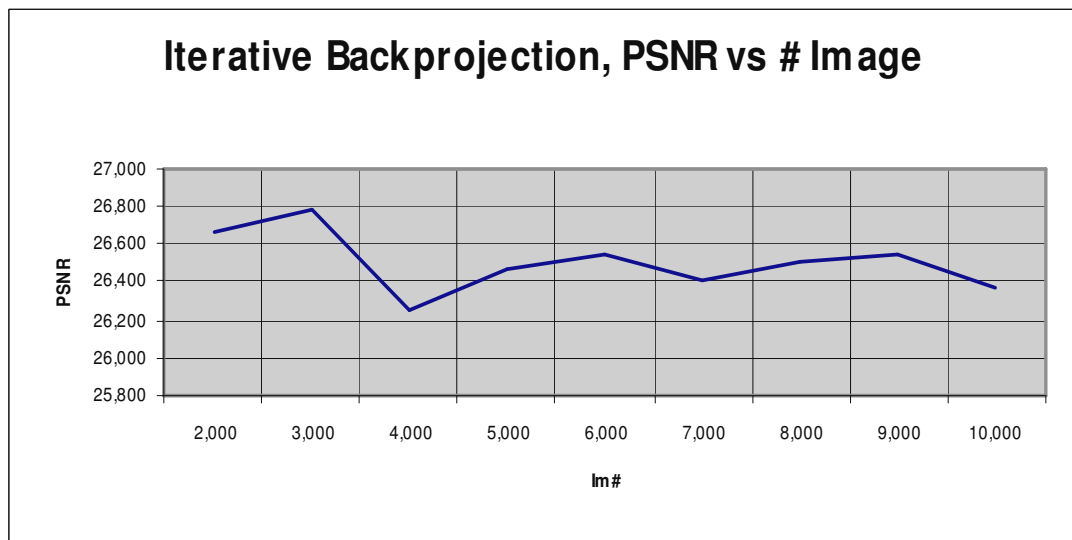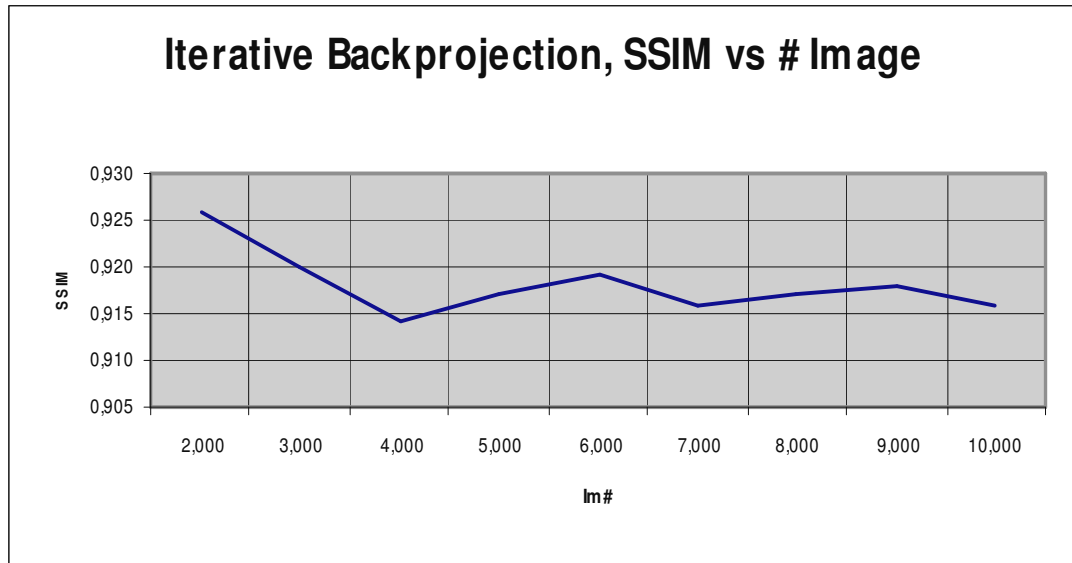
Figure 3.22 Effect of the image quantity on IBP method

As long as the new coming images contain false, inaccurate or similar data to the present values there will not be a drastic improvement in the resultant image quality for all of the methods including IBP.
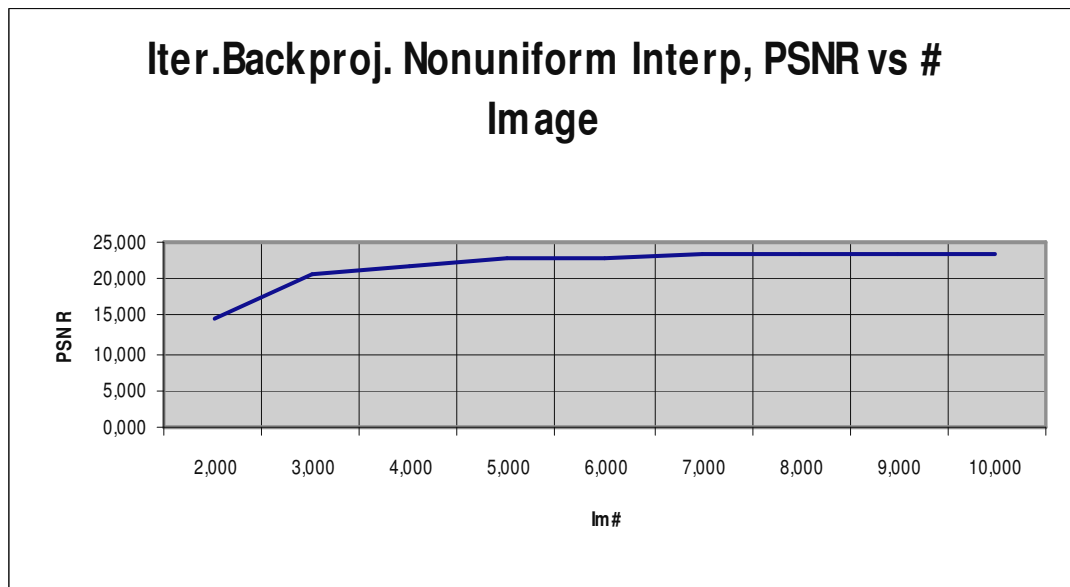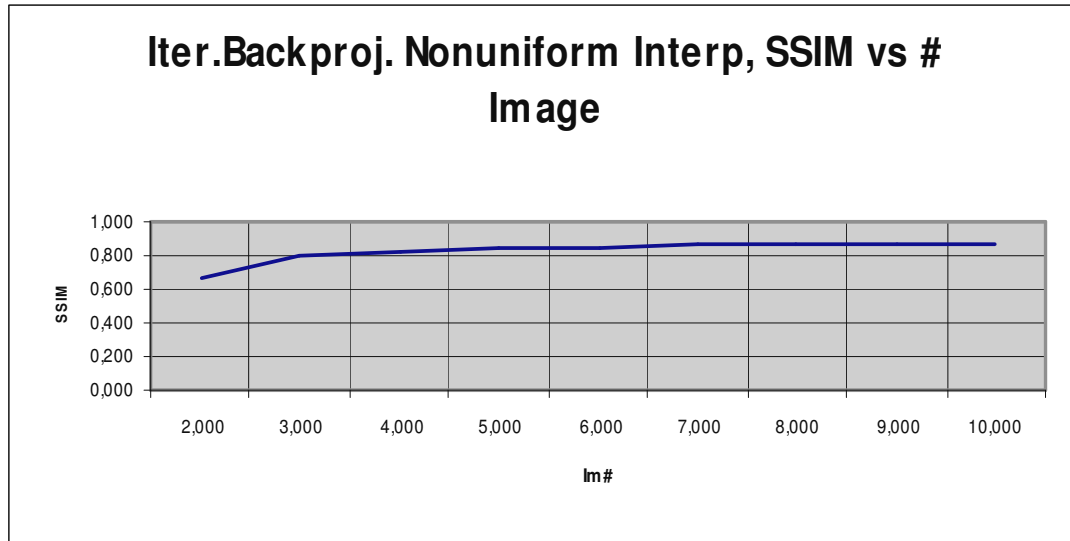
Figure 3.23 Effect of the image quantity on IBP with nonuniform interpolation

The IBP with nonuniform Interpolation result improves as the added images increases, this is because, the first iteration of the algorithm is rather low quality, every added image improves the quality by transferring new details to the result. After some extent, the improvement will slow down and stops due to the lack of new information addition to the image, since every image we add, will have a higher possibility of being similar to the existing ones.
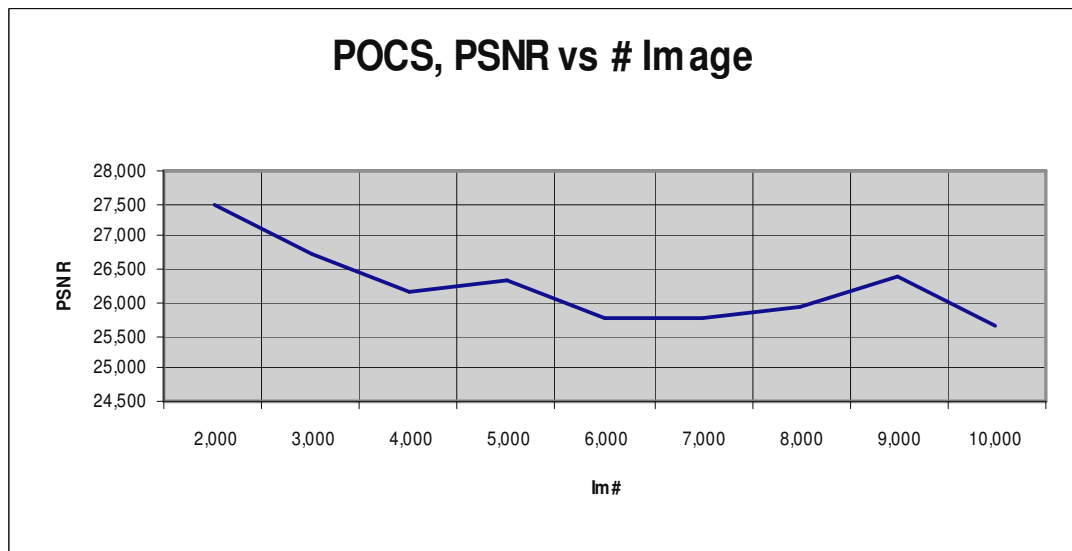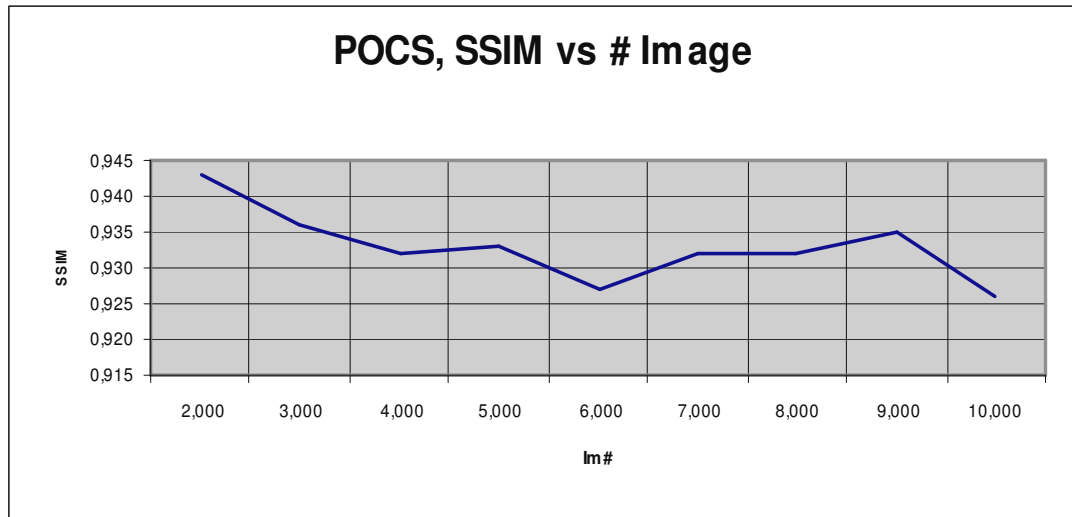
Figure 3.24 Effect of the image quantity on POCS

As in the previous cases, POCS algorithm also does not show an improvement with new images.

Almost all of the methods remain similar as new images added. This may be a result of the synthetic image production. Theoretically, the new images will have new details and information on them and as the image quantity increases, the image quality must be improved, similarly.

3.3.6.6. ITERATION NUMBER vs. IMAGE QUALITY

As the final test the effect of the order of iterations, will be examined. The images used are test1.bmp images and they are registered by Keren algorithm.
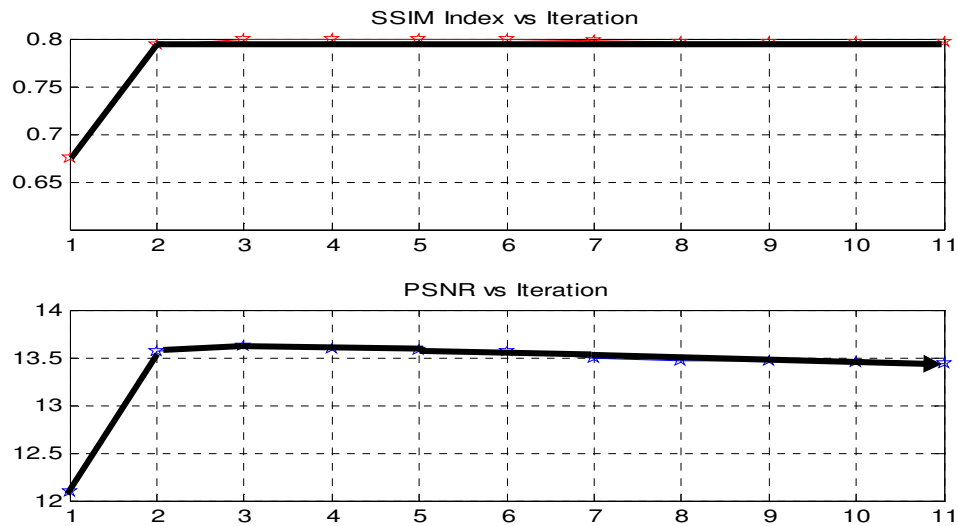
Figure 3.25 Iteration vs Image Quality graph for POCS
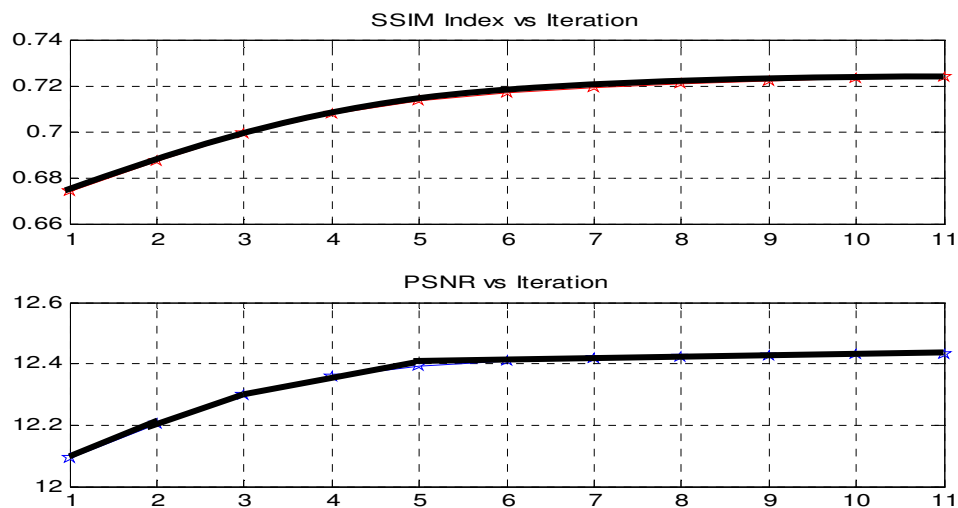
Figure 3.26 Iteration vs Image Quality graph for IBP
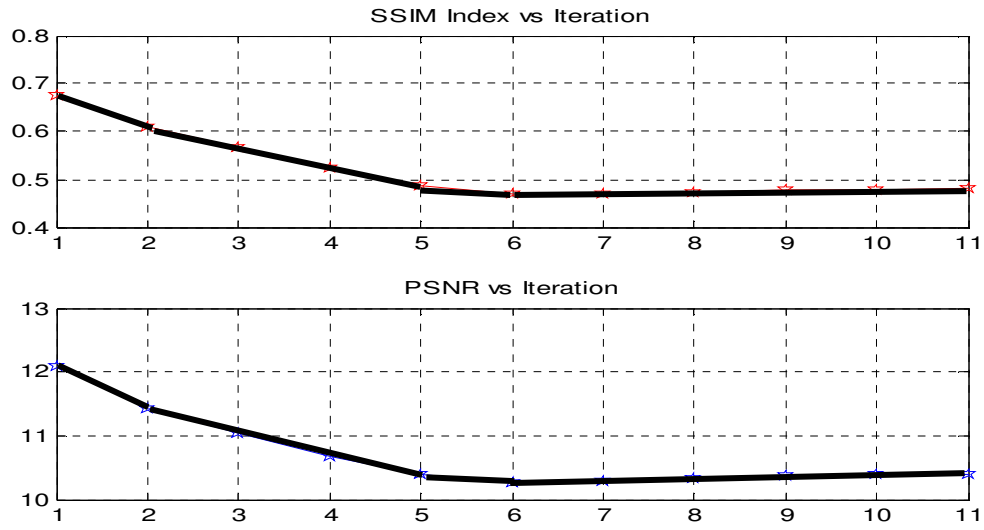
Figure 3.27 Iteration vs Image Quality graph IBP with
Nonuniform Interpolation

The increasing number of iterations improves POCS and IBP methods. Nevertheless, the improvement stops at the end of the second run for the POCS algorithm. This is because the algorithm converges rapidly. Moreover, just after the convergence, the artifacts are present and the artifacts cannot be distinguished from the fine details so the algorithm seems to disturb the image quality as the iterations grows, but visually the details are more visible. For the IBP method, the algorithm converges more slowly as every run reduces the error level slightly. However, for the IBP with Nonuniform Method the error rapidly increases as both SSIM and PSNR indexes are decreased in a few runs. Therefore, the nature of this method seems to multiply the misregistrations and noise instead of minimizing. As a result, for both IBP and POCS methods, we may increase the number of iterations but as the information available is used completely the images may begin to loose quality instead of preserving or increasing the quality level.

CHAPTER 4

DISCUSSIONS

The super resolution is the name of a methodology that contains a wide range of image processing tools in it. Registration algorithms, fusion methodologies, restoration techniques are all the subjects of super resolution. This study especially focuses on the effects of registration and fusion techniques on resolution enhancement by implementing some well-known algorithms in the literature and experimenting on them.

Image registration is a crucial step of SR methodology and it is a well-known state of art. During the study, five registration methods are examined. Three of them are the frequency domain methods, and the other two are spatial domain techniques. The stability of the spatial domain techniques makes us to prefer them instead of the frequency domain counterparts. The main reasons to this preference are the fact that the spatial domain techniques are computationally efficient and give us very reliable results. However, the frequency domain techniques are unstable and may converge to a false value. Since the registration is the key point in SR, we usually choose the RANSAC for the video frames and the algorithm of Keren et al for synthetic images. Since the video frames may contain many kinds of transformations, thet are not only planar but also in three-dimesional (zooming, shearing etc.). Keren method cannot handle these motions and does not give us a useful result. Therefore we preferred to use RANSAC method. On the other hand, synthetic images may be too small sized to find enough corners for RANSAC algorithm to have a correct solution; as a result, we prefered to use method of Keren et al. in synthetic images.

In addition, examinations show that the RANSAC methodology can give results covering whole bunch of transformation probabilities, but the noise present on the image may make the algorithm fail. At that point, the use of Keren will find the transformations effectively if there is no scaling present on the images. As a result, the use of both methods and then deciding to use one of them may be solution depending on the visual assessment of the video or the image sequence.

Secondly, the image fusion techniques are of our interest. The techniques of covering all information available and combine them into a HR image are examined. In the meantime, the single frame interpolation techniques are also considered since they work very effectively and give good results without any concern about the weight of the computational load. The results of the examinations on the techniques show that, three important factors are affecting the performance of multiframe superresolution restoration methods. These are:

1. Accurate and precise subpixel-resolution motion information is required.
2. The model must contain the knowledge of imaging system degradations to overcome the possible problems occurs during image acquisition.
3. The fusion method must include well-defined priori information.

The effectiveness of the registration techniques is vital for the sake of the SR product. Especially set theoretic POCS method and the nonuniform interpolation method use the information of the all images very effectively and even a slight registration error causes the product to loose efficiency. POCS forces the HR image to be bounded by a range of pixels determined from LR images. If there are registration errors, this causes the bounds to be placed at false pixel locations, leading to ghosting and ringing effects on the image. As well as POCS, nonuniform interpolation also uses the false pixels without any error correction, which leads us to poor quality in the product.

The Iterative Back Projection method turned out to be one of the most successful methods in our study. The reason to this is its effectiveness in elimination of the various errors at the backprojection stage. IBP process can be equated to solution of

a jigsaw puzzle. Each puzzle piece is gathered from different LR observations. These pieces are superposed first and glued to each other to remove the edge inconsistencies. The back projection kernel smooths the reconstruction error.

Generally, the multi-frame methods perform very well in video frames. The improvement of images is clear especially for the IBP and POCS methods. The number of images does not strongly affect the solution in case of using synthetic images. The quality saturates eventually. Visual quality improvement with higher number of images is more evident in video frames.

Additionally, noise is another problem for all methods. We may attempt to remove the noise before the application of super resolution. Nevertheless, this may cause the loss of detail that we need for the super-resolution reconstruction. A second approach is applying noise-filtering algorithms after the super-resolution stage. This can be more suitable if the reconstructed details can be preserved after filtering. The Addition methods, and IBP with non-uniform interpolation are effective in noise removal; but cannot produce details as good as POCS, nonuniform interpolation and IBP. This leads them to the lack of performance in QM. On the other hand, even IBP and especially POCS is also noise prone, they are able to come over the noise better than any other algorithm.

As the result of the experiments, the combination of image registration methods; Keren or RANSAC, and the image fusion methods; IBP or POCS, will leads us to quite satisfactory results compared to other multi-frame and single-frame methods.

CHAPTER 5

CONCLUSIONS

In this thesis, we have examined methods of image reconstruction from multiple images. Our goal is combining multiple low-resolution images to construct a single high-resolution image. The problem is complicated by nature due to the unknown point spread function of imaging sensors, varying imaging conditions and due to the need of sub-pixel level relative motion estimation. All these factors make super-resolution problem a difficult inverse problem with a non-unique solution.

We have observed that image registration or the estimation of the motion parameters is the performance-determining step of super-resolution. Misaligned images may cause loss of detail or ghosts instead of improving the resolution. Therefore, one may prefer single frame interpolation methods when the image set cannot be properly registered.

Under tolerable noise conditions as long as a good registration can be established, POCS and iterative back projection (IBP) methods provide better results than interpolation based single frame methods. The gains are marginal according to PSNR or SSIM metric; but quality improvement is visually distinguishable. The gains are especially evident for video sequences.

We have noted that POCS is more suitable for synthetic data or for images with strong edge content, while the iterative back projection (IBP) method is suitable for real world sequences such as the ones captured by an ordinary video camera.

In the presence of noise and misregistrations, the performance of POCS and iterative back projection degrades. The degradation of POCS is more severe than iterative back projection. This is due to inherent pixel bounding operation of POCS and its critical dependence of motion estimation. Therefore, addition of noise results in double threat to POCS. The iterative back projection is also affected by the registration errors, but due to its back projection operation the error is smoothed at every iteration leading to less obtrusive results.

Qualitatively speaking, the super resolution methods combine the details of low quality images to extract information on unknown pixels of high quality image. The signal detail is sensitive to the noise level in the system. If the noise in the system is above a certain threshold, it can be better to use single frame methods with low-pass filtering to have better quality images. It is not clear where this threshold is for the super-resolution algorithms present in the literature. An interesting study can be the determination of the breakpoint or point of no-gain for the super-resolution methods. A second interesting study can be the joint denoising and super-resolution of images. This can be valuable if the detail and noise can separated during the super-resolution process. However, before such a study, the effectiveness of super resolution should be clearly illustrated under noiseless conditions.

REFERENCES

[1] P. Vandewalle, L. Sbaiz, M. Vetterli, and S. S¨usstrunk, "Super-resolution from highly   undersampled images," in IEEE International Conference on Image Processing, September 2005.

[2] P. Vandewalle, S. S¨usstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," accepted to EURASIP Journal on Applied Signal Processing, Special Issue on Super-Resolution Imaging, 2005.

[3] D. Capel and A. Zisserman, "Computer vision applied to super resolution," IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 75–86, May 2003.

[4] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacement," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, June 1988, pp. 742–746.

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

[6] C. J. Harris and M. Stephens. "A combined corner and edge detector" In Proc. Alvey Vision Conf., pages 147–151, 1988.

[7] B. Marcel, M. Briot, and R. Murrieta, "Calcul de translation et rotation par la transformation de Fourier," Traitement du Signal, vol. 14, no. 2, pp. 135–149, 1997.

[8] L. Lucchese and G. M. Cortelazzo, "A noise-robust frequency domain technique for estimating planar roto-translations," IEEETransactions on Signal Processing, vol. 48, no. 6, pp. 1769–1786, June 2000.

[9] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, vol. 24, no. 6, pp. 381–395, June 1981.

[10] H. Stark and P. Oskoui, "High resolution image recovery from image plane arrays, using convex projections," J. Opt. Soc. Am. A, vol. 6, pp. 1715-1726, 1989.

[11] A.M. Tekalp, M.K. Ozkan, and M.I. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), San Francisco, CA., vol. 3, Mar. 1992, pp. 169-172.

[12] S. Borman and R.L. Stevenson, "Super-resolution from image sequences A review," in Proc. 1998 Midwest Symp. Circuits and Systems,1999, pp. 374-378.

[13] Z. Wang and A. C. Bovik, "A universal image quality index," IEEESignal Processing Letters, vol. 9, pp. 81–84, Mar. 2002.

[14] M. Irani and S. Peleg, "Improving resolution by image registration," CVGIP: Graphical Models and Image Proc., vol. 53, pp. 231-239, May 1991.

[15] R.G. Keys, ``Cubic convolution interpolation for digital signal processing,'' IEEE Trans. Acoust., Speech, and Signal Process., vol.29, pp. 1153--1160, 1981.

[16] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical review," IEEE Signal Processing Mag., vol. 20, pp. 21–36, May 2003.

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

[18] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "The SSIM Index for Image Quality Assessment" http://www.cns.nyu.edu/~lcv/ssim/ 2003

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

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

[21] S. P. Kim and W. Su, "Recursive high-resolution reconstruction of blurred multiframe images," in Proceedings of the IEEE International Conference on

Acoustics Speech and Signal Processing, Toronto, Canada, May 1991, vol. 4, pp. 2977–2980.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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