ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

HIGH-RESOLUTION IMAGE RECONSTRUCTION USING SUPER-RESOLUTION AND DECONVOLUTION TECHNIQUES

HE YU

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2009

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

To my family, for all your love.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Assistant Professor Kim-Hui Yap, for his patient guidance, helpful instructions on this research and continuous encouragement through the time of my study towards the Doctor of Philosophy degree. His knowledge, wisdom, kindness and spirit have benefited me greatly.

My sincere gratitude also goes to Associate Professor Lap-Pui Chau, my co-supervisor for his guidance, valuable advice and kind encouragement.

Further, I would like to thank my fellow colleagues Dr. Chen Li and Mr. Tian Jing for their help in my research work. I am also thankful to the technicians of the Media Technology Lab for their kind assistance, support and understanding.

Finally, I am very grateful to my parents for teaching me the importance of hard work and perseverance. I would like to thank my sister and my girlfriend, whose love and support are always the greatest inspiration for me. Appreciation is also extended to all my friends who have given me the strength to go on.

Thanks again to all the people who have helped me in one-way or another.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Contents

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Summary

This thesis investigates the techniques for high-resolution (HR) image reconstruction through image super-resolution (SR) and deconvolution algorithms. In many imaging applications, HR images are desired, but may not be available due to practical limitations. Using HR reconstruction, it is possible to restore high-frequency content, reduce noise, and increase spatial resolution when hardware modification is unrealizable.

In the first part of this thesis, we address the multi-frame SR from two points of view: blur identification and motion estimation. Multi-frame SR is the fusion of a set of LR images to produce a HR image. A main issue arising from the multi-frame SR problem is the availability of the knowledge on the blurring occurred during the image formation process. In many real-world applications, it is often difficult to know the blurring completely *a priori*. In view of this, we present a new soft maximum *a posteriori* (MAP) estimation framework to perform joint blur identification and HR image reconstruction. The proposed method incorporates a soft blur prior that estimates the relevance of the best-fit parametric blur model, and induces reinforcement learning towards it. An iterative scheme based on alternating minimization is developed to estimate the blur and HR image progressively. Another critical step in multi-frame SR is accurate motion estimation of the LR images, particularly when the motion is not purely translational between the LR images. In view of this, we present a new framework that performs simultaneous image registration and HR image reconstruction in Chapter 4. This is promising as more accurate motion parameters can be determined, thereby enhancing the performance of the HR reconstruction. Further, it is worth noting that unlike most algorithms that focus on the translational motion model, the proposed method adopts a more generic motion model that includes both translation as well as rotation. An iterative scheme is developed to solve the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

arising nonlinear least squares problem.

In the second part of this thesis, we develop a new learning-based single-frame face SR method, which attempts to estimate missing HR details from a single face image by learning from training examples. The SR problem is challenging, as there is less observed data available and we have to take into account the specific features of human faces. An improved principle component analysis (PCA)-based method is used to synthesis a global HR face, followed by a patch-based residue prediction step for compensating detailed facial features. Compared with the conventional PCA-based method, the new method can be applied to new data more effectively, and a better HR image that bears closer resemblance to the original face while preserving usual characteristics of a human face. Furthermore, the estimation of the residual patch does not depend on a single nearest neighbor in the training set. Instead, it depends simultaneously on multiple nearest neighbors in a way similar to locally linear embedding for manifold learning.

The last part of this thesis deals with blind color image deconvolution, which aims to restore a single degraded color image at the same pixel resolution, given limited or no prior knowledge of the burring function. It is considered as a multi-input multi-output (MIMO) deconvolution problem, where intrachannel and interchannel blurring are taken into consideration. We present a new algorithm to address this problem based on a hybrid framework of single-input single-output (SISO) and single-input multi-output (SIMO) models in Chapter 6. We emphasize the importance of exploiting both spatial and spectral correlations in color image deconvolution and use a combination of SISO and SIMO blind algorithms to alleviate the intractable original MIMO problem.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
AM	Alternating minimization
AR	Autoregressive
ARMA	Autoregressive moving average
AWGN	Additive white Gaussian noise
CCD	Charged-coupled device
CFA	Color filter array
CFT	Continuous Fourier transform
CG	Conjugate gradient
DFT	Discrete Fourier transform
DR	Double regularization
EM	Expectation maximization
EVAM	Eigenvector-based algorithm
FP	Fixed point
GCV	Generalized cross-validation
HR	High-resolution
IBD	Iterative blind deconvolution
IBP	Iterative backprojection
KL	Karhunen-Loeve
k-NN	*k*-nearest neighbor
LLE	Locally linear embedding
LR	Low-resolution
MA	Moving average
MAP	Maximum *a posteriori*
MC	Multi-channel
MCMC	Markov Chain Monte Carlo
MCSC	Minimum cyclic-shift correlation

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

MIMO	Multi-input multi-output
ML	Maximum likelihood
MMSE	Minimum mean squared error
MRF	Markov random field
NAS-RIF	Nonnegative and support constraints-recursive inverse filtering
NLS	Nonlinear least squares
NMSE	Normalized mean squared error
PCA	Principle component analysis
PDE	Partial differential equation
PDF	Probability density function
POCS	Projection-onto-convex-sets
PSF	Point spread function
PSNR	Peak signal-to-noise ratio
RDWT	Redundant discrete wavelet transform
SAR	Simultaneously autoregressive
SC	Single-channel
SIMO	Single-input multi-output
SISO	Single-input single-output
SNR	Signal-to-noise ratio
SR	Super-resolution
TLS	Total least squares
TV	Total variation
VP	Variable projection
VQ	Vector quantization

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Figures

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Tables

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 1

Introduction

1.1 Motivation

Digital imaging devices are growing in popularity and commonly used in various imaging applications in recent years. Although a natural scene has infinitely many levels of details, a typical digital imaging device can only capture the scene at a limited range of spatial resolution. In many applications, there is a demand for obtaining high spatial resolution images. In general, the term "*resolution*" of an image often refers to the number of pixels in a digital image. Thus, it is often named as "*pixel resolution*". However, the concept of "*spatial resolution*" is more than this. It is also related to the ability to distinguish details in an image. For example, we take a digital image and scale it up by repeating its pixel data. It is evident that the spatial resolution of the image has not been increased. Therefore, images with higher spatial resolution not only have higher spatial density of pixels, but also have stronger resolving power. In other words, the details in the high spatial resolution image should be more discernable. For simplicity, the term "*resolution*" appearing in the rest of the thesis is referred to as "*spatial resolution*".

In many imaging applications, high quality and high-resolution (HR) images are desired, but may not be available due to practical limitations. Blurring, frequency aliasing and noise are common degradations found in digital imaging systems. For example, in a traffic surveillance application, a video-based vehicle tracking system tracks moving vehicles through a camera. The original scene is shown in Fig. 1.1(a). The captured video frame is usually distorted by

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

motion blurring due to rapid vehicle motion, or by optical blurring since the camera is fixed so that the sensor plane may not be in focus. The effect of blurring can be observed in Fig. 1.1(b). During the spatial sampling of the continuous scene, the tracking system may potentially produce an aliased low resolution (LR) image in Fig. 1.1(c). This often takes place when the target vehicle is far away from the camera. Moreover, it is noted that the imaging sensor array, such as charge-coupled device (CCD), is subject to various noise sources including the thermal and shot noise. The noisy effect is particularly evident under the environment of low lighting as shown in Fig. 1.1(d). Finally, the blurred, aliased and noisy LR image is shown in Fig. 1.1(e). It is obvious that under less than ideal situations, the image quality of the captured images through the imaging devices could be degraded severely.

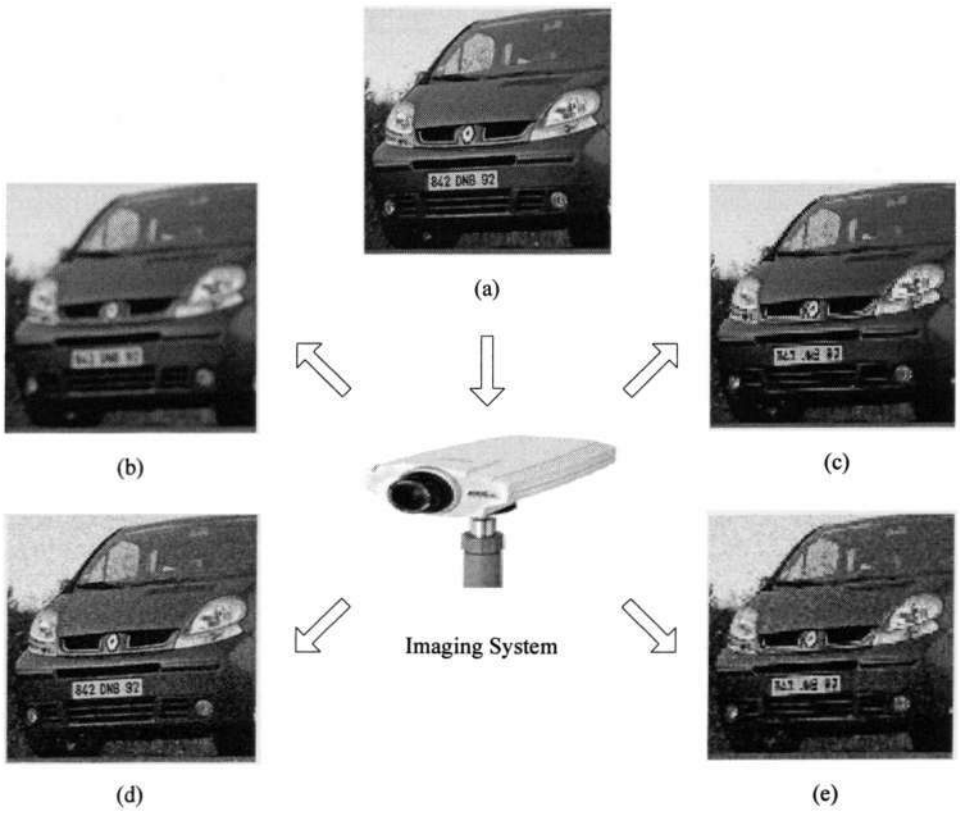


Fig. 1.1: Effects of various imaging degradations. (a) Original image, (b) Blurred image, (c) Aliased image, (d) Noisy Image, (e) Blurred, aliased and noisy LR image.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

To avoid the aforementioned degradations, the most direct solution is to improve the hardware by using an advanced optical system or CCD with more pixels. However, the advanced hardware can only improve the image quality to a certain level due to the practical constraints. The presence of shot noise, which is unavoidable in any imaging system, is associated with an upper limit on the resolution of CCD [1]. This upper limit arises from the fact that while reducing the sensor pixel size increases the resolution, the signal strength is correspondingly decreased, leading to a severe noise degradation of the image quality. Aside from the approaching limit due to the noise effect, cost is another concern in using high precision optics and image sensors in many commercial applications. Finally, another impediment to the use of the advanced hardware regarding HR image is that the imaging is often done under less than ideal situations, such as military surveillance or remote sensing. Hence, hardware modification is not always a feasible option. This motivates the study of employing software approaches for HR image reconstruction to overcome these limitations of the hardware.

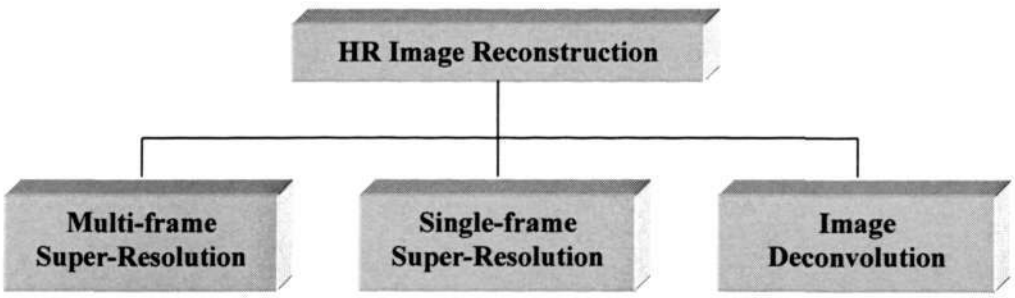


Fig. 1.2: The focused subjects for HR image reconstruction.

This thesis will investigate how to reconstruct a HR image through the use of image processing algorithms. We focus on two topics involved in HR image reconstruction, namely image super-resolution (SR) and deconvolution as shown in Fig. 1.2. The objective of image SR

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

is to reconstruct HR images from aliased, blurred and noisy LR images, while the main goal of image deconvolution is to restore a blurred and noisy image, without increasing the pixel resolution. It is noted that image SR can generally be divided into two categories, multi-frame SR and single-frame SR. Multi-frame SR is the fusion of a number of LR images to produce a HR image. An illustration of multi-frame SR is shown in Fig. 1.3. The LR images with subpixel shift consist of an overlapping content of a scene, thus there is new information in the observations that can be used to reconstruct a HR image. On the other hand, the objective of single-frame SR is to estimate missing HR details from a single observed LR image, which could be learned from training set that consists of similar images to the target HR image. Fig. 1.4 illustrates a general process for single-frame face SR, in which a large database of facial images is used as the training set. Thus single-frame SR is often expressed in terms of "*learning-based image SR*" or "*example-based image SR*" [2]. Not surprisingly, image SR and deconvolution are closely related problems. In fact, image deconvolution is simply a special case of image SR.

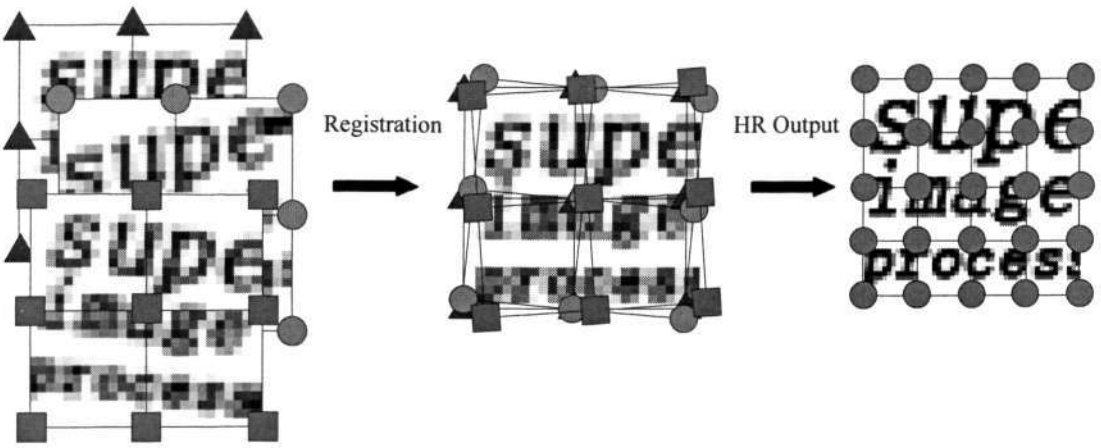


Fig. 1.3: Illustration of multi-frame SR.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

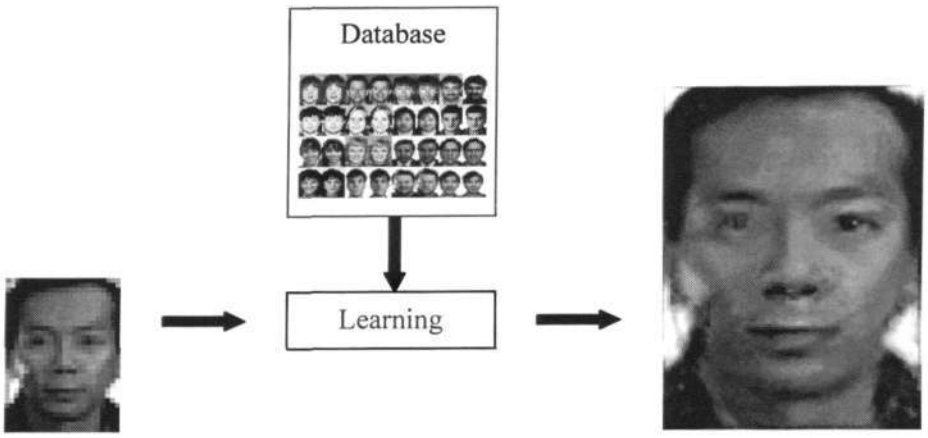


Fig. 1.4: Illustration of single-frame SR.

HR image reconstruction has been shown to be useful in many applications including astronomical imaging, medical imaging, video surveillance, and television signal conversion, amongst others. The first encounter with digital image deconvolution in the engineering community is in the area of astronomical imaging. Imaging systems in satellites are often subject to blurring for some reasons such as out-of-focus optics, spinning and tumbling of spacecraft, and atmospheric turbulence. The obtained images also suffer from the noise arising from the electronic components in the imaging system and broadcast transmission effects. Therefore, image deconvolution techniques are required so as to restore the blurred astronomical images. In addition, in remote image sensing, under-sampling is a common problem for an infrared imaging system. The detector arrays in the cameras are not sufficiently dense for the desired optics to meet the Nyquist criterion [3]. For example, a real infrared "tank" image is shown in Fig. 1.5. It is observed that there are obvious aliasing artifacts around the tanks. One way to overcome the under-sampling problem is to exploit multiple frames from an image sequence. This is possible if there is relative motion between the scene and the sensor during the image sequence acquisition. Thus, multi-frame SR techniques can be employed to provide HR images.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Fig. 1.5: A real infrared "tank" image[1].

In the area of medical imaging, HR image reconstruction certainly plays an important role. The ultrasound imaging, for example, is one of the most widely used imaging modality in the clinical examination. The resolution of ultrasonic images is not only limited by the dimensions of the sound beam from the acoustic aperture, but also degraded by the speckle noise. Fig. 1.6 illustrates one ultrasonic image of a fetus. Enhancing ultrasonic resolution can increase the accuracy of diagnostic decisions, and lessen the burden of well-trained radiographers as well.

Another important application of HR reconstruction is video surveillance. One of the most frequent needs for HR image reconstruction arises when viewing poor quality videos. We have introduced a surveillance application in Fig. 1.1. In the scenario, a selected region on the license plate of the blurred, aliased and noisy image in Fig. 1.1(e) is enlarged in Fig. 1.7. Clearly, the HR reconstruction for the degrade image are needed so as to recognize the car number. Another

[1] The image is shared by the website: http://www.soe.ucsc.edu/~milanfar/software/sr-images.html

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

surveillance example is shown in Fig. 1.8. A suspect in black shirt, who is using a stolen credit card, is caught on the surveillance video in a convenience store. Due to the large distance between the camera and the suspect, the face is too small and unrecognizable. If traditional bicubic interpolation is used to estimate the HR face, the result is shown in Fig. 1.8 (c). We are unable to establish the identity of the suspect. Thus advanced SR technology is desirable so as to produce a recognizable face of the suspect to assist the police investigation.

Fig. 1.6: An ultrasonic image of a fetus in uterus.

It is worth noting that most existing applications of HR image reconstruction are focused on specialized imaging products. HR image reconstruction is becoming a mainstream technique in image processing. Our task in this thesis is to develop efficient HR reconstruction algorithms that are useful for various imaging applications.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

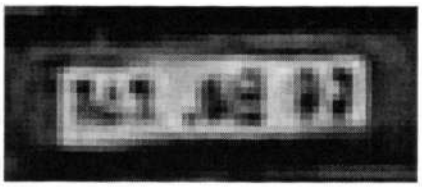
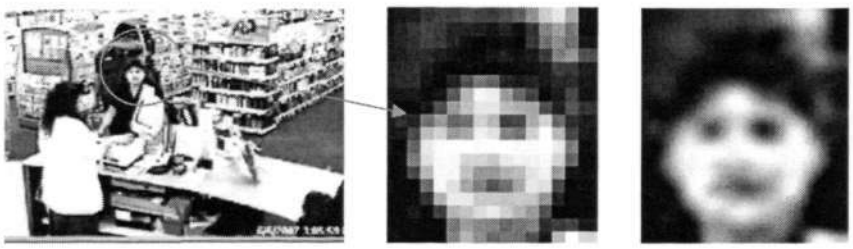


Fig. 1.7 The selected enlarged region on the license plate of the degraded image in Fig. 1.1(e).



(a) (b) (c)

Fig. 1.8: Application for HR image reconstruction in video surveillance. (a) Image captured by the surveillance camera, (b) Enlarged suspect's face, (c) Reconstructed face by using bicubic interpolation.

1.2 Objective

HR image reconstruction attempts to reconstruct HR images from the degraded images in imaging systems. The objective of this thesis is to develop efficient HR reconstruction algorithms through the use of image SR and deconvolution techniques. The forward relationship between the HR image and the corresponding degraded image is illustrated in Fig. 1.9. Following the work in [4] and [5], the general mathematical forward model for HR image reconstruction is linear in form:

$$\boldsymbol{g} = \mathbb{C}\boldsymbol{f} + \boldsymbol{n} \tag{1.1}$$

where $\boldsymbol{g}, \boldsymbol{f}$ and $\boldsymbol{n}$ are the vectors representing the discrete and lexicographically ordered degraded image, original image, and noise, respectively. The matrix $\mathbb{C}$ in the model

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

characterizes the imaging system, which consists of several degradations that affect the quality of the captured images. Different formations of $\mathbb{C}$ could be adopted according to different problems. This will become clear in the following subsections.

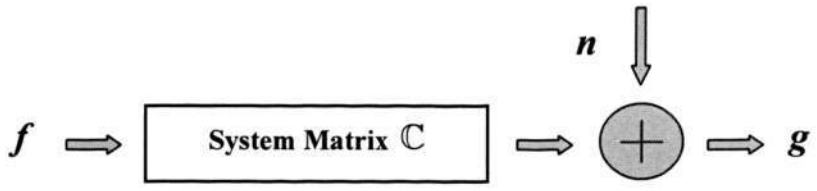


Fig. 1.9: Forward relationship between the HR image $\boldsymbol{f}$ and the corresponding degraded image $\boldsymbol{g}$.

The HR reconstruction problem can be mathematically stated as given $\boldsymbol{g}$, find a best estimate of $\boldsymbol{f}$ by inversing the forward model. It is noted that the estimation of $\boldsymbol{f}$ is a difficult ill-posed inverse problem resulting from the high condition number of the matrix $\mathbb{C}$. Since the system matrix $\mathbb{C}$ is typically ill-conditioned [6], huge and sparse, direct pseudo-inverse technique could make the solution intractable and sensitive to noise, as the uniqueness and stability of the solutions cannot be guaranteed. Thus, to tackle this problem, we should constrain the solution space according to *a prior* knowledge on the form of the solution. Proper prior constraints could turn the HR reconstruction problem into a well-posed one by limiting the space of acceptable HR images.

1.2.1 Multi-frame Super-resolution

Multi-frame SR is the task of estimating a HR image from a set of blurred, aliased and noisy LR images. In this case, the imaging system could include warping, blurring and downsampling processes. Fig. 1.10 illustrates the image degradation process for multi-frame SR problem. Then the k-th LR image can be modeled as:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\boldsymbol{g}_k = \boldsymbol{D}\boldsymbol{H}_k\boldsymbol{S}_k\boldsymbol{f} + \boldsymbol{n}_k, \quad k = 1,2,\cdots,N \tag{1.2}$$

where N is the number of the observed degraded images. $\boldsymbol{g}_k$ and $\boldsymbol{n}_k$ are the vectors representing the discrete and lexicographically ordered k-th degraded image and noise, respectively. $\boldsymbol{S}_k$ represents the warping of the HR image for the k-th LR image. $\boldsymbol{H}_k$ represents the blurring process, such as out-of-focus blurring or atmosphere turbulence blurring, during capturing the k-th LR image. $\boldsymbol{D}$ denotes the downsampling operator. The objective of multi-frame SR is to solve the above set of linear equations to reconstruct $\boldsymbol{f}$.

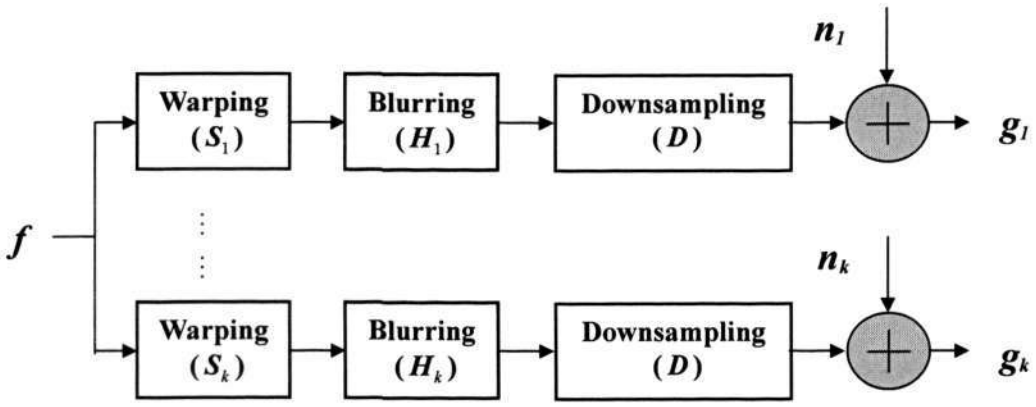


Fig. 1.10: Image degradation process for multi-frame SR.

A main issue arising in this formulation is the availability of the knowledge on $\boldsymbol{H}_k$ and $\boldsymbol{S}_k$. Currently, most SR works assume that the blurring occurred during the image acquisition $\boldsymbol{H}_k$ is negligible or can be characterized fully *a priori*. This assumption, however, is impractical, as it is sometimes difficult to know the blurring function completely *a priori*. Therefore, this motivates the study of blind image SR, where the term "*blind*" denotes that there is limited or no knowledge of the blurring function during the HR estimation. Blind image SR attempts to incorporate blur identification into the SR image reconstruction.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

On the other hand, it should be pointed out that multi-frame SR is possible due to the presence of motion. A critical step in multi-frame SR is accurate registration (motion estimation) of the LR images, particularly when the motion is not purely translational between the LR images. The objective of image registration is to identify the pixel-to-pixel mapping between LR images in their overlapping region, i.e. find the information of $\boldsymbol{S}_k$. Generally, most current SR algorithms ignore registration errors and assume the estimated motion parameters by existing registration methods to be error-free. Nevertheless, due to the presence of aliasing in the captured LR images, most existing registration algorithms for aliased images still experience registration errors. Therefore, this motivates the study of multi-frame SR, which takes into account the impact of unreliable initial registration while using a more flexible motion model.

1.2.2 Single-frame Super-resolution

Single-frame SR, also commonly known as learning-based image SR, attempts to estimate missing HR details from a single image by learning from training examples. Fig. 1.11 illustrates the imaging process for single-frame SR. It is observed that the imaging process does not include the warping operation, as there is only a single observation of the scene available. The observed LR image can be modeled as follows:

$$\boldsymbol{g} = \boldsymbol{DHf} + \boldsymbol{n} \tag{1.3}$$

where $\boldsymbol{H}$ denotes the blurring operator. The task of single-frame SR is to solve the above linear equation with respect to $\boldsymbol{f}$, with the help of a training database that consists of similar images to the target HR image. Another problem related to single-frame SR is image interpolation that is also used to increase the resolution of a single image. However, the missing pixel information for interpolation is estimated from the surrounding pixels of the observed LR image. The

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

quality of an image magnified from an aliased LR image is inherently limited, as there is no new information that can be used. For this reason, image interpolation is not considered as the SR technique. To overcome the limitation of the interpolation technique, single-frame SR utilizes a training database to allow SR reconstruction of the single observation.

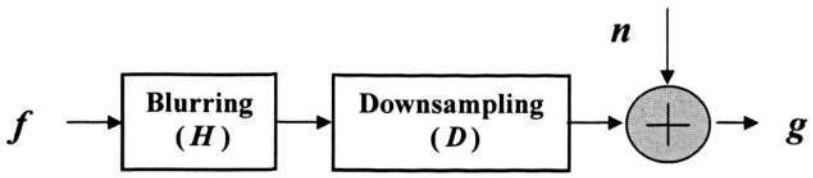


Fig. 1.11: Image degradation process for single-frame SR.

In many imaging applications, of particular interest is to render HR face images from LR ones. This is especially useful in a surveillance system where the resolution of a face image can be low in video. One example has been shown in Fig. 1.8. In this thesis, we focus our attention on single-frame face SR, which aims to estimate the missing facial features from a single LR face image by learning from a large database of face images. However, the face SR has its unique challenges when compared with the SR for generic images, because people are familiar with facial features. A small error in a reconstructed HR face might be unacceptable to human perception, such as an asymmetry of the eyes or noses, whereas the errors in the textured regions of a reconstructed HR generic image usually go unnoticed. Therefore, the single-frame face SR problem is challenging, as there is less observed data available and we have to take into account the specific features of human faces.

1.2.3 Image Deconvolution

The objective of image deconvolution, also known as image restoration, is to restore a blurred

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

and noisy image, without increasing the pixel resolution. The diagram of the degradation process is shown in Fig. 1.12. We can formulate the blurred image as:

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{n} \tag{1.4}$$

From the formulation, it can be seen that image SR and deconvolution are closely related problems. Unlike (1.2) and (1.3), $\boldsymbol{g}$ and $\boldsymbol{f}$ in (1.4) have the same dimension. In image SR, HR images could be reconstructed in three stages, namely (i) image registration from LR images, (2) image fusion and followed by (3) image deconvolution. Thus, image deconvolution can also be considered as one stage of image SR. This thesis focuses on blind color image deconvolution, which is an inverse problem that attempts to estimate the original color image from a single degraded blurred color image at the same pixel resolution, given limited or no prior knowledge of the blurring function. In this case, $\boldsymbol{g}$ and $\boldsymbol{f}$ represent the observed degraded color image and original color image, respectively.

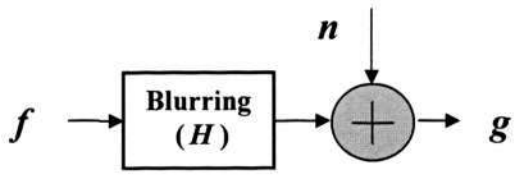


Fig. 1.12: Image degradation process for image deconvolution.

Most current algorithms on blind image deconvolution focus on the restoration of grayscale images. To the best of our knowledge, there is relatively little work addressing blind color image deconvolution. Most common approaches apply the grayscale image deconvolution algorithm to each of the color channels independently. However, these methods neglect the correspondence between the color channels, which should be utilized during the deconvolution process. Therefore, this motivates the study of blind color image deconvolution, which takes

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

into account the special characteristics of color images.

1.3 Contribution

The main contribution of this thesis is the development of a set of HR image reconstruction algorithms, which cover the techniques of image SR and deconvolution. For each topic, we propose the algorithms and verify them using experiments. The main contribution of this thesis can be summarized as follows:

1) Blind Super-resolution Image Reconstruction

We propose a new algorithm to address blind image SR problem, which aims to fuse multiple blurred LR images to render a HR image when there is limited information about the blurring function. We present a new soft maximum *a posteriori* (MAP) estimation framework to perform joint blur identification and HR image reconstruction. The proposed method incorporates a soft blur prior that estimates the relevance of the best-fit parametric blur model, and induces reinforcement learning towards it. It is assumed that the actual blur satisfies up to a certain degree of parametric structure, as there are many well-known parametric blurs in practical applications. Further, it can be tailored flexibly to include other blur types if some prior parametric knowledge of the blur is available.

2) Joint Image Registration and Super-resolution Image Reconstruction

A new algorithm is developed to integrate image registration into multi-frame SR. A critical step in multi-frame SR is accurate registration of the LR images, particularly when the motion is not purely translational between the LR images. In view of this, we present a new framework that performs simultaneous image registration and HR image reconstruction. This is promising as more accurate motion parameters can be determined, thereby enhancing the performance of

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the HR reconstruction. Further, it is worth noting that unlike most algorithms that focus on the translational motion model, the proposed method adopts a more generic motion model that includes both translation as well as rotation. An iterative scheme is developed to solve the arising nonlinear least squares problem.

3) Learning-based Super-resolution Face Image Reconstruction

We present a new learning-based single-frame face SR method, which attempts to estimate missing HR details from a single image by learning from training examples. The SR problem is challenging, as there is less observed data available and we have to take into account the specific features of human faces. A common issue in conventional methods is that, when the given LR face is a new face significantly different with those in the face database, the obtained result is less than satisfactory. To alleviate this difficulty, we develop an improved principle component analysis (PCA)-based method to render a HR face by incorporating an estimated initial HR face in the training set. The reconstructed HR face is able to preserve usual facial features. To further enhance the detailed local information, we develop a patch-based residue prediction approach by using a manifold learning method named locally linear embedding (LLE). Experimental results show that the proposed method is effective in performing single-frame face SR.

4) A hybrid Model Framework for Blind Color Image Deconvolution

We present a new hybrid model framework to address blind color image deconvolution, which aims to restore a single degraded color image at the same pixel resolution, given limited or no prior knowledge of the burring function. Conventional methods based on single-input single-output (SISO) model experience suboptimal results as each color channel is processed independently. On the other hand, there are limitations on the practicality of using multi-input multi-output (MIMO) model in solving this problem as the color channels are usually highly

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

correlated. In view of these constraints, we propose a new framework to solve blind color image deconvolution by first decomposing the color channels into wavelet subbands, and performing image deconvolution using a hybrid of SISO and single-input multi-output (SIMO) models.

1.4 Organization

The remainder of this thesis is organized as follows:

In Chapter 2, we provide the background and literature review in HR image reconstruction. Many popular and recent HR reconstruction methods are presented.

In Chapter 3, we propose a new algorithm to address blind SR image reconstruction. The new method performs joint blur identification and HR image reconstruction from the observed multiple LR images. It develops a soft MAP estimation framework that incorporates likelihood and prior modeling of the HR image and the blur. Experimental results show that the new method is effective in performing blind SR image reconstruction where there is limited information about the blurring function.

In Chapter 4, we present a new algorithm to integrate image registration into multi-frame SR. We adopt a more generic motion model that includes both translation as well as rotation. A new framework using a nonlinear least squares technique is developed to perform simultaneous image registration and HR image reconstruction. Experimental results show that the new method is effective in performing image registration and SR for simulated as well as real-life images.

In Chapter 5, we investigate the single-frame face SR algorithm that estimates missing facial features by learning from a large database of face images. We develop a new framework by employing the PCA technique, and a manifold learning method named LLE. Experimental

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

results show the effectiveness of the proposed method in performing single-frame face SR.

In Chapter 6, we present a new algorithm to address blind color image deconvolution based on a hybrid framework of SISO and SIMO models. We emphasize the importance of exploiting both spatial and spectral correlations in color image deconvolution and use a combination of SISO and SIMO blind algorithms to alleviate the original intractable MIMO problem. Experimental results show that the proposed method is able to provide good restored color images when there is limited information about the blurring function.

Finally, Chapter 7 summarizes all the works in this thesis. Several promising avenues of future research for HR image reconstruction are given.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 2

Literature Review

This chapter provides the background and literature review in HR image reconstruction. We will review many popular and recent HR reconstruction methods. The approaches are classified into three categories: multi-frame SR, single-frame SR and blind image deconvolution.

2.1 Multi-frame Super-resolution Methods

Several articles have surveyed different classical multi-frame SR methods and compared their performances. In 1998, Borman and Stevenson [7] presented a complete overview of different techniques that address the problem of SR video restoration. Later, other review articles [8], [9] and book [10] had followed. In the following, we will review some popular and recent methods, which could be categorized into several categories as shown in Fig. 2.1. On top, the multi-frame SR algorithm can be divided into frequency domain-, spatial domain-, and wavelet domain-based methods. Earlier SR techniques are mostly developed in the frequency domain, but as their degradation model are more limited, later research has tended to concentrate on spatial domain formulation.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

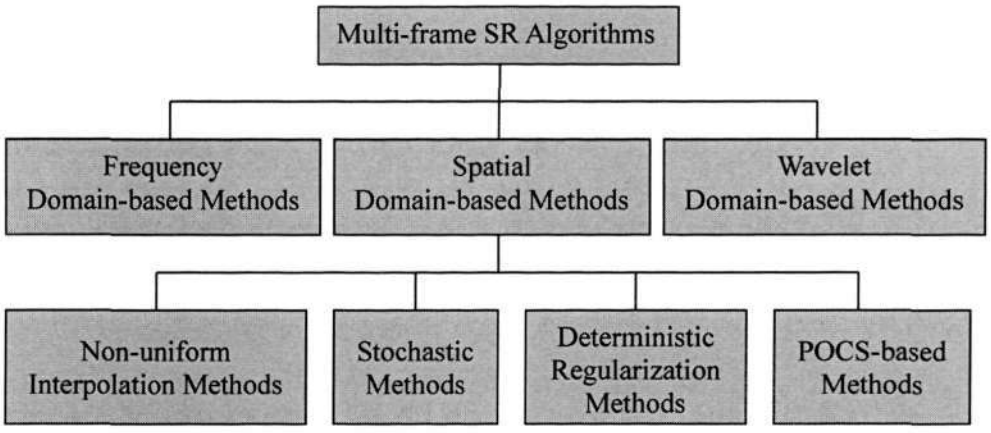


Fig. 2.1: The category of multi-frame SR algorithms.

2.1.1 Frequency Domain-based Multi-frame Super-resolution Methods

The first SR work was proposed by Tsai and Huang in [11], motivated by the need for improved resolution images from Landsat image data. Their method is centered on the shifting and aliasing properties of the continuous and discrete Fourier transform. It formulates a linear equation in (2.1) to estimate the Fourier coefficients of the HR image using the least-squares criterion.

$$\tilde{\boldsymbol{G}} = \boldsymbol{\Phi}\tilde{\boldsymbol{F}} \tag{2.1}$$

where $\tilde{\boldsymbol{G}}$ is a column vector with the elements of the discrete Fourier transform (DFT) coefficients of the LR images, $\tilde{\boldsymbol{F}}$ is a column vector with the samples of the continuous Fourier transform (CFT) of the continuous HR image, and $\boldsymbol{\Phi}$ relates $\tilde{\boldsymbol{G}}$ to $\tilde{\boldsymbol{F}}$. The method is computationally attractive. However, it does not take into account the blurring and noise effects during the image acquisition.

An extension of [11] for blurred and noisy images was proposed by Kim et al. [12]. They developed a weighted recursive least-squares algorithm for solving the linear equations in the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

frequency domain. They further refined their work in [13], [14] by addressing the issue of the ill-posedness of the inverse problem by adopting the Tikhonov regularization method. The SR work is formulated as a minimization problem:

$$\text{Minimize:} \quad \| \tilde{\boldsymbol{G}} - \boldsymbol{\Phi}\tilde{\boldsymbol{F}} \|^2 + R(\tilde{\boldsymbol{F}}) \tag{2.2}$$

where $\|\bullet\|$ denotes the L_2 norm and $R(\tilde{\boldsymbol{F}})$ is the regularization functional. Even though this method stabilizes the solution of the inverse problem, the results may suffer from the errors in the observation model, which arise from motion estimation or blur identification.

To alleviate this difficulty, Bose et al. [15] formulated a new observation model to solve SR problem using the method of total least squares (TLS) [16], which is known to provide robustness to errors in both the system matrix and observations. The minimization problem is transformed into:

$$\begin{aligned} &\text{Minimize:} && \|[\boldsymbol{N}\ \boldsymbol{E}]\|_F \\ &\text{subject to:} && \tilde{\boldsymbol{G}} - \boldsymbol{N} = [\boldsymbol{\Phi} + \boldsymbol{E}]\tilde{\boldsymbol{F}} \end{aligned} \tag{2.3}$$

where $\|\bullet\|_F$ denotes the Frobenius norm. $\boldsymbol{N}$ and $\boldsymbol{E}$ are the errors in the observation $\tilde{\boldsymbol{G}}$ and system matrix $\boldsymbol{\Phi}$, respectively. The recursive TLS algorithm in [17] was adopted to solve (2.3) efficiently. In contrast to the previous frequency domain approaches that use a band-limited image model, recent paper [18] adopted a non-band-limited stationary spectral model. Then the SR problem is considered to be finding a minimum mean squared error (MMSE) reconstruction solution for the generalized under-sampling problem. Although the aforementioned frequency domain-based methods are computationally attractive, the constraint of global translational motion model limits their application. Further, it is also difficult to apply *a prior* knowledge in the frequency domain for regularization.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.1.2 Spatial Domain-based Multi-frame Super-resolution Methods

Having examined a number of frequency domain-based multi-frame SR approaches, we look at some spatial domain-based SR methods in this subsection. Since a more general degradation model can be considered in the spatial domain, most researches in recent literature tend to concentrate on spatial domain techniques.

2.1.2.1 Non-uniform Interpolation Methods

An intuitive spatial domain-based method is the non-uniform interpolation method [1], [19], [20], [21]. The registered LR images can be considered as non-uniformly spaced sample points. These points could be interpolated and resampled on the HR sampling grids. Then deblurring is performed on the interpolated HR image. The advantage of the non-uniform interpolation methods is its low computational cost, however, the methods are suboptimal as they oversimplify the relationship between image deblurring and SR reconstruction. This causes the errors occurred during the reconstruction of the blurred HR image to affect the subsequent HR image deconvolution. Further, these methods require the strict assumption that the blurs are identical for all the observed LR images.

2.1.2.2 Iterative Backprojection (IBP) Methods

To alleviate the difficulty in the non-uniform interpolation methods, some methods assume that the blurring function can be estimated or characterized fully a *prior* before the SR process. In [22], Irani and Peleg proposed an iterative backprojection (IBP) method. It starts with an initial estimate of the SR image, compares the projected LR results with the observed images, and updates the HR estimate based on the predicted error. The scheme to update the HR estimation can be expressed by:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$f^{i+1} = f^i + \sum_{k=1}^{N} H^{BP}(g_k - DH_k S_k f^i) \tag{2.4}$$

where f^i denotes the estimated HR image in the i-th iteration. H^{BP} is a back projection matrix that projects the predicted error $(g_k - DH_k S_k f^i)$ onto the HR estimation. Similar to [22], Shah and Zakhor utilized a Landweber reconstruction method [23] to solve the SR problem [24]. They presented a novel approach for resolving problems caused by unreliable motion estimates in the SR restoration. It can be shown that the IBP method is similar to maximum likelihood (ML) estimation [4]. The advantage of the IBP method is that a more general motion model could be adopted [25]. However, the method is unable to ensure unique solution due to the ill-posed nature of the inverse problem. Further, the choice of H^{BP} is difficult to determine.

2.1.2.3 Stochastic Methods

Tom and Katsaggelos [26], [27] solved the SR problem using three steps: registration of the LR images, restoration of these LR images, followed by an interpolation step that yields the SR reconstruction. The work is built on the ML formulation developed in [28], [29] for multi-channel blur identification and image restoration. A problem with this approach is that performing image restoration on the LR images is not effective because the sampling rate in the formation of LR images is often lower than the Nyquist rate. The observed LR images hence contain significant aliasing artifacts, leading to inadequate blind restoration.

Image SR is a typical ill-posed inverse problem. An effective method for solving this problem is the use of *a prior* knowledge, which constrains the HR image space. It is known that Bayesian estimation methods include *a priori* constraint in the form of prior probability density functions. Bayesian formulation solves the SR problem by maximizing the *a posteriori* (MAP) conditional probability $P(f \mid g_1, g_2, \cdots, g_N)$. To estimate the HR f, the MAP estimator can be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

written,

$$\begin{aligned} \boldsymbol{f} &= \arg\ \max_{\boldsymbol{f}}\{\boldsymbol{P}(\boldsymbol{f} \mid \boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_N)\} \\ &= \arg\ \max_{\boldsymbol{f}}\{\boldsymbol{P}(\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_N \mid \boldsymbol{f})\boldsymbol{P}(\boldsymbol{f})\} \\ &= \arg\ \max_{\boldsymbol{f}}\{\ln\left(\boldsymbol{P}(\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_N \mid \boldsymbol{f})\right) + \ln\left(\boldsymbol{P}(\boldsymbol{f})\right)\} \end{aligned} \tag{2.5}$$

It is observed that (2.5) includes *a priori* constraint in the second term represented by $\boldsymbol{P}(\boldsymbol{f})$, resulting in the stability of the HR estimation being guaranteed. If we drop the second term in (2.5), the problem formulation is equivalent to that of the ML estimation [4].

The distinct feature of the MAP methods is that they will define the prior probability $\boldsymbol{P}(\boldsymbol{f})$, which should reflect the characteristics of the HR image. A general model for the prior distribution can be defined as a Markov random field (MRF) that is characterized by Gibbs distribution [30]. Commonly, for simplicity in the computation, the Gaussian MRF is often used for image prior [31], [32]. In [33], [34] and [35], Tian and Ma exploited a Markov Chain Monte Carlo (MCMC) technique [36] to generate samples from the posterior probability distribution to compute the MAP estimation. To further enhance the edge information in the reconstructed HR image, an edge adaptive MCMC SR approach was further proposed in [37]. Some researchers [38], [39] also modeled the HR image as a simultaneously autoregressive (SAR) model. This simple model requires the estimation of one parameter, which can be computed iteratively using expectation maximization (EM) method. In [40], the image models of both the MRF and SAR were discussed. The paper presented a technique for SR imaging of a scene from several observations at different camera zoom settings. The parameters in the model of the MRF or SAR are learned from the most zoomed observation. Then a picture of the entire scene at a resolution corresponding to the most zoomed observation is obtained. However, either the MRF or SAR prior enforces smoothness uniformly across the image, leading to blurred edges in the reconstructed HR image. To address this problem, Woods and Galatsanos [41] proposed a

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

non-stationary image prior model that provides the flexibility to model spatially varying correlation. In [42] and [43], Schultz and Stevenson adopted a discontinuity-preserving model using the Huber-Markov Gibbs prior, resulting in sharper edges in the reconstructed image.

2.1.2.4 Deterministic Regularization Methods

There is another class of techniques named deterministic regularization approaches, which formulate the SR problem in a way similar to the MAP estimator. The deterministic approach aims to estimate the HR image $\boldsymbol{f}$ by minimizing the following cost function:

$$\boldsymbol{f} = \arg\min_{\boldsymbol{f}} \left\{ \sum_{k=1}^{N} \| \boldsymbol{g}_k - \mathbb{C}_k \boldsymbol{f} \|_p^p + U(\boldsymbol{f}) \right\} \tag{2.6}$$

where $\|\bullet\|_p$ denotes the L_p $(p=1,2)$ norm. The first term in (2.6) represents the data fidelity to the observed data. $U(\boldsymbol{f})$ is the regularization term imposing constraints on the estimated HR image. Smooth constraint known as Tikhonov regularization is often adopted in formulation $\| \boldsymbol{L}\boldsymbol{f} \|_2^2$, where $\boldsymbol{L}$ denotes the high-pass filter. When $p=2$ in the data fidelity term and the Tikhonov regularization is used, it can be shown that the cost function in (2.6) is equal to the MAP estimation in (2.5) where Gaussian process is used to model both the observed noise and image prior. Hardie et al. used this deterministic regularization method for SR reconstruction in [3]. They defined an observation model that incorporates the knowledge of the optical system and the detector array. Both gradient descent and conjugate-gradient (CG) optimization procedures are used to minimize the cost function. Nguyen et al. [44] further addressed the computational issues by using circulant-type preconditioners to accelerate the CG convergence for the Tikhonov regularized SR problem. They adopted the generalized cross-validation (GCV) criterion [45] to calculate the regularization parameters for the underdetermined systems. Later, a very fast SR algorithm for pure translational motion and common space

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

invariant blur was developed in [46]. Farsiu et al. [47] further proposed a fast and robust SR algorithm by using L_1 norm, both for the regularization and the data fidelity terms. It is shown that this method is robust to errors from the estimation of motion or blurs, and results in images with sharp edges.

2.1.2.5 Projection-onto-convex-sets (POCS)-based Methods

Another prominent approach in the spatial domain for SR reconstruction is based on the idea of projection-onto-convex-sets (POCS). POCS-based methods restrict the feasible solution space of SR to the intersection of the different constraints. These constraints include data fidelity to observation, finite energy and statistical characteristics of specific images (faces or texts). We define the constraint set as $\mathbb{Q}_k$ for each constraint. Then the estimated HR image is restricted to $\cap_{k=1}^{m}\mathbb{Q}_k$ $(1 \le k \le m)$ and it can be obtained recursively as follows [48]:

$$\boldsymbol{f}^{i+1} = \boldsymbol{P}_m \boldsymbol{P}_{m-1} \cdots \boldsymbol{P}_1 \boldsymbol{f}^i \tag{2.7}$$

where $\boldsymbol{P}_k$ is the k-th projection operator that projects an image onto the convex set $\mathbb{Q}_k$. Based on this idea, Stark and Oskoui [48] first proposed a POCS-based method in SR. To extend their method, Tekalp et al. considered the observation noise and sensor blurring in [49]. Later Patti et al. [50] further incorporated the motion blur occurring during the aperture time of the camera in the observation model. The problem of space varying blur and arbitrary sampling lattices is also considered. This extension of the observation models tends to suffer from the problem of reconstruction artifacts caused by unreliable motion estimates, since local motion estimates are generally less reliable than global motion estimates. Eren et al. [51] alleviated this difficulty by introducing two new concepts, a validity map and a segmentation map. The algorithm enables an object-based SR reconstruction to be reliable in the presence of registration errors. In [52], Patti and Altunbasak enhanced the previous POCS-based methods by using a high-order

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

interpolation method. They included edge adaptive constraints similar to the MAP estimation to produce sharp edges and reduce edge ringing. Recent paper [53] also aimed to reduce the Gibbs artifacts present on the edges of the HR image obtained by the POCS-based methods. Overall, it can be seen that the POCS-based methods are better than the IPB or ML-based methods, since nonlinear constraints can be integrated into the reconstruction process. Compared with the MAP-based methods, the POCS-based methods are easier to incorporate *a priori* constraints. However, they have the disadvantages of non-uniqueness of solution and slow convergence. To combine the benefits of the MAP or ML estimators, and the advantages of the POCS-based methods, Elad and Feuer [4] proposed a hybrid algorithm to ensure a single optimal solution for the HR image reconstruction.

2.1.2.6 Joint multi-frame SR methods

In most multi-frame SR algorithms, the blurring process is assumed to be either negligible or can be characterized fully *a priori*. In many practical situations, however, it is difficult to eliminate the blurring or to know the blurring function completely. Therefore, it is necessary to incorporate blur identification into the SR algorithms. Currently, there are a few works formulating image SR into a disjoint process of SR reconstruction followed by blind image deconvolution on the blurred HR image [20], [46]. The advantage of the approach is its low complexity and computational cost. Nevertheless, these methods require the assumption that the blurs are identical for all the observed LR images. To the best of our knowledge, there are still very limited works on blind image SR. These include the example-based image learning method [54] and the generalized cross-validation method [55]. However, these methods can only handle images degraded by specific parametric blur models. Recently, He *et al.* proposed a regularization framework for simultaneous estimation of both HR image and blurs [56]. However, the method does not utilize potential parametric blur information, leading to

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

underutilization of the information available.

Another class of SR methods takes registration errors into account. Registation is a critical step in multi-frame SR. In [57], a regularized constrained total least-squares method was proposed to minimize a nonconvex cost function. Similar to [31], the cost function is projected onto the image and motion parameter domains one at a time, and minimized iteratively. However, it has been shown in [58] that the optimization using the alternating minimization (AM) approach tends to be trapped at local minima. Woods et al. [59] proposed algorithms for joint estimation of HR image and motion parameters using Bayesian and MAP formulation. Similar to the idea of the AM framework, EM is employed to solve the problem.

Recently, some image SR methods based on the principle of variable projection (VP) have been developed in [60] and [61] . In [60], Chung et al. proposed a nonlinear cost function, and estimate the registration parameters and the HR image using the Gauss-Newton method. The method, up to a certain extent, shares many similarities with the AM method, as the cost function is projected onto the image and motion parameter domains iteratively. In [61], a VP-based SR method was also developed by Robinson et al. In their method, the registration parameters for general motion model are first estimated by minimizing a VP functional. Then the estimates of the motion parameters are used to reconstruct the HR image using ML or MAP estimation. The paper further describes an efficient implementation of the minimization process in the Fourier domain. In [62] and [63], two subspace methods were developed for registration of totally aliased signals. After the motion parameters are estimated, the HR image is reconstructed in the frequency domain by using the least squares method. These methods, nevertheless, do not consider the effect of uniform kernel during the downsampling process of LR image formation. Some other researchers [64], [65] attempted to model the registration errors using Gaussian noise and proceed to develop a regularized adaptive HR estimation

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

method. However, these methods can only handle global translational shift. This precondition constrains the application of these methods.

2.1.3 Wavelet Domain-based Multi-frame Super-resolution Methods

Another class of the SR methods is centered on wavelet-based technique. Nguyen and Milanfar first proposed an efficient wavelet-based SR reconstruction algorithm in [66]. They exploited the interlacing structure of the sampling grid in SR and derive a computationally efficient wavelet interpolation for interlaced two-dimensional data. Bose *et al* in [67] presented an image-formation system composed of a set of identical CCD sensor arrays which are shifted with respect to each other by subpixel displacement. As the assumption that perfect subpixel displacements can be achieved is impractical, they considered that blur functions in the multisensor HR image reconstruction are periodically shift-variant and are formed by sampling and summing different shift-invariant blurring operators. Based on the image system in [67], Zuowen Shen *et al* [68] extended the algorithm in [69] to propose a wavelet deblurring algorithm for spatially variant blurs for HR image reconstruction. They represented the different shift-invariant blurring operators by tensor products of different low-pass filters and designed a dual filter for each low-pass filter associated with the corresponding blurring operator. In [70], they further developed a minimally supported biorthogonal wavelet system for arbitrary resolution ratio. Experimental results showed that the new wavelet-based algorithm performs better than the one designed in [68] with fewer artifacts.

Recently, it has been shown that second-generation wavelets (SGWs) are inherently more suited for image SR than first-generation wavelets (FGWs) [66]. SGWs replace dilations and translations with an entirely spatial domain lifting scheme based on the operations of splitting, prediction, and updating. In [71], SGWs were firstly introduced to solve image SR by Bose et al. HR image reconstruction with noise reduction is restricted to semi-regular sampling lattices

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(tensor product of 1D irregular sampling lattices). Consequently, the LR images are displaced only by translations from a reference LR image. Intead of only using translational motion model, 2D prediction and update operators are used to handle arbitrary sampling lattices in [72]. To extend their work [72], Chappalli et al. investigated the effect of the threshold level on the reconstructed image quality in SGW [73]. The selection of the optimal threshold involves a tradeoff between the impact of blur and noise.

2.2 Single-frame Super-resolution Methods

Most of SR literatures focus on multi-frame SR. With only one image, it is generally difficult to estimate the missing pixels with little prior information about the image. Single image resolution enhancement by estimating the missing pixel information from the surrounding pixels of the observed LR image is often referred to as *image interpolation*. Although this field has been extensively studied [74], [75], [76], the quality of the magnified image is inherently limited as there is no new information that can be used for HR reconstruction. The higher the scaling factor, the smoother the magnified image seems to be. In contrast to the existing interpolation methods, the learning-based single-frame SR methods use the learned knowledge from training data to increase image resolution. Generally, the LR images can be classified as generic images and domain-specific images, e.g. face or text images. According to this classification, single-frame SR methods can be categorized into two groups as shown in Fig. 2.2.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

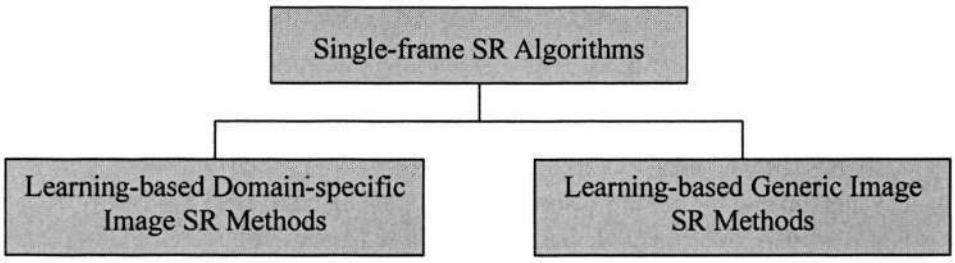


Fig. 2.2: The category of single-frame SR algorithms.

2.2.1 Learning-based Generic Image Super-resolution Methods

Learning-based generic image SR methodology has been drawing research attention in recent years. It aims to increase the resolution of single generic (natural) image by use of a database of HR natural images as training samples. The most popular methods are based on the idea of *pattern matching* technique [2], [77], [78]. The training database consists of a set of HR images and a corresponding set of LR versions obtained by applying degradation matrix on each HR image and adding noise. The basic idea is that given a LR pattern, e.g. small patches of LR images, seek similar LR examples in the database. The corresponding HR components could then be used for the HR reconstruction as they provide the high-quality content that fits the observation. Two learning-based SR methods were developed in [2]. The first employs a Markov network [77] to infer high frequency components. The second approach is a one-pass algorithm by finding the nearest neighbor of the medium frequency component in the training set and transferring the corresponding high frequency component to the interpolated LR image. In [79], Hertzmann et al. also presented a learning-based method for performing SR, in the context of analogies between images. Based on the idea in [77], Sun et al. [80] only replaced the patches of detected primal sketches, and then applyed the IBP method [22] to ensure reconstruction constraints. Inspired by recent locally linear embedding (LLE) manifold

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

learning method [81], Chang et al. in [82] assumed that the feature spaces for LR and HR face images share similar local manifold structure. Then the SR method through neighbor embedding was proposed. Extending their work, Chang et al. [83] formulated the single-frame SR problem as a direct energy minimization problem by using a three-layer Markov network to realize the manifold assumption.

Instead of encoding a fixed relationship between pairs of HR and LR image patches as described above, an entirely different way of exploiting training examples is to use them for training the parameters in image priors. The approaches introduce the training samples into inverse problems and exploit the examples to tune some parameters that control HR image priors. In [84], a sparse derivative prior was learned in the training database. The max-product belief propagation used in [77] is adopted to compute the MAP estimate. Roth and Black [85] developed a method for learning rich MRF image priors by exploiting ideas from sparse image coding. They modeled the prior probability of an image in terms of a random field with overlapping cliques, whose potentials are represented as a Product of Experts [86]. In contrast to the previous methods that learn the prior based on a set of HR images, the work in [87] uses all pairs of training images, the HR images and the corresponding degraded ones. The method enables the MAP estimation on all the LR training images and aims to minimize the reconstruction errors. Then the parameters controlling the HR image prior can be obtained, leading to the MMSE.

Another alternative approach to single-frame SR is based on the unsupervised learning technique. The algorithm in [88] relies on a generic model of HR images that is obtained from kernel principal component analysis (PCA), which is capable of capturing part of the higher-order statistics of images. In [89], Jiji et al. proposed a single-frame SR algorithm using the contourlet transform. The contourlet coefficients at finer scales of the unknown HR image

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

are learned locally from a set of HR training images. Extending their work, the paper in [90] further studies the usefulness of different local and global learning-based single-frame SR reconstruction techniques in handling three specific tasks, namely, deblurring, denoising and anti-aliasing.

2.2.2 Learning-based Domain-specific Image Super-resolution Methods

It is noted that the use of training examples becomes much more effective when handling domain-specific images, such as text or face images. Based on the idea in [91], Dalley et al. [92] proposed a single-frame text SR method by adopting a full-Bayesian approach with an explicit noise model. In [93], a novel algorithm for text SR was developed by interpolation with a variable linear filter. The filter is modeled as a linear function of edge patterns, whose parameters can be trained in a text database. Pickup et al. [94] presented a texture image prior in the form of a distribution function based on texture samples. In [95], an efficient SR scheme for scanned documents was proposed. The algorithm starts by assigning each location in the LR image several HR patches. These are found as the nearest neighbors in the training set. Then the patches are used for the definition of an image prior expression, and later merged into a global MAP penalty function.

Baker and Kanade first proposed the term "*hallucinating faces*" for learning-based face image SR [96]. They proposed an algorithm to learn *a prior* on the spatial distribution of face image gradients. However, the method does not consider some global properties of a face, giving rise to suboptimal result, such as an asymmetry of the eyes. To overcome this shortcoming, a two-step approach to hallucinating faces was developed by using both global parametric and local nonparametric models in [97]. However, the degradation model in this method has to be determined explicitly, which is impractical in many applications. To alleviate this difficulty, Wang and Tang in [98] developed a face hallucination method by employing

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

eigentransformation. They used a linear combination of the HR images in the training set to render the HR image, while retaining the same weights as those in the LR image space. Although this algorithm has low complexity, it loses some detailed facial information. The use of PCA techniques for image zooming purposes has been investigated in [99]. It is assumed that the principal components remain unchanged across the scale. The method is applicable only to zooming up of images of a specific class of objects such as faces or fingerprints. A different method for enhancing the resolution of LR facial images using an error back projection method based on top-down learning was proposed in [100]. The approach represents a face by a linear combination of prototypes of shape and texture. Inspired by the generic image SR algorithm in [88] to further incorporate the knowledge of the imaging process, Chakrabarti et al. [101] used the kernel PCA for deriving prior knowledge about the face class. Valuable prior information is extracted in a computationally efficient manner and used with a MAP framework for improving the quality of the SR face. Recently, Stephenson and Chen [102] have proposed an adaptive MRF method for learning-based face SR. They adopted adaptive observation and transition functions, which are region-dependent. In [103], Liu et al. developed an associative learning method for learning dependencies between LR face data and HR data in the hidden parameter space rather that the observation space. Although the aforementioned learning-based algorithms are able to provide good visual results, they are usually not able to handle the real-time applications, since its performance depends on the training database that is usually huge, leading to larger computational cost.

2.3 Blind Image Deconvolution Methods

There is a large amount of work in the literature on blind image deconvolution. In the following subsection, we will review the most popular methods in three categories as shown in Fig. 2.3.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Most current algorithms on blind image deconvolution focus on image restoration of grayscale images. Some good reviews have been provided in [5], [104] and [105].

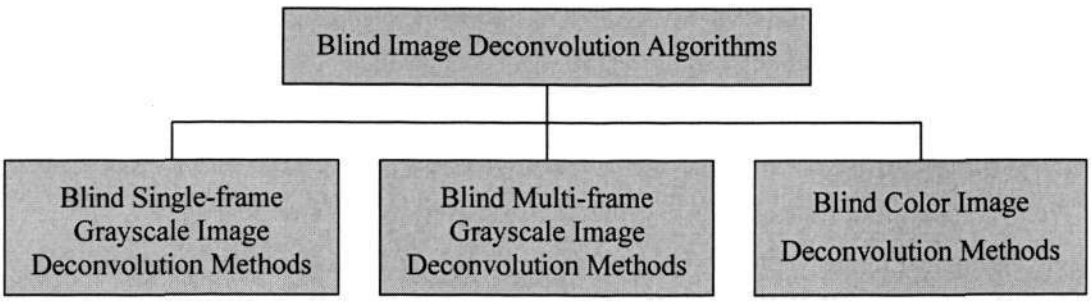


Fig. 2.3: The category of image deconvolution algorithms.

2.3.1 Blind Single-frame Grayscale Image Deconvolution Methods

Blind single-frame grayscale image deconvolution attempts to estimate the original grayscale image from a single observation of blurred grayscale image, given limited or no prior knowledge of the blurring function. To address this problem, some parametric methods were developed by using the autoregressive moving average (ARMA) parameter estimation techniques [106], [107], [108], [109]. These methods model the original image and blur as autoregressive (AR) and moving average (MA) processes, respectively. Based on this model, the objective of image deconvolution is to estimate the ARMA parameters. In [106], [107] and [108], the ML estimation technique was used to derive the restoration filters by estimating the blurs, variance of the noise, and the AR model coefficients of the original image. EM techniques are employed to optimize the cost function. Reeves et al. further proposed an alternative ARMA estimation algorithm based on the technique of GCV [109]. The GCV method could provide better results and has higher computational complexity when compared with the ML method. The aforementioned parametric methods are less sensitive to the noise.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

However, the performance of the methods is susceptible to high-dimensional parametric vector, leading to poor local convergence. Moreover, the AR process may also be inadequate in modeling real-life images that consist of numerous non-homogeneous regions.

In contrast to the parametric methods, some researchers have developed nonparametric deterministic image constraints techniques to address blind image deconvolution. Ayers and Dainty proposed an iterative blind deconvolution (IBD) method by using the Wiener-based filter in frequency domain to restore the image and identify the blur alternately [110]. The method is popular for its computational complexity. However, the major shortcoming of IBD is its sensitivity towards the initial image estimate, thus leading to algorithmic instability. The uniqueness and convergence properties are uncertain. Another class of iterative methods include nonnegative and support constraints recursive inverse filtering (NAS-RIF) algorithms [111] [112]. The NAS-RIF methods involve recursive filtering of the blurred image to minimize a predictive cost function. However, the main drawback of the aforementioned iterative methods is that they require the image object to have a known support against a uniform background. This is inflexible as it conflicts with many real-life images.

Recently, inspired the work of learning-based single-frame SR works [77], [113], a vector quantization (VQ)-based image restoration algorithm was proposed [114]. A number of codebooks are designed corresponding to various versions of the blurring functions. Given a noisy and blurred image, one of the codebooks is chosen based on a similarity measure, therefore providing the identification of the blur. In [115] and [116], the Bayesian formulation was used for image deconvolution where the point spread function (PSF) was assumed partially known. Extending their works, a new algorithm based on a variational approximation was developed in [117]. The method has the benefits of the Bayesian model while bypassing some of its difficulties.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.3.2 Blind Multi-frame Grayscale Image Deconvolution Methods

Unlike single-frame deconvolution, blind multi-frame grayscale image deconvolution attempts to estimate the original grayscale image from multiple blurred grayscale images of the same imaging scene. Using multiple measurements, the restoration algorithm can exploit the redundancy present in the observations. In principle it can achieve performance not obtainable from a single measure. One of the earliest intrinsic multi-frame deconvolution [118] was designed particularly for images blurred by atmospheric turbulence. Extending the work in [119], Pai and Bovik [120] proposed a restoration algorithm based on a subspace approach. In [121], a method based on the greatest common divisor was proposed. It has been shown that the algorithm is computationally efficient and relatively robust to noise.

Some efforts focus on blur identification that reduces the blind problem to the nonblind restoration. Harikumar et al. proposed an elegant one-step eigenvector-based algorithm (EVAM) which accomplishes image restoration and blur identification by finding a minimum eigenvector of a multi-channel (MC) condition matrix [122]. Giannakis et al. [123] developed another multi-frame deconvolution method based on Bezout's identity of coprime polynomials, which finds restoration filters and restores the original image by convolving the filters with the observed images. Unfortunately, both the EVAM method and Giannakis's method are susceptible to the noise and even for a moderate noise level, the restoration may break down. To overcome this disadvantage, Sroubek and Flusser proposed an algorithm that combines the TV technique with the EVAM in order to offer a solution that is more robust toward noise [124]. More recently, a generalization of the Bussgang blind equalization algorithm to the iterative MC deconvolution was developed in [125]. A Bayesian nonlinear estimator, acting in the spatial domain, has been derived for the case of uncorrelated images, while for highly correlated natural images a different approach to the nonlinearity design, based on the representation of

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the image in a properly defined edge domain, has been described. All the above-mentioned multi-frame deconvolution methods contain a common assumption. They require all the PSFs in each blurred image to be coprime, which sometimes may be untrue in practice.

2.3.3 Blind Color Image Deconvolution Methods

Most current algorithms on blind image deconvolution focus on image restoration of grayscale images. However, there has been little investigation of blind color image deconvolution. We first review some classical color image restoration methods in the past. These methods assume that the blurs or PSFs are known prior to image restoration.

Hunt and Kubler derived a MMSE multi-channel filter by introducing Karhunen-Loeve (KL) transform in decorrelating the color channels, so that each channel can be processed independently. This method, however, requires the assumption that the MC correlation describing the interchannel and intrachannel relationship is separable [126]. The use of this assumption is restrictive, and the method cannot handle the case when interchannel degradation is present. Other researchers [127] had proposed techniques to alleviate this separability condition by utilizing the information of interchannel correlation. Set theoretic and constrained least squares restoration had been studied in [128]. Kalman filtering was employed in [129] to perform color image restoration. A regularized framework of color image restoration using cross-validation had been proposed in [130] by extending the work for grayscale images in [131], while a MC restoration technique in frequency domain had been developed in [132]. Schultz and Stevenson developed a MAP-based color image restoration method [133]. A Gibbs prior over a MRF containing spatial and spectral clique functions has been selected as the image model. By extending the work in [127], Angelopoulos and Pitas further discussed the issues for the design of MC Wiener filters both in the spatial and frequency domain [134]. However, these techniques assume that the intra- and interchannel blurs or PSFs are known prior to image

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

restoration. This precondition, unfortunately, imposes rigid constraints on the practicality of these methods. Currently the amount of works on blind color image restoration is still very limited. Chow et al. [135] extended the framework of grayscale image deconvolution to blind color image restoration. A three-dimensional (3D) Laplacian operator is employed to regularize the color channels of the image. An iterative EM algorithm has been extended to the multi-channel and has been used for intra- and interchannel blur identification and image restoration [28], [29]. This approach assumes Gaussian stochastic models for both the image and the noise statistics. However, the assumption may not be realistic, as the Gaussian distribution cannot characterize a vast majority of images, whose histograms may not even be unimodal.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 3

Blind Image Super-resolution

3.1 Introduction

The objective of multi-frame SR is to fuse a set of aliased, blurred and noisy LR images so as to produce a HR image. This chapter mainly investigates blind image SR that is a process to perform multi-frame SR with limited or no knowledge of the blurring occurred during the image formation process. The observed LR images may be blurred due to many reasons, such as motion blurring due to running objects, or optical blurring since the sensor plane may not be in focus. Conventional SR image reconstruction algorithms assume the blurring to be either negligible [11], [20] or can be characterized fully *a priori* [4], [22], [44]. This assumption, however, is impractical as it is often difficult to eliminate the blurring completely in some applications or to know the blurring function completely *a priori*. Therefore, this motivates the study of blind image SR. It attempts to incorporate blur identification into the SR algorithms.

Most multi-frame SR algorithms either assume the blurring functions are known *a priori* or can be estimated reliably using blind image deconvolution on the LR images prior to SR image reconstruction. Often it is impractical to have full knowledge of the blurring function in some applications such as remote sensing. Further, performing blind image deconvolution on the LR images may not be effective because the sampling rate in the formation of LR images is often lower than the Nyquist rate. Thus, the observed LR images hence contain significant aliasing artifacts, leading to inadequate blind deconvolution.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

In view of this, we propose a new soft maximum *a posteriori* (MAP) framework to address blind image SR. The main contribution of this chapter is that we introduce and incorporate a new soft parametric blur prior to solve the SR problem. As opposed to the conventional notion of hard decision [54], [55], the new method proposes a soft blur learning scheme that attempts to integrate the parametric information of blurring function into the algorithm. The best-fit parametric model of the blurring function is estimated and incorporated into the blur prior to induce reinforcement learning towards it. This chapter also derives a total variation (TV) prior in the vector-matrix form for image regularization, which is an extension of the fixed point (FP) scheme [136]. The vector-matrix form of the TV prior simplifies the minimization of the quadratic cost function in the MAP estimation. A MAP framework based on alternating minimization (AM) is then developed to estimate the blur and the HR image iteratively. Conjugate gradient (CG) optimization is used to improve the computational efficiency of the algorithm. In contrast to the SR methods [20], [46], the new algorithm offers a joint process of blur identification and HR image reconstruction. This provides a unified framework where a more reliable blur identification from the HR image can be achieved, which in turn renders a better HR image reconstruction. Experimental results in later sections show that the new method is effective in performing blind image SR where there is limited information about the blurring functions.

The rest of this chapter is organized as follows. The problem formulation of blind image SR is introduced in Section 3.2. The soft MAP framework for blind SR reconstruction is presented in Section 3.3. Experimental results on simulated and real-life images are presented and discussed in Section 3.4. A brief summary is given in Section 3.5.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.2 Problem Formulation

Let us start with the standard approach similar to the papers [46], [47] to model the acquired k-th LR image $g_k(m,n)$ of sized $M_g \times N_g$ $(1 \le k \le N)$ by shifting the HR image $f(x,y)$ of sized $M_f \times N_f$ with a translational motion vector $\boldsymbol{s} = (s_{xk}, s_{yk})$, blurring by point spread functions (PSFs) h_k and h_c, downsampling the result to the resolution of the observed image by a decimation factor of γ, followed by additive noise n_k. Similar to most work on SR [8], [46], it is assumed in this chapter that the imaging blur is spatial shift-invariant and the warping between different LR images is translational. For tractability, the process can be linearly expressed as:

$$g_k(m,n) = \left(f(x+s_{xk}, y+s_{yk}) \otimes h_k \otimes h_c\right)\downarrow + n_k(m,n) \tag{3.1}$$

where $\otimes$ is the two dimensional convolution operator representing the blurring process, $\downarrow$ is the decimation operator, and n_k denotes the additive white Gaussian noise (AWGN). h_k represents the external blurring such as motion blurring in each LR image and h_c represents the effect of spatial integration of light intensity over a square surface region to simulate image acquisition by the sensors during the downsampling process. h_c takes the form of a uniform PSF with support $(\gamma \times \gamma)$, where γ is dependent on the desired HR. Here we consider the decimation factors in the horizontal and vertical directions to be the same. Then the decimation factor γ can be represented by $\gamma = \frac{M_f}{M_g} = \frac{N_f}{N_g}$.

In this work, it is considered that there is limited information on h_k. The objective of blind image SR is to reconstruct the HR image $f(x,y)$ from N observed LR images with little or no information about the blurring functions. Using the commutativity property of block circulant matrices and convolution, the SR formulation in (3.1) can be expressed in the vector-matrix

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

form as:

$$\boldsymbol{g} = \boldsymbol{W}\boldsymbol{f} + \boldsymbol{n} = \boldsymbol{U}\boldsymbol{h} + \boldsymbol{n} \tag{3.2}$$

where $\boldsymbol{f}$ denotes the lexicographically ordered original image f, $\boldsymbol{g} = \left[\boldsymbol{g}_1^T \cdots \boldsymbol{g}_N^T\right]^T$, $\boldsymbol{h} = \left[\boldsymbol{h}_1^T \cdots \boldsymbol{h}_N^T\right]^T$ and $\boldsymbol{n} = \left[\boldsymbol{n}_1^T \cdots \boldsymbol{n}_N^T\right]^T$ are the vectors representing the discrete, concatenated and lexicographically ordered g_k, h_k and n_k, respectively. The matrices of $\boldsymbol{W}$ and $\boldsymbol{U}$ are given as:

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_1 \\ \vdots \\ \boldsymbol{W}_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}\boldsymbol{H}_c\boldsymbol{H}_1\boldsymbol{S}_1 \\ \vdots \\ \boldsymbol{D}\boldsymbol{H}_c\boldsymbol{H}_N\boldsymbol{S}_N \end{bmatrix}, \quad \boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_1 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{U}_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}\boldsymbol{H}_c\boldsymbol{S}_1\boldsymbol{F} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{D}\boldsymbol{H}_c\boldsymbol{S}_N\boldsymbol{F} \end{bmatrix}$$

where $\boldsymbol{D}$ denotes the decimation operator which is considered to be the same for all LR images. $\boldsymbol{S}_k$ is a block circulant matrix that represents the translational warping of the HR image f for the k-th LR image. $\boldsymbol{H}_c, \boldsymbol{H}_k$ and $\boldsymbol{F}$ are the corresponding matrices constructed from h_c, h_k and f, respectively [47], [124], [137]. The matrices $\boldsymbol{H}_c$ and $\boldsymbol{H}_k$ are all block circulant as well.

3.3 A New MAP Framework to Blind Super-resolution

In this work, we propose a new MAP-based framework to address stochastic blind image SR reconstruction. We assume that $\boldsymbol{f}$, $\boldsymbol{h}$ and $\boldsymbol{g}$ are random vectors with associated probability density functions (pdfs) of $p(\boldsymbol{f})$, $p(\boldsymbol{h})$ and $p(\boldsymbol{g})$, respectively. To estimate the HR image and the blurs, we will determine the arguments that maximize the *a posteriori* probability $p(\boldsymbol{f}, \boldsymbol{h} | \boldsymbol{g})$. Appling the Bayesian theorem, we have

$$p(\boldsymbol{f}, \boldsymbol{h} | \boldsymbol{g}) = \frac{p(\boldsymbol{g} | \boldsymbol{f}, \boldsymbol{h}) p(\boldsymbol{f}) p(\boldsymbol{h})}{p(\boldsymbol{g})} \tag{3.3}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

We will ignore $p(\boldsymbol{g})$ in the estimation of $\boldsymbol{f}$ and $\boldsymbol{h}$ since it is a common term that does not affect the maximization of $p(\boldsymbol{f},\boldsymbol{h}|\boldsymbol{g})$. The objective is to model the likelihood $p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{h})$ and the priors $p(\boldsymbol{f})$ and $p(\boldsymbol{h})$ so as to enable the MAP estimation. Assume that the noise n_k is AWGN with variance ${\sigma_k}^2$, the likelihood $p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{h})$ can be expressed as:

$$p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{h}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{g}-\boldsymbol{W}\boldsymbol{f})^T \boldsymbol{C}^{-1}(\boldsymbol{g}-\boldsymbol{W}\boldsymbol{f})\right\} \tag{3.4}$$

As each LR observation is independent, the covariance matrix $\boldsymbol{C}$ for the noise $\boldsymbol{n}$ can be expressed by $\boldsymbol{C}=\begin{bmatrix} \boldsymbol{C}_1 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{C}_N \end{bmatrix}$, where $\boldsymbol{C}_k$ denotes the covariance matrix of the noise $\boldsymbol{n}_k$ for the k-th LR image. Thus, $\boldsymbol{C}_k$ can be represented by $\sigma_k^2\boldsymbol{I}$. For simplicity, the same noise variance $\sigma^2 = {\sigma_k}^2$ is assumed for all channels [32], [46]. Thus, $\boldsymbol{C}$ can be represented as $\sigma^2\boldsymbol{I}$. In the following subsections, we will discuss the formulation of appropriate priors $p(\boldsymbol{f})$ and $p(\boldsymbol{h})$.

3.3.1 High-resolution Image Prior

The *prior* of the HR image can be chosen based on whether it is domain-specific or general-purpose. Domain-specific applications refer to example-based super-resolution [96], which uses the characteristics of specific domain (e.g. faces or text) to perform learning. On the other hand, it is difficult to construct learning samples for general-purpose applications. In this chapter, we will consider the general-purpose SR algorithms. A classical HR image prior is given as $p(\boldsymbol{f}) \propto \exp\left\{-\int |\nabla f|^2\right\}$, where ∇f denotes the gradient of f [138]. The prior essentially imposes smoothness constraint on the images, since the condition is satisfied in many real-life images. However, this *prior* tends to over-penalize the gradients at the edges,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

giving rise to insufficient recovery of visual clarity at these regions. Therefore, the TV technique [136] is introduced here to overcome this shortcoming.

We employ the TV norm $\boldsymbol{T}_1 = \int |\nabla f|$ instead of the $\boldsymbol{L}_2$ norm $\int |\nabla f|^2$. By using the discretization of the image gradient ∇f, an approximation of the TV norm can be formulated as:

$$\begin{aligned} \boldsymbol{T}_1 &= \sum_x \sum_y |f(x+1,y) - f(x,y)| + |f(x,y+1) - f(x,y)| \\ &= \| \boldsymbol{U}\boldsymbol{f} \|_1 + \| \boldsymbol{V}\boldsymbol{f} \|_1 \end{aligned} \tag{3.5}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the first order derivative operators in the x and y directions, respectively. It can be shown that we need to solve the nonlinear partial differential equations (PDEs) in the MAP estimation [136]. Similar to the work in [124], we reformulate the TV norm as below to circumvent the difficulty of solving nonlinear PDEs.

$$\begin{aligned} \boldsymbol{T}_1 &= \sum_x \sum_y \frac{1}{\varepsilon(\underset{\sim}{x}^i)} (f^{i+1}(x+1,y) - f^{i+1}(x,y))^2 + \frac{1}{\varepsilon(\underset{\sim}{y}^i)} (f^{i+1}(x,y+1) - f^{i+1}(x,y))^2 \\ &= \boldsymbol{f}^T \boldsymbol{T}(\varepsilon) \boldsymbol{f} \qquad \left(\boldsymbol{T}(\varepsilon) = \boldsymbol{T}_x(\varepsilon) + \boldsymbol{T}_y(\varepsilon)\right) \end{aligned} \tag{3.6}$$

where $\varepsilon^i(\underset{\sim}{x})$ and $\varepsilon^i(\underset{\sim}{y})$ are defined as $\varepsilon(\underset{\sim}{x}^i) = \sqrt{(f^i(x+1,y) - f^i(x,y))^2 + \eta}$ and $\varepsilon(\underset{\sim}{y}^i) = \sqrt{(f^i(x,y+1) - f^i(x,y))^2 + \eta}$, respectively. Here '$i$' is the i-th iteration of the AM loop, which will be explained in Section 3.3.3. $\eta > 0$ is the regularization constant [139], which is required in the smooth regions where the absolute values of the gradients in the x and y directions are zero or very close to zero. The construction of $\boldsymbol{T}(\varepsilon)$ is outlined in Appendix A of this Chapter. It is noted that the values of $\varepsilon^i(\underset{\sim}{x})$ and $\varepsilon^i(\underset{\sim}{y})$ are calculated by using the estimated HR image f^i in the previous iteration of the AM loop and fixed for the estimation of f^{i+1}. Further, the gradient of the image could also include two additional diagonal directions apart

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

from the x and y directions.

The *prior* pdf of the HR image can be obtained as follows:

$$\begin{aligned} p(\boldsymbol{f}) &\propto \exp\left\{-\frac{1}{2}\int |\nabla f|\right\} \\ &\propto \exp\left\{-\frac{1}{2}\boldsymbol{f}^T\boldsymbol{T}(\varepsilon)\boldsymbol{f}\right\} \end{aligned} \tag{3.7}$$

In this work, we derive the TV prior in the vector-matrix form as an image prior, which simplifies the minimization of the quadratic cost function in the MAP estimation.

3.3.2 New Soft Blur Prior

In this section, we will propose a new reinforcement blur estimation framework. As opposed to the conventional notion of hard decision [54], [55], the new method proposes a soft blur learning scheme that attempts to integrate the parametric information of the blurring function into the algorithm. The best-fit parametric model of the blurring function is estimated and incorporated into the blur prior to induce reinforcement learning towards it.

We first define a blur function set comprised of different parametric blur types $b_i(\theta)$ with their defining parameters θ, where i is used to index different parametric blur models. In this work, we will focus on some commonly observed blur models, namely the Gaussian, the out-of-focus and the linear blurs. The learning set Θ is defined as follows:

$$\Theta = \left\{ \begin{array}{l} \text{Out-of-focus blurs: } b_1(r) = \dfrac{1}{\pi r^2} \quad (if\ \sqrt{x^2+y^2} \le r) \\ \text{Gaussian blurs: } b_2(\sigma) = \exp(-(x^2+y^2)/(2\sigma^2)) \\ \text{Linear blurs: } b_3(a,b) = b - a\sqrt{x^2+y^2} \end{array} \right\} \tag{3.8}$$

where a, b, r and σ are model parameters of blurs. It is worth noting that the construction of Θ

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

is flexible as blur structures can be added, removed or modified based on the prior parametric information of the blurring functions.

The idea behind the framework is that it is well known that most real-life blur h_k satisfy, up to a certain degree of, parametric structures in most practical applications. Therefore, we will try to find the best-fit parametric blur model $b_j(\theta_k)$ with respect to the current blur estimate h_k, and induce reinforcement learning towards it. This will integrate the parametric information progressively throughout the HR image reconstruction. The estimation of $b_j(\theta_k)$ is outlined as follows:

1. Find the best-fit parametric arguments θ_k corresponding to each blur type $b_i(\theta)$ for each estimated channel blur ($\hat{h}_k$):

For *all k* **do**

For *all i* **do**

$$\theta_k = \arg\min_{\theta}(\| \hat{h}_k - b_i(\theta) \|) \tag{3.9}$$

end For

end For

2. Determine the best-fit blur model $b_j(\theta_k)$ for each estimated channel blur ($\hat{h}_k$):

For *all k* **do**

For *all i* **do**

$$j = \arg\min_{i}(\| \hat{h}_k - b_i(\theta_k) \|) \tag{3.10}$$

end For

end For

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The distance function $\|\bullet\|$ is the L_2 norm and '^' denotes the currently computed estimate. The minimization problem in (3.9) can be solved using the *fminbnd* function in MATLAB Optimization Toolbox. The algorithm is based on golden section search and parabolic interpolation in [140]. The best-fit blur model $b_j(\theta_k)$ is determined by searching for the minimum of the Euclidean distance between each $b_i(\theta_k)$ with the estimated $\hat{h}_k$ in (3.10). Assuming the errors between the actual h_k and the estimated $b_j(\theta_k)$ are additive, white and Gaussian, the proposed blur *prior* is given by:

$$p(\boldsymbol{h}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{h}-\boldsymbol{b})^T(\boldsymbol{h}-\boldsymbol{b})\right\} \tag{3.11}$$

where $\boldsymbol{b}=[\boldsymbol{b}_1 \cdots \boldsymbol{b}_k \cdots \boldsymbol{b}_N]^T$ and $\boldsymbol{b}_k$ is the parametric blur vector formed by $b_j(\theta_k)$. The new blur prior is a reinforcement learning term that induces learning toward the best-fit soft parametric model. The new framework can learn the relevance of current estimated blurs with respect to the parametric models, and integrates this knowledge progressively into the scheme.

3.3.3 Alternating Minimization for MAP Estimation

Expressing the *a posteriori* $p(\boldsymbol{f},\boldsymbol{h}|\boldsymbol{g})$ using (3.3), (3.4), (3.7), (3.11), and taking the logarithm, the MAP estimation of the HR image and the blurs can be formulated as the minimization of the following cost function:

$$A(\boldsymbol{f},\boldsymbol{h}) = \frac{1}{2}\left\{(\boldsymbol{g}-\boldsymbol{W}\boldsymbol{f})^T(\boldsymbol{g}-\boldsymbol{W}\boldsymbol{f}) + \lambda \boldsymbol{f}^T\boldsymbol{T}(\varepsilon)\boldsymbol{f} + \beta(\boldsymbol{h}-\boldsymbol{b})^T(\boldsymbol{h}-\boldsymbol{b})\right\} \tag{3.12}$$

where λ and β are the regularization parameters. It is noted that $A(\boldsymbol{f},\boldsymbol{h})$, as a joint function of several variables, is not convex. However, if we project $A(\boldsymbol{f},\boldsymbol{h})$ into $\boldsymbol{f}$ or $\boldsymbol{h}$ iteratively, $A(\boldsymbol{f},\boldsymbol{h})$ becomes convex with respect to $\boldsymbol{f}$ or $\boldsymbol{h}$. This implies that the minima of the projected

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

cost function in each domain can be reached. Therefore, an AM procedure is developed to search for the minimum of the function. Even though in general the proposed method may experience local convergence, however, as we have demonstrated in various experiments in Section 3.4, the initial estimates are sufficient to produce satisfactory convergence and results in the proposed algorithm. In order to explain the algorithm more effectively, we use the equivalent form of $\boldsymbol{Uh}$ instead of $\boldsymbol{Wf}$ in (3.2) when projecting $A(\boldsymbol{f},\boldsymbol{h})$ onto the blur domain. The solution of the minimization problem can be determined by solving $(\partial A/\partial \boldsymbol{h}) = (\partial A/\partial \boldsymbol{f}) = 0$. In short, the AM algorithm can be summarized as follows:

$$\boldsymbol{h}^{i+1} = \arg\min_{\boldsymbol{h}} A(\boldsymbol{h} \mid \boldsymbol{f}^i) = (\boldsymbol{U}^T\boldsymbol{U} + \beta\boldsymbol{I})^{-1}(\boldsymbol{U}^T\boldsymbol{g} + \beta\boldsymbol{b}) \tag{3.13}$$

$$\boldsymbol{f}^{i+1} = \arg\min_{\boldsymbol{f}} A(\boldsymbol{f} \mid \boldsymbol{h}^{i+1}) = (\boldsymbol{W}^T\boldsymbol{W} + \lambda\boldsymbol{T}(\varepsilon^i))^{-1}\boldsymbol{W}^T\boldsymbol{g} \tag{3.14}$$

We derive the CG optimization method to solve (3.13) and (3.14). The CG optimization utilizes conjugate direction instead local gradient to search for the minima. Therefore, it can achieve faster convergence when compared with steepest descent method [47]. It also requires less storage requirement and computational complexity when compared with Quasi-Newton method. On the other hand, the blur in each channel is estimated separately as they may be independent. The mathematical formulations of the HR image estimation and the PSF estimation are derived in Tables 3.1 and 3.2, respectively.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table 3.1: Summary of CG optimization for HR image reconstruction.

We denote:

$$\begin{aligned} \boldsymbol{d}^i(\boldsymbol{f}^i) &= \frac{\partial A(\boldsymbol{f}^i \mid \boldsymbol{h}^i)}{\partial \boldsymbol{f}^i} = (\boldsymbol{W}^T\boldsymbol{W} + \lambda \boldsymbol{T}(\varepsilon))\boldsymbol{f}^i - \boldsymbol{W}^T\boldsymbol{g} \\ &= \{\sum_k (\boldsymbol{S}_k{}^T \boldsymbol{H}_k{}^T \boldsymbol{H}_c{}^T \boldsymbol{D}^T \boldsymbol{D}\boldsymbol{H}_c\boldsymbol{H}_k\boldsymbol{S}_k) + \lambda \boldsymbol{T}(\varepsilon)\}\boldsymbol{f}^i - \sum_k (\boldsymbol{S}_k{}^T \boldsymbol{H}_k{}^T \boldsymbol{H}_c{}^T \boldsymbol{D}^T \boldsymbol{g}_k) \end{aligned} \tag{3.15}$$

Step 1. Initialize the conjugate vector by setting:

$$\boldsymbol{v}^0 = -\boldsymbol{d}^0(\boldsymbol{f}^0) \tag{3.16}$$

Step 2. At the $(i+1)$-th iteration

Update the i-th HR image:

$$\boldsymbol{f}^{i+1} = \boldsymbol{f}^i + \psi^i \boldsymbol{v}^i, \quad \text{where } \psi^i = \frac{\boldsymbol{d}(\boldsymbol{f}^i)^T \boldsymbol{d}(\boldsymbol{f}^i)}{\boldsymbol{v}^{iT}(\boldsymbol{W}^T\boldsymbol{W} + \lambda \boldsymbol{T}(\varepsilon))\boldsymbol{v}^i} \tag{3.17}$$

Step 3. Update the i-th conjugate vector:

$$\boldsymbol{v}^{i+1} = -\boldsymbol{d}^{i+1}(\boldsymbol{f}^{i+1}) + \eta^i \boldsymbol{v}^i, \quad \text{where } \eta^i = \frac{\boldsymbol{d}(\boldsymbol{f}^{i+1})^T \boldsymbol{d}(\boldsymbol{f}^{i+1})}{\boldsymbol{d}(\boldsymbol{f}^i)^T \boldsymbol{d}(\boldsymbol{f}^i)} \tag{3.18}$$

Step 4. Go to step 2 until convergence or a maximum number of iterations is reached.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table 3.2: Summary of CG optimization for blur estimation.

We denote:

$$\begin{aligned} \boldsymbol{d}^i(\boldsymbol{h}_k^{\,i}) &= \frac{\partial A(\boldsymbol{h}_k^{\,i} \mid \boldsymbol{f}^i)}{\partial \boldsymbol{h}_k^{\,i}} \\ &= (\boldsymbol{F}^T \boldsymbol{S}_k^{\,T} \boldsymbol{H}_c^{\,T} \boldsymbol{D}^T \boldsymbol{D} \boldsymbol{H}_c \boldsymbol{S}_k \boldsymbol{F} + \beta \boldsymbol{I}) \boldsymbol{h}_k^{\,i} - (\boldsymbol{F}^T \boldsymbol{S}_k^{\,T} \boldsymbol{H}_c^{\,T} \boldsymbol{D}^T \boldsymbol{g}_k + \beta \boldsymbol{b}_k^{\,i}) \end{aligned} \tag{3.19}$$

Step 1. Initialize the conjugate vector by setting:

$$\boldsymbol{z}^0 = -\boldsymbol{d}^0(\boldsymbol{h}_k^{\,0}) \tag{3.20}$$

Step 2. At the $(i+1)$-th iteration

Update the i-th $\boldsymbol{h}_k$:

$$\boldsymbol{h}_k^{\,i+1} = \boldsymbol{h}_k^{\,i} + \varphi^i \boldsymbol{z}^i, \quad \text{where } \varphi^i = \frac{\boldsymbol{d}(\boldsymbol{h}_k^{\,i})^T \boldsymbol{d}(\boldsymbol{h}_k^{\,i})}{\boldsymbol{z}^{iT} (\boldsymbol{U}_k^{\,T} \boldsymbol{U}_k + \beta \boldsymbol{I}) \boldsymbol{z}^i} (\boldsymbol{U}_k = \boldsymbol{D} \boldsymbol{H}_c \boldsymbol{S}_k \boldsymbol{F}) \tag{3.21}$$

Step 3. Update the i-th conjugate vector:

$$\boldsymbol{z}^{i+1} = -\boldsymbol{d}^{i+1}(\boldsymbol{h}_k^{\,i+1}) + \rho^i \boldsymbol{z}^i, \quad \text{where } \rho^i = \frac{\boldsymbol{d}(\boldsymbol{h}_k^{\,i+1})^T \boldsymbol{d}(\boldsymbol{h}_k^{\,i+1})}{\boldsymbol{d}(\boldsymbol{h}_k^{\,i})^T \boldsymbol{d}(\boldsymbol{h}_k^{\,i})} \tag{3.22}$$

Step 4. Go to step 2 until convergence or a maximum number of iterations is reached.

The schematic overview of the proposed method is given in Fig. 3.1. The procedure of the proposed soft MAP algorithm on blind image SR is explained as follows. Given the LR images, we first perform image registration[2] by using the phase correlation method proposed by Foroosh et al. in [141]. After the initialization, we perform the AM algorithm for joint estimation of the blurs and the HR image by using (3.13) and (3.14) iteratively. During the blur estimation, the proposed scheme determines and updates the best-fit blur model in the AM iteration. It progressively integrates the parametric information of blur structures into the SR

[2] The effect of the registration errors in the blind SR work is discussed in Appendix B of this chapter.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

image reconstruction.

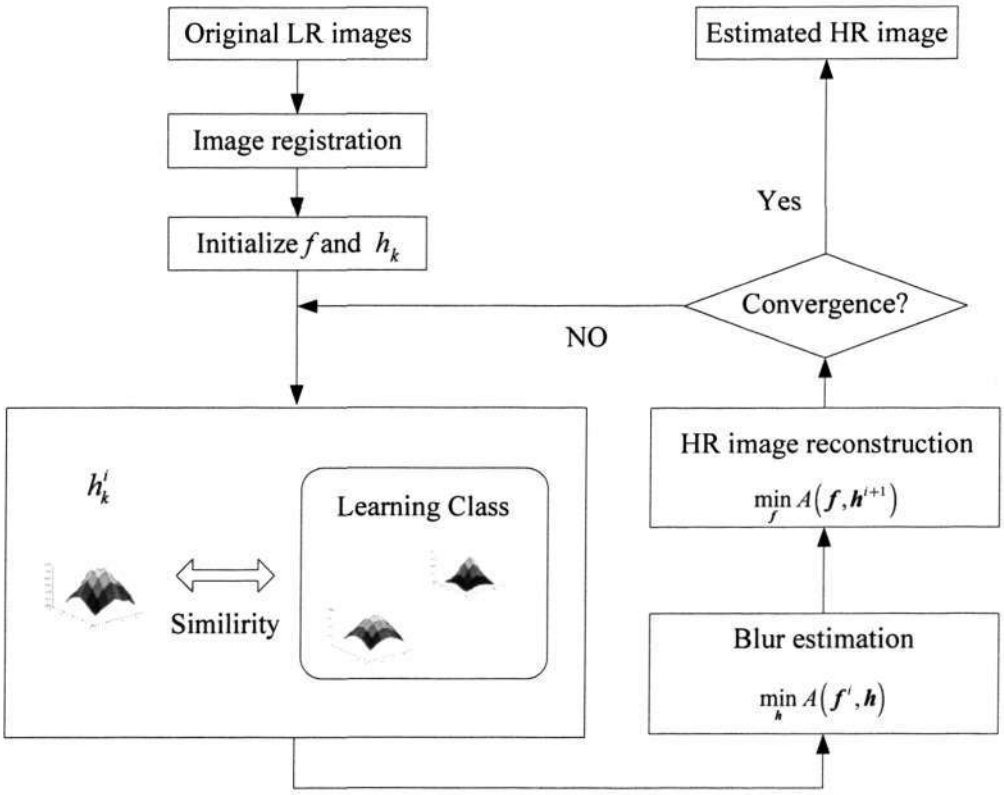


Fig. 3.1: Schematic diagram of the proposed algorithm.

3.3.4 Initialization and Selection of Regularization Parameters

It is useful if we can have good initial estimates of the SR image and the blurs, as this will enhance the convergence rate of the algorithm. The HR image $\boldsymbol{f}^0$ is initialized based on cubic interpolation of the LR images. The following algorithm can obtain good initial estimates of blurs. We use a classical blur prior $p(\boldsymbol{h}_k) \propto \exp\{-\boldsymbol{h}_k{}^T \boldsymbol{L}^T \boldsymbol{L} \boldsymbol{h}_k\}$ where $\boldsymbol{L}$ is the Laplacian operator, to ensure the blur smoothness. Then we minimize the following cost function to obtain the initial blur estimates:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\boldsymbol{h}_k^{\,0} = \arg\min_{\boldsymbol{h}_k} A(\boldsymbol{f}^0, \boldsymbol{h}_k) = (\boldsymbol{U}_k^{\,T}\boldsymbol{U}_k + \alpha_k \boldsymbol{L}^T\boldsymbol{L})^{-1}(\boldsymbol{U}_k^{\,T}\boldsymbol{g}_k) \tag{3.23}$$

A few papers have addressed the problem of estimating the optimal regularization parameters for image super-resolution [32], [64], [65]. However, the complexities of these methods make them incompatible with the simplicity of the proposed algorithm. Furthermore, the selection issue of regularization parameters, is up to a certain extent, heuristic. This is due to the lack of a widely accepted quantitative measure of image quality to decide rigorously which estimate is the best. We follow the algorithm [124] to estimate the regularization parameters λ, β and α_k. The basic idea behind this method is the original image and correct blurs should satisfy (3.13), (3.14) and (3.23) when converged. The equations containing the regularization parameters are solved using the least squares method [124]. The algorithm described here gives an order-of-magnitude estimate for the regularization parameters. The experimental results show that the algorithm is robust towards different regularization parameters so long as they fall within a reasonable range.

3.4 Experiments

In this section, we illustrate the performance of the proposed method on two different sets of data: simulated and real-life images. To evaluate the performance, we used the following well-known metrics in this work: normalized mean squared error (NMSE) for the estimated blurs and peak signal-to-noise ratio (PSNR) for the reconstructed HR image. They are defined as follows [142]:

$$NMSE(\hat{\boldsymbol{h}}) = 100\frac{\|\hat{\boldsymbol{h}} - \boldsymbol{h}\|^2}{\|\boldsymbol{h}\|^2} \tag{3.24}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$PSNR(\hat{f}) = 10\log_{10}\frac{255^2}{\frac{1}{M_f \times N_f}\| f - \hat{f} \|^2} \tag{3.25}$$

Generally, good algorithm is reflected by low NMSE and high PSNR. Nevertheless, the best performance measure remains human inspection of the reconstructed HR images.

3.4.1 Blind Super-resolution on Simulated Images with Additive Noise

The first experiment is based on the simulated images. In this experiment, we created a sequence of LR frames from a HR image. The "Text" image in Fig. 3.2(a) was selected as the test image. To generate the LR images, the HR image was shifted by random translations of (0,0), (0.15,0.75), (0.65,0.25), (1,1) pixels, and blurred by four out-of-focus blurs with different radii (2.0, 2.5, 3.0, 3.5). The results were then blurred by 2×2 uniform blur as a decimation factor of 2 was chosen. A sample of the LR images degraded by the out-of-focus blur with a radius of 2.0, scaled up to the same size as the HR image, is shown in Fig. 3.2(b). Fig. 3.2(d) shows the HR image reconstructed by the interpolation-deblurring method highlighted in [8]. To implement the method, we first performed image registration using [141] for all the given LR images. Then we used cubic interpolation [20] to reconstruct the blurred HR image. The result is shown in Fig. 3.2(c). Next, blind image deconvolution method in [136] was performed to obtain a deblurred HR image. It is observed that the obtained result is less than satisfactory. When the blurs are not identical for all the observed LR images, the result is unsatisfactory as shown in this experiment. The proposed algorithm was run to perform blind SR image reconstruction. We adopted $\lambda = 10^{-5}$, $\alpha_k = 10^3$ and $\beta = 50$ for the regularization parameters. The simulations showed that the algorithm was robust toward different scaling coefficients so long as they fall within a reasonable range. The iterative AM algorithm was terminated when the maximum number of 25 iterations was reached. The result obtained using the proposed

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

algorithm is given in Fig. 3.2(f). We also compare our result with the reconstructed image in Fig. 3.2(g) when the exact blurs are known. This is achieved by using the exact blurs in (3.14) to reconstruct the HR image. Comparing our result with the reconstructed HR image in Fig. 3.2(g), we observe that despite our algorithm not having the knowledge of the exact blurs, it can still recover the HR image effectively. Our result is almost as good as the reconstructed image with the exact blurs. The PSNR of the reconstructed HR image using the proposed method is plotted in Fig. 3.5. It is observed that our result converges toward the case when the exact blurs are known. In other words, the PSNR values by both methods are very close after convergence. This demonstrates the effectiveness of the proposed scheme in blind SR reconstruction. In order to illustrate the performance of the blur identification, Fig. 3.6 shows the NMSE of the identified blurs by the proposed method. The NMSE tends to zero as the number of AM iterations increase. These objective performance measures further reconfirm our subjective evaluation of the reconstructed image. The satisfactory image quality is due to the proposed soft MAP framework performing joint blur identification and HR image reconstruction. We also applied the algorithm in [56] on the degraded LR images. The reconstructed HR image is shown in Fig. 3.2(e). It is observed that the proposed method offered better result than the method in [56]. In the blur estimation, the proposed method renders a NMSE of 0.002 as compared to NMSE of 3.5 offered by the method in [56]. This is because the proposed method utilizes the parametric blur information, hence resulting in better blur estimation. The effectiveness of the blur estimation in turn leads to the good performance of HR image reconstruction.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

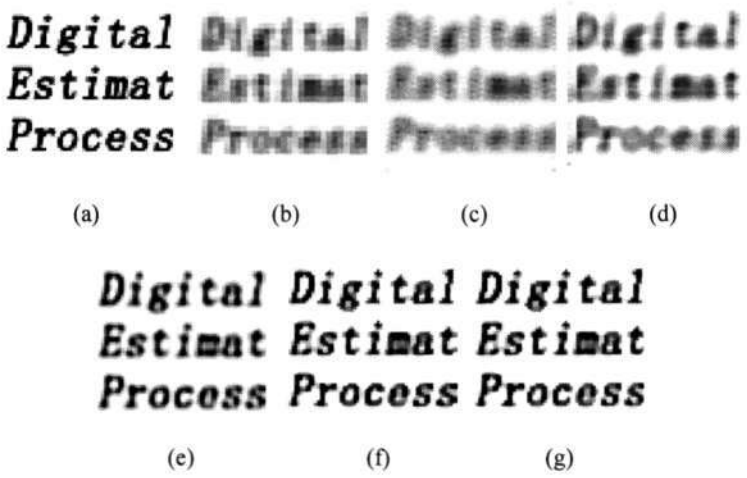


Fig. 3.2: Blind image SR on simulated LR images degraded by out-of-focus blur without noise. (a) Original HR image, (b) A sample of the scaled-up LR images degraded by the out-of-focus blur with a radius of 2.0, (c) Reconstructed image using cubic interpolation without deblurring, (d) Reconstructed image using the interpolation-deblurring method, (e) Reconstructed image using the method in [56], (f) Reconstructed image using the proposed blind SR algorithm, (g) Reconstructed image using exactly known blur.

In order to illustrate the effectiveness of the proposed method to handle LR images degraded by AWGN at different noisy levels, AWGN is added on the LR images obtained in the previous experiment to produce two sets of signal-to-noise ratios (SNRs) at 30dB and 20dB, respectively. It is noted that a SNR of 20dB for image SR can be considered as a very noisy environment. The same experiment as before was conducted again. We adopted $\lambda = 5*10^{-5}$, $\alpha_k = 10^3$ and $\beta = 50$ for SNR = 30dB and $\lambda = 5*10^{-4}$, $\alpha_k = 5*10^3$ and $\beta = 500$ for SNR = 20dB. The reconstructed HR images using the proposed algorithm and the method which assumes the exact blurs are known for different noise levels are given in Figs. 3.3(e)-(f) and Figs. 3.4(e)-(f), respectively. The PSNR of the reconstructed HR images and the NMSE of the estimated blurs by the proposed method are plotted in Figs. 3.5 and 3.6, respectively. The HR images reconstructed by the interpolation-deblurring method [8] for

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

different noise levels are given in Fig. 3.3(c) and Fig. 3.4(c), respectively. The HR images reconstructed by the algorithm in [56] for different noise levels are given in Fig. 3.3(d) and Fig. 3.4(d), respectively. From Figs. 3.3 and 3.4, it is observed that the proposed method produces the best results when compared with the results obtained using the interpolation-deblurring method [8] and the algorithm in [56]. It is noted that the reconstructed HR image by the proposed method is almost as good as the method which assumes the exact blurs are known for the same noise level. These results demonstrate the effectiveness of the proposed method in performing blind image SR. From Fig. 3.6, it is observed that the proposed method can achieve accurate blur estimation for different noise levels up to 20dB SNR. From Fig. 3.5, it is noted that the performance of the HR image reconstruction drops slightly for moderate noise level (e.g. 30dB SNR) and starts to decline for high noise level (e.g. 20dB SNR). Nevertheless, this observation is in agreement with previous studies on blind image restoration and super-resolution [143], [144]. We also computed the biases and variances of the blur estimates for each channel, which are shown in Figs. 3.7-3.8, where the x axis represents the coefficients of the channel blurs ordered lexicographically and the y axis represents the values of biases/variances for the channel blurs. It can be seen that the bias for each channel blur is small and tends to zero.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

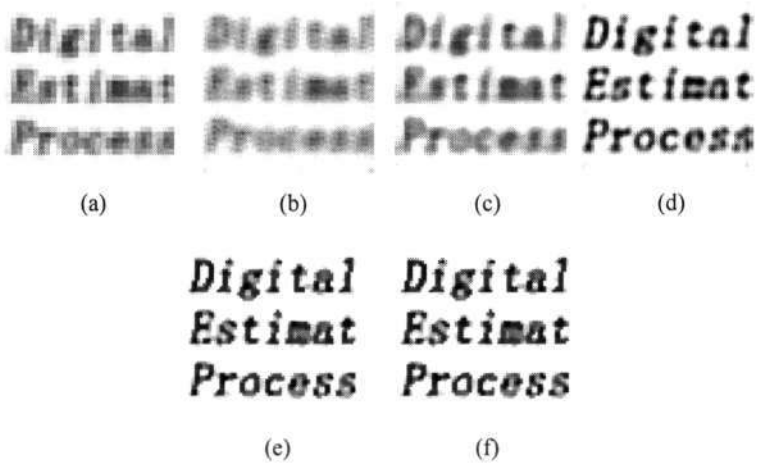


Fig. 3.3: Blind image SR on LR images degraded by out-of-focus blur with 30dB additive noise. (a) A sample of the scaled-up LR images degraded by the out-of-focus blur with a radius of 2.0, (b) Reconstructed image using cubic interpolation without deblurring, (c) Reconstructed image using the interpolation-deblurring method, (d) Reconstructed image using the method in [56], (e) Reconstructed image using the proposed blind SR algorithm, (f) Reconstructed image using exactly known blur.

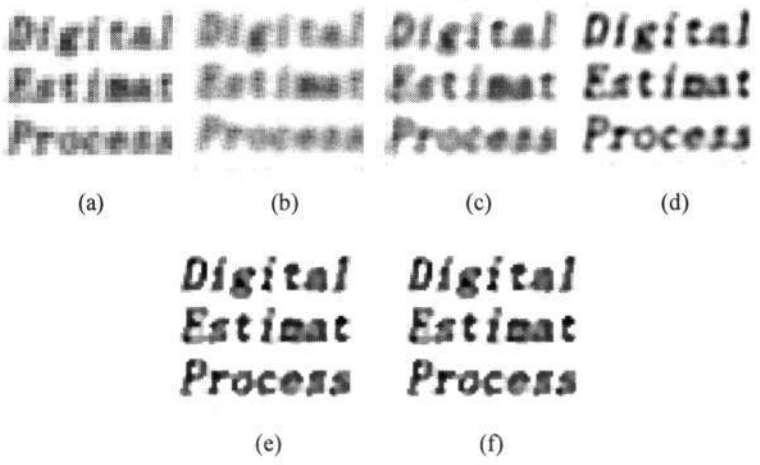


Fig. 3.4: Blind image SR on LR images degraded by out-of-focus blur with 20dB additive noise. (a) A sample of the scaled-up LR images degraded by the out-of-focus blur with a radius of 2.0, (b) Reconstructed image using cubic interpolation without deblurring, (c) Reconstructed image using the interpolation-deblurring method, (d) Reconstructed image using the method in [56], (e) Reconstructed image using the proposed blind SR algorithm, (f) Reconstructed image using exactly known blur.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

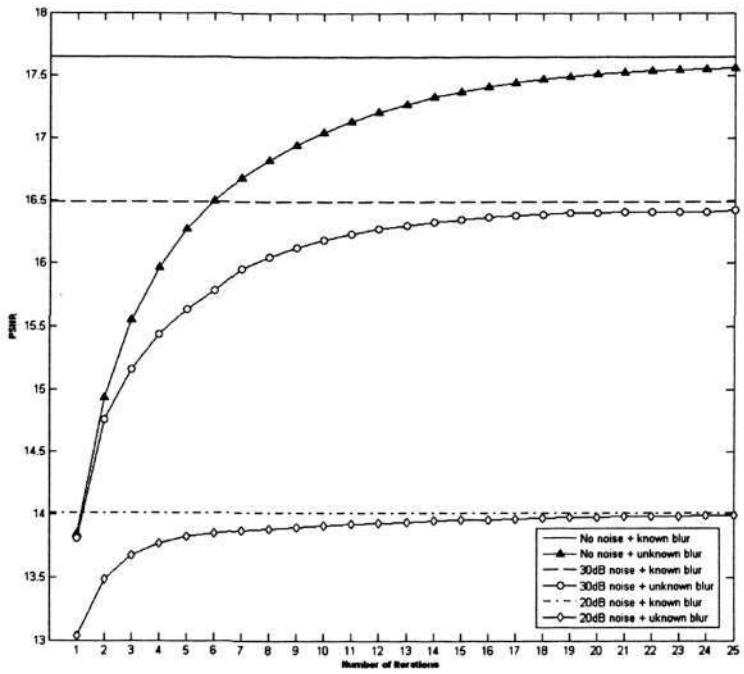


Fig. 3.5: PSNR of the reconstructed HR image.

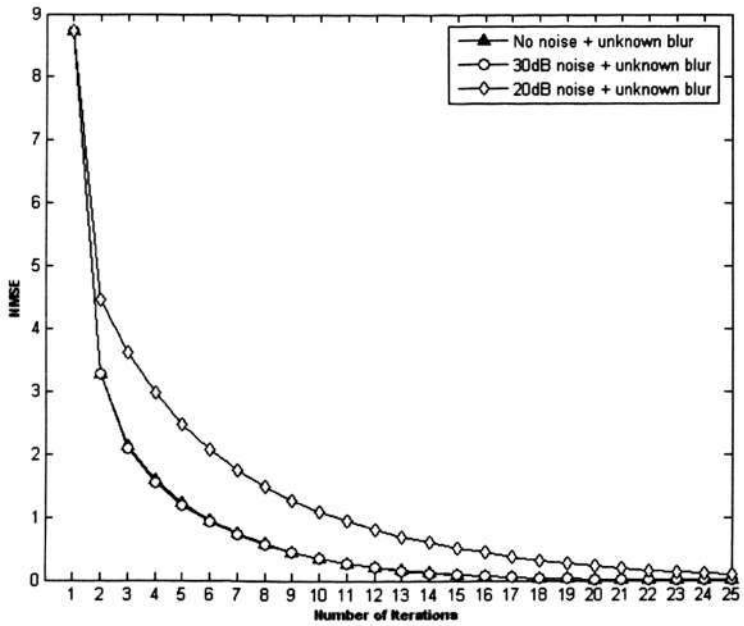
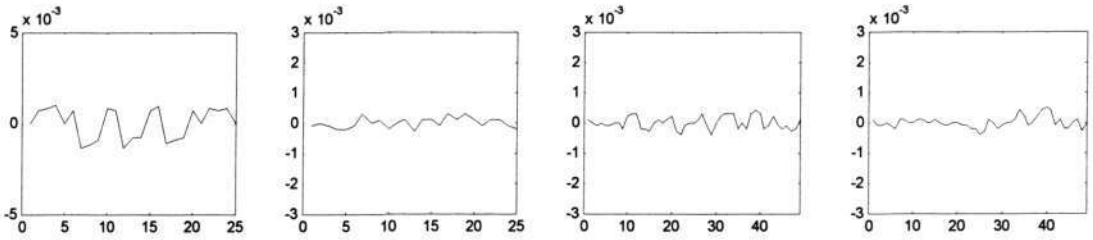


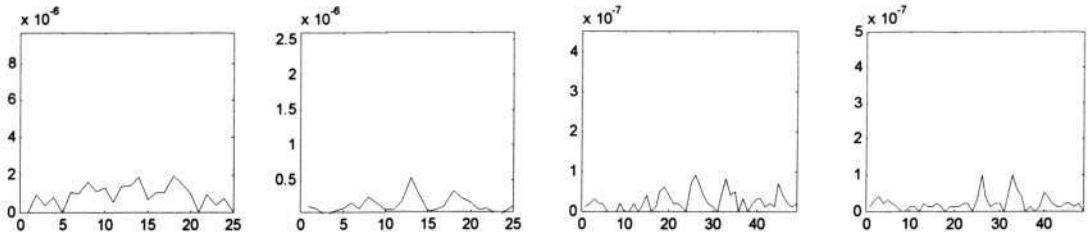
Fig. 3.6: NMSE of the estimated blurs.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



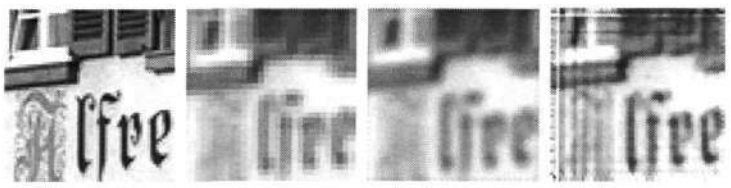
(a) (b) (c) (d)

Fig. 3.7: Bias of the blur estimate for each channel (h_1-h_4)



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

Fig. 3.8: Variance of the blur estimate for each channel (h_1-h_4)



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

(e) (f) (g)

Fig. 3.9: Blind image SR on LR images degraded by identical blurs with 35dB additive noise. (a) Original HR image, (b) A sample of the scaled-up LR images, (c) Reconstructed image using cubic interpolation without deblurring, (d) Reconstructed image using the interpolation-deblurring method, (e) Reconstructed image using the method in [56], (f) Reconstructed image using the proposed blind SR algorithm, (g) Reconstructed image using exactly known blur.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Next we study the capability of the proposed algorithm to handle the blurs not belonging to the predefined classes in Fig. 3.9. The "building" image in Fig. 3.9(a) was selected as the test image. A 5×5 nonstandard exponential blur was used for all the blurred channels. The blurs are identical for all the observed LR images. The nonstandard blur is given by $h_k(x,y)=\xi\exp(-\sigma\sqrt{x^2+y^2})$, where $\sigma=0.4$ is the decay factor, and ξ is the normalizing constant. The blurred LR image was further degraded by a 35 dB AWGN as shown in Fig. 3.9(b). Fig. 3.9(d) shows the HR image reconstructed by the interpolation-deblurring method. There are noticeable artifacts in the reconstructed HR image, especially near the wall. This is due to the disjoint process where the SR is followed by deblurring. Although the blurs in each channel are identical, the errors occurred during the reconstruction of the blurred HR image still affect the subsequent HR image deconvolution. The reconstructed image by the proposed algorithm is given in Fig. 3.9(f). Comparison with the results in Fig. 3.9(e) using the method in [56] and Fig. 3.9(g) where the exact blurs are known demonstrates the effectiveness of the proposed method.

Finally, we will provide an analysis on the computational complexity of the proposed method. As the exact number of arithmetic operations required is difficult to determine for the proposed method, we will use computational time to show the complexity of the proposed method. We used the method in [56] as the benchmark for comparison. In order to provide a good indication of the computational time, we used the above simulation, in which we reconstruct 64×64 "text" HR images based on noiseless LR images and noisy LR images at SNRs of 30dB and 20dB. The simulation environments are given as follows: Windows XP, MATLAB 7.1, CPU P4-3.4GHz, and 1G RAM. The average computational time is 96.4s for the proposed method, and 114.4s for the method in [56]. This shows that the proposed method has better computational time. It should also be noted that the proposed method is implemented in

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

MATLAB, which is an interpreter-programming platform with lower processing speed. The computational time of the experiments can be improved if the algorithm is implemented in a compiler language such as C.

3.4.2 Blind Super-resolution on Real-life Images

The real-life experiment was conducted by capturing four LR "calendar" images using a Canon PowerShot S60 camera under the wide-angle setting with relative shifting. This gives rise to four LR images that experience blurring. The degraded LR images are shown in Fig. 3.10(a). We chose a decimation factor of 2 so that the estimated HR image will have twice the resolution of the LR images. The phase correlation algorithm in [141] was used to estimate the relative shifts among the LR images. Next the proposed soft MAP framework was run to perform blind SR image reconstruction. We adopted $\lambda = 10^{-3}$, $\alpha_k = 10$ and $\beta = 100$ for regularization parameters in this experiment.

The reconstructed HR image using the proposed method is given in Fig. 3.10(e). It can be seen that the overall sharpness of the image has been recovered. Significant amount of details has also been restored near the number region. There is also no noticeable ringing artifact in the restored image. We compared the performance of the proposed algorithm with the interpolation-deblurring approach [8] and the method in [56]. The results are presented in Figs. 3.10 (c-d), respectively. Comparison reveals that our approach is superior in handling real-life blind SR image reconstruction, as it is able to recover more visual clarity from the blurred LR images particularly near the textual region.

3.5 Summary

In this chapter, we propose a new algorithm to address blind SR image reconstruction. The new

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

method performs joint HR image reconstruction and blur identification from the observed LR images. It develops a MAP framework that incorporates likelihood and prior modeling of the HR image. The blur prior, on the other hand, involves estimating the best-fit parametric blur model, and induces reinforcement learning towards it. Experimental results show that the new method is effective in multi-frame SR where there is limited information about the blurring functions.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

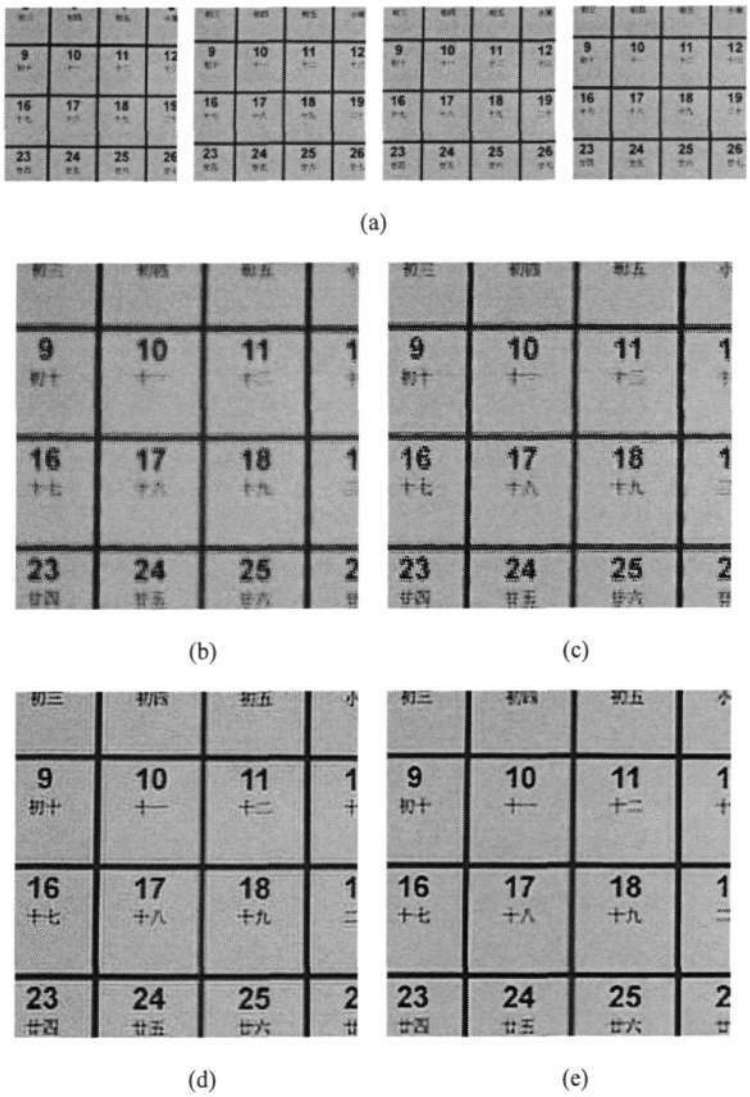


Fig. 3.10: Blind image SR on real-life images. (a) Four LR images, (b) Reconstructed image using cubic interpolation without deblurring, (c) Reconstructed image using the interpolation-deblurring method, (d) Reconstructed image using the method in [56], (e) Reconstructed image using the proposed blind SR algorithm.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Appendix A: Construction of $T(\varepsilon)$

We will first explain how to construct $\boldsymbol{T}_x(\varepsilon)$ based on the gradient of the image in the x direction. $\boldsymbol{T}_x(\varepsilon)$ should satisfy the following condition:

$$\sum_x \sum_y \frac{1}{\varepsilon(\underset{\sim}{x})}(f(x+1,y)-f(x,y))^2 = \boldsymbol{f}^T \boldsymbol{T}_x(\varepsilon)\boldsymbol{f} \tag{3A.1}$$

Adopting the scheme in [136], the coefficient $\varepsilon(\underset{\sim}{x})$ can be calculated by using the estimated HR image f in the previous iteration of the AM loop. Suppose the values of $\frac{1}{\varepsilon(\underset{\sim}{x})}$ at each lexicographical location are given by $\{a,b,c\cdots d,e\}$, then $\sum_x \sum_y (f(x+1,y)-f(x,y))^2$ can be expressed in the vector-matrix form as follows:

$$\sum_x \sum_y (f(x+1,y)-f(x,y))^2 = \|\boldsymbol{Uf}\|_2^2 = \boldsymbol{f}^T \boldsymbol{U}^T \boldsymbol{Uf}$$

$$\text{where } \boldsymbol{U} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3A.2}$$

The left hand side of equation (3A.1) can be computed as follows:

$$\sum_x \sum_y \frac{1}{\varepsilon(\underset{\sim}{x})}(f(x+1,y)-f(x,y))^2 = \boldsymbol{f}^T \boldsymbol{U}^T \boldsymbol{ZUf} \tag{3A.3}$$

where $\boldsymbol{Z}$ is a diagonal matrix. It is given as $\boldsymbol{Z} = diag\{a,b,c\cdots d,e\}$. Then $\boldsymbol{T}_x(\varepsilon)$ can be computed by $\boldsymbol{T}_x(\varepsilon) = \boldsymbol{U}^T \boldsymbol{ZU}$. The extension to the y direction follows the similar derivation. Then, $\boldsymbol{T}(\varepsilon)$ can be obtained by: $\boldsymbol{T}(\varepsilon) = \boldsymbol{T}_x(\varepsilon) + \boldsymbol{T}_y(\varepsilon)$, where $\boldsymbol{T}_y(\varepsilon)$ is constructed by using $\varepsilon(\underset{\sim}{y})$. The advantage of using this scheme is that it can effectively circumvent the difficulty of solving the nonlinear PDEs arising out of the TV norm. Further information on this scheme can

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

be found in [124].

Appendix B: Effect of Registration Errors

To analyze the effect of the registration errors in the proposed blind SR method, the SR formulation in (3.1) can be re-expressed in vector-matrix form as follows:

$$\begin{aligned} \boldsymbol{g}_k &= \boldsymbol{W}_k \boldsymbol{f} + \boldsymbol{n}_k \\ &= (\hat{\boldsymbol{W}}_k + \Delta\hat{\boldsymbol{W}}_k)\boldsymbol{f} + \boldsymbol{n}_k \\ &= \hat{\boldsymbol{W}}_k \boldsymbol{f} + \hat{\boldsymbol{n}}_k \qquad (\hat{\boldsymbol{n}}_k = \Delta\hat{\boldsymbol{W}}_k \boldsymbol{f} + \boldsymbol{n}_k) \end{aligned} \tag{3B.1}$$

where $\hat{\boldsymbol{W}}_k$ is the estimate for $\boldsymbol{W}_k$. We use $\Delta\hat{\boldsymbol{W}}_k$ to represent the difference between $\hat{\boldsymbol{W}}_k$ and $\boldsymbol{W}_k$ due to the registration errors. Following the SR works in [64] and [65], the residual $\hat{\boldsymbol{n}}_k$ can be approximated as Gaussian noise. Therefore, the registration errors in the proposed blind SR method can be considered as the introduction of extra Gaussian noise in the SR formulation. Currently, there are several registration algorithms for the translational movement which are able to provide satisfactory performance at the subpixel level, such as the phase correlation-based method in [141]. Experiments have shown that the estimate of the motion parameters obtained using the existing registration technique [141] is sufficient to produce satisfactory results for the HR image reconstruction and blur estimation.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 4

Joint Image Registration and Super-resolution

4.1 Introduction

In multi-frame SR problem, the LR images are typically shifted up to subpixel level, and hence the information available in each LR image can be extracted and combined to obtain a HR image. Registration is a critical step in multi-frame SR, particularly when the motion is not purely translational between the LR images. Conventional SR algorithms [145] assume the motion parameters are known *a priori* (such as the image-formation system using multiple CCD sensor arrays in [67]). This assumption, however, is only practical in certain applications.

Current SR techniques [4], [8], [9], [42], [146] are commonly performed in two disjoint stages, namely (i) image registration from LR images, followed by (ii) inverse estimation that integrates image fusion and deblurring into a single step. Generally, these SR algorithms ignore registration errors and assume the estimated motion parameters by existing registration methods to be error-free. Nevertheless, due to the presence of aliasing in the captured LR images, most existing registration algorithms for aliased images still experience registration errors. For instance, a frequency domain-based algorithm in [20] is developed for registering a set of aliased LR images with application to SR. Rotational and translational parameters are estimated based on the low-frequency part of the LR images. Nevertheless, accurate registration

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in the LR domain remains difficult to achieve, giving rise to suboptimal results in the reconstructed HR images [62]. In the previous chapter, we have analyzed the effect of the registration errors. However, the analysis can be only applicable to global translational shift motion model. This motivates the study of progressive image SR in this chapter, which takes into account the impact of unreliable initial registration while using a more flexible motion model that includes both translation and rotation.

This chapter proposes a new framework for simultaneous image registration and HR image reconstruction. The main contribution of this chapter is two-fold. Firstly, the proposed method integrates image registration and SR into a single estimation process. As opposed to the current two-stage SR methods [20] that perform registration on the LR images, the image registration in the proposed method is performed iteratively using the progressively estimated HR image. This is promising as more accurate motion parameters can be determined, thereby enhancing the performance of the HR reconstruction. Further, the new method can overcome the shortcoming of the iterative alternating minimization (AM) framework [31], as the cost function is not projected onto the image and motion parameter domains one at a time. Instead, an iterative scheme based on a nonlinear least squares (NLS) method is developed to estimate the motion parameters and the HR image simultaneously. Although there is no guarantee of global convergence through a complete mathematical study, various experiments have demonstrated that our algorithm is effective in achieving satisfactory convergence and results.

The second contribution of this chapter is that a more flexible motion model that consists of translational and rotational motion is developed. As opposed to the translational motion model used in other SR methods, the adopted motion model is more realistic. It is noted that with this new model, a more challenging problem arises. The motion operator $\boldsymbol{S}_k$ is no longer Toeplitz, thus, $\boldsymbol{S}_k \boldsymbol{f}$ is not linear with respect to motion (rotation and translation) paramters. This chapter

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

will address this issue by deriving the Jacobian matrix for our SR problem. Experimental results in later sections show that the proposed method is effective in performing image registration and SR for simulated as well as real-life images.

The rest of this chapter is organized as follows. The problem formulation of image SR taking account of registration errors is introduced in Section 4.2. An iterative algorithm using the NLS method is developed in Section 4.3. Experimental results on simulated and real-life images are presented and discussed in Section 4.4. A brief summary is given in Section 4.5.

4.2 Problem Formulation

The SR formulation in (3.2) can then be re-expressed in a vector-matrix form as:

$$\boldsymbol{g} = \boldsymbol{W}(\boldsymbol{\alpha})\boldsymbol{f} + \boldsymbol{n} \tag{4.1}$$

where the matrix $\boldsymbol{W}(\boldsymbol{\alpha})$ is formed by nonlinear, differentiable functions of an unknown motion parametric vector $\boldsymbol{\alpha}$. Without loss of generality, we have assumed the first LR image to be the referenced image. Hence, $\boldsymbol{W}(\boldsymbol{\alpha})$ can be rewritten as:

$$\boldsymbol{W}(\boldsymbol{\alpha}) = \begin{bmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2(\boldsymbol{\alpha}_2) \\ \vdots \\ \boldsymbol{W}_N(\boldsymbol{\alpha}_N) \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}\boldsymbol{H}_1 \\ \boldsymbol{D}\boldsymbol{H}_2\boldsymbol{S}_2(\boldsymbol{\alpha}_2) \\ \vdots \\ \boldsymbol{D}\boldsymbol{H}_N\boldsymbol{S}_N(\boldsymbol{\alpha}_N) \end{bmatrix}; \boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_N \end{bmatrix} = \begin{bmatrix} [\cos\theta_2, \sin\theta_2, s_{x2}, s_{y2}]^T \\ \vdots \\ [\cos\theta_N, \sin\theta_N, s_{xN}, s_{yN}]^T \end{bmatrix}$$

where $\boldsymbol{D}$ is the decimation operator. $\boldsymbol{H}_k$ represents the blurring process, such as optics or sensor blurring. $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)$ $(2 \le k \le N)$ represents the geometric motion operator for the k-th LR image. In this work, we consider the initial estimated motion parameters contain some errors. This is a more realistic assumption, as accurate registration for SR is difficult to achieve especially at the beginning of the algorithm. The objective of image SR, in this context, is to

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

reconstruct the HR image $\boldsymbol{f}$ from N observed LR images with the unknown motion parametric vector $\boldsymbol{\alpha}$. As the first LR image is used as the reference, we need to estimate $3\times(N-1)$ unknown motion parameters and the HR image.

4.3 A Nonlinear Least Squares-based Super-resolution Algorithm

4.3.1 Iterative Simultaneous High-resolution Estimation and Registration

Current disjoint two-stage SR methods perform image registration on the LR images, followed by HR reconstruction. The disadvantage of these methods is that they rely on the rough initial registration heavily, hence reducing the performance of subsequent HR reconstruction. To address this difficulty, we propose an effective iterative algorithm that integrates image registration and SR into a single estimation process.

The motion model that is considered in this work includes both translation as well as rotation. A challenge arising as a result of this is that the motion operator $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)$ is no longer Toeplitz. Thus, traditional methods cannot address image SR effectively [147]. In view of this, an NLS-based method is proposed. In this section, we will explain how the nonlinear parametric estimation problem is formulated and solved to simultaneously estimate the motion parameters and the HR image.

Generally, the SR image reconstruction is an ill-posed problem. To address this issue, we adopt the MAP estimation method in Section 3.3. Different from (3.3), we drop the pdf $p(\boldsymbol{h})$, since we assume the blurring as known *a priori*. Then it can be shown that the estimate of the HR image $\boldsymbol{f}$ can be obtained by minimizing the following cost function:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$A(\boldsymbol{\alpha}, \boldsymbol{f}) = \frac{1}{2}(\boldsymbol{g} - \boldsymbol{W}(\boldsymbol{\alpha})\boldsymbol{f})^T(\boldsymbol{g} - \boldsymbol{W}(\boldsymbol{\alpha})\boldsymbol{f}) + \frac{\lambda}{2}\boldsymbol{f}^T\boldsymbol{T}(\varepsilon)\boldsymbol{f} \tag{4.2a}$$

$$= \frac{1}{2}\left\| \begin{matrix} r_1(\boldsymbol{f}) \\ \vec{r}(\boldsymbol{\alpha}, \boldsymbol{f}) \\ \sqrt{\lambda}\boldsymbol{L}(\varepsilon)\boldsymbol{f} \end{matrix} \right\|^2 \tag{4.2b}$$

where $\|\cdot\|$ denotes the L_2 norm. $r_1(\boldsymbol{f}) = \boldsymbol{g}_1 - \boldsymbol{W}_1\boldsymbol{f}$ and $\vec{r}(\boldsymbol{\alpha}, \boldsymbol{f}) = \vec{\boldsymbol{g}} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha})\boldsymbol{f}$ are defined as the residual vectors, where $\vec{\boldsymbol{g}} = \left[\boldsymbol{g}_2^T, \ \cdots \ \boldsymbol{g}_N^T\right]^T$ and $\vec{\boldsymbol{W}}(\boldsymbol{\alpha}) = \left[\boldsymbol{W}_2^T(\boldsymbol{\alpha}_2), \ \cdots \ \boldsymbol{W}_N^T(\boldsymbol{\alpha}_N)\right]^T$. λ is the regularization parameter, which provides a compromise between the first and second terms in (4.2a). To choose λ, we follow the algorithm in [124] to obtain an order-of-magnitude estimate. The basic idea behind this method is that the original image should satisfy the equation $\partial A / \partial \boldsymbol{f} = 0$. Then the equation is solved using the least squares method. The definition of $\boldsymbol{T}(\varepsilon)$ can be referred in Chapter 3. Its pseudo-decomposition is given by $\boldsymbol{T}(\varepsilon) = \boldsymbol{L}(\varepsilon)^T\boldsymbol{L}(\varepsilon)$. In the proposed method, it is unnecessary to compute $\boldsymbol{L}(\varepsilon)$ explicitly during the minimization process. This will become clear in Section 4.3.2.

As the proposed method adopts the motion model comprising the translation and rotation, the minimization problem in (4.2b) is linear with respect to $\boldsymbol{f}$ but nonlinear with respect to $\boldsymbol{\alpha}$. In order to solve this optimization problem, we develop an NLS-based approach to estimate the motion parameters and the HR image simultaneously. We extend the principle of the nonlinear parametric estimation algorithm in [148] to derive a linear approximation for $\vec{r}(\boldsymbol{\alpha}, \boldsymbol{f})$. Let $\Delta\boldsymbol{f}$ represent a small change in the HR image $\boldsymbol{f}$, and $\Delta\boldsymbol{\alpha}$ a small change in the motion vector $\boldsymbol{\alpha}$. Ignoring the second order term $\boldsymbol{J}(\Delta\boldsymbol{f}, \boldsymbol{\alpha})\Delta\boldsymbol{\alpha}$ since its value is generally small with small $\Delta\boldsymbol{f}$ and $\Delta\boldsymbol{\alpha}$, the residual vector $\vec{\boldsymbol{r}}(\boldsymbol{\alpha}, \boldsymbol{f})$ can be linearized with respect to $\Delta\boldsymbol{\alpha}$ and $\Delta\boldsymbol{f}$ as follows:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$
\begin{aligned}
&\vec{\boldsymbol{r}}(\boldsymbol{\alpha}+\Delta\boldsymbol{\alpha}, \boldsymbol{f}+\Delta\boldsymbol{f}) \\
&= \vec{\boldsymbol{g}} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha}+\Delta\boldsymbol{\alpha})(\boldsymbol{f}+\Delta\boldsymbol{f}) \\
&= \vec{\boldsymbol{g}} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha}+\Delta\boldsymbol{\alpha})\boldsymbol{f} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha}+\Delta\boldsymbol{\alpha})\Delta\boldsymbol{f} \\
&= \vec{\boldsymbol{g}} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha})\boldsymbol{f} - \boldsymbol{J}(\boldsymbol{f}, \boldsymbol{\alpha})\Delta\boldsymbol{\alpha} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha})\Delta\boldsymbol{f} - \boldsymbol{J}(\Delta\boldsymbol{f}, \boldsymbol{\alpha})\Delta\boldsymbol{\alpha} \\
&\simeq \vec{\boldsymbol{r}}(\boldsymbol{f}, \boldsymbol{\alpha}) - \boldsymbol{J}(\boldsymbol{f}, \boldsymbol{\alpha})\Delta\boldsymbol{\alpha} - \vec{\boldsymbol{W}}(\boldsymbol{\alpha})\Delta\boldsymbol{f}
\end{aligned}
\tag{4.3}
$$

where $\boldsymbol{J}(\boldsymbol{\alpha}, \boldsymbol{f})$ is the Jacobian of $\vec{\boldsymbol{W}}(\boldsymbol{\alpha})\boldsymbol{f}$ with respect to $\boldsymbol{\alpha}$. We will discuss how to construct $\boldsymbol{J}(\boldsymbol{\alpha}, \boldsymbol{f})$ in Section 4.3.2. Therefore, given the current estimate of the HR image $\boldsymbol{f}$ and the motion vector $\boldsymbol{\alpha}$, the minimization problem in (4.2b) can be modified as:

$$
\min_{\Delta\boldsymbol{\alpha}, \Delta\boldsymbol{f}} \left\| \begin{pmatrix} \boldsymbol{0} & \boldsymbol{W}_1 \\ \boldsymbol{J}(\boldsymbol{\alpha}, \boldsymbol{f}) & \vec{\boldsymbol{W}}(\boldsymbol{\alpha}) \\ \boldsymbol{0} & \sqrt{\lambda}\boldsymbol{L}(\varepsilon) \\ \boldsymbol{R} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \Delta\boldsymbol{\alpha} \\ \Delta\boldsymbol{f} \end{pmatrix} + \begin{pmatrix} -\boldsymbol{r}_1(\boldsymbol{f}) \\ -\vec{\boldsymbol{r}}(\boldsymbol{\alpha}, \boldsymbol{f}) \\ \sqrt{\lambda}\boldsymbol{L}(\varepsilon)\boldsymbol{f} \\ \boldsymbol{0} \end{pmatrix} \right\|^2 \tag{4.4}
$$

Compared with (4.2b), we have added a new term $\| \boldsymbol{R}\Delta\boldsymbol{\alpha} \|^2$ into (4.4). This is the regularization functional that introduces stability into the solution $\Delta\boldsymbol{\alpha}$. $\boldsymbol{R}$ is an adaptive regularization matrix, which is given by $\boldsymbol{R} = \sqrt{\beta}\boldsymbol{I}$, where β is the regularization parameter in the motion parameter domain and $\boldsymbol{I}$ is a $4(N-1)\times 4(N-1)$ identity matrix. β is chosen similar to the idea given in [147] to ensure the stability of the solutions. By using the proposed method, the original problem for the direct estimation of $\boldsymbol{\alpha}$ and $\boldsymbol{f}$ by minimizing (4.2b) has been transformed into the minimization problem for the increment $\Delta\boldsymbol{\alpha}$ and $\Delta\boldsymbol{f}$ in (4.4). The proposed iterative algorithm to perform simultaneous image registration and HR reconstruction is summarized in Table 4.1.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table 4.1: Summary of the proposed algorithm.

Step 1. Initialize $\boldsymbol{\alpha}^0$ by using an existing image registration method, then fix $\boldsymbol{\alpha} = \boldsymbol{\alpha}^0$, compute $\boldsymbol{f}^0$ by minimizing (4.2a). This is equivalent to solving the following:

$$\left(\boldsymbol{W}^T(\boldsymbol{\alpha}^0)\boldsymbol{W}^T(\boldsymbol{\alpha}^0) + \lambda \boldsymbol{T}(\varepsilon)\right)\boldsymbol{f}^0 = \boldsymbol{W}^T(\boldsymbol{\alpha}^0)\boldsymbol{g} \tag{4.5}$$

Step 2. Construct $\boldsymbol{W}_1$, $\vec{\boldsymbol{W}}(\boldsymbol{\alpha}^0)$, $\boldsymbol{J}(\boldsymbol{\alpha}^0, \boldsymbol{f}^0)$, $\boldsymbol{r}_1(\boldsymbol{f}^0)$ and $\vec{\boldsymbol{r}}(\boldsymbol{\alpha}^0, \boldsymbol{f}^0)$.

Step 3. At the i-th iteration

Solving (4.4) for $\Delta\boldsymbol{\alpha}^i$ and $\Delta\boldsymbol{f}^i$ based on the following equivalent equation:

$$\boldsymbol{P}^T\boldsymbol{P}\begin{pmatrix}\Delta\boldsymbol{\alpha}^i \\ \Delta\boldsymbol{f}^i\end{pmatrix} = \boldsymbol{P}^T\begin{pmatrix}\boldsymbol{r}_1(\boldsymbol{f}^i) \\ \vec{\boldsymbol{r}}(\boldsymbol{\alpha}^i, \boldsymbol{f}^i) \\ -\sqrt{\lambda}\boldsymbol{L}(\varepsilon)\boldsymbol{f}^i \\ \boldsymbol{0}\end{pmatrix}, \quad \text{where } \boldsymbol{P} = \begin{pmatrix}\boldsymbol{0} & \boldsymbol{W}_1 \\ \boldsymbol{J}(\boldsymbol{\alpha}^i, \boldsymbol{f}^i) & \vec{\boldsymbol{W}}(\boldsymbol{\alpha}^i) \\ \boldsymbol{0} & \sqrt{\lambda}\boldsymbol{L}(\varepsilon) \\ \boldsymbol{R} & \boldsymbol{0}\end{pmatrix} \tag{4.6}$$

Step 4. Update the estimates $\boldsymbol{\alpha}^{i+1}$ and $\boldsymbol{f}^{i+1}$ with:

$$\boldsymbol{\alpha}^{i+1} = \boldsymbol{\alpha}^i + \Delta\boldsymbol{\alpha}^i, \ \boldsymbol{f}^{i+1} = \boldsymbol{f}^i + \Delta\boldsymbol{f}^i.$$

Step 5. Impose the following condition:

$$-1 \le \cos\theta_i, \sin\theta_i \le 1, \ \ 0 \le f(x, y) \le 255.$$

Step 6. Update $\vec{\boldsymbol{W}}(\boldsymbol{\alpha}^{i+1})$, $\boldsymbol{J}(\boldsymbol{\alpha}^{i+1}, \boldsymbol{f}^{i+1})$, $\boldsymbol{r}_1(\boldsymbol{f}^{i+1})$ and $\vec{\boldsymbol{r}}(\boldsymbol{\alpha}^{i+1}, \boldsymbol{f}^{i+1})$.

Step 7. Go to Step 3 and update $i = i+1$ until convergence or a maximum number of iterations is reached.

During the initialization, the HR image is reconstructed by minimizing (4.2a) based on the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

initialized $\boldsymbol{\alpha}^0$. The minimization problem is equivalent to solving (4.5). The closed-form solution $\boldsymbol{f}^0$ in (4.5) requires inversion of the matrix $\boldsymbol{W}^T(\boldsymbol{\alpha}^0)\boldsymbol{W}^T(\boldsymbol{\alpha}^0)+\lambda\boldsymbol{T}(\varepsilon)$, which is computationally intensive. To solve this problem, a numerical approach using conjugate gradient (CG) optimization is adopted. The CG optimization method has been discussed in detail in Chapter 3. Similarly in step 3, it is difficult to invert the matrix $\boldsymbol{P}^T\boldsymbol{P}$. Hence, the CG optimization method is again adopted due to its fast convergence. In this work, the dimension of the unknown motion parametric vector $\boldsymbol{\alpha}$ is much smaller than the dimension of the unknown HR image $\boldsymbol{f}$. Hence, its computational cost is almost similar to the traditional SR algorithms where the estimated $\boldsymbol{\alpha}$ is considered to be accurate or known *a priori*.

As opposed to the current two-stage SR methods that perform registration on the LR images at the first stage, the image registration in the proposed method is performed iteratively using the progressively estimated HR image. This is promising as more accurate motion parameters can be determined, thereby enhancing the performance of the HR reconstruction. Further, the new method can overcome the shortcoming of the iterative AM framework as the cost function is not projected onto the image and the motion parameter domains one at a time. Instead, an iterative NLS-based scheme is developed to estimate the motion parameters and the HR image simultaneously and progressively. The theoretical justification for the proposed algorithm is presented in Appendix A of this chapter.

4.3.2 Derivation of the Jacobian Matrix

The main challenge in the development of the NLS method lies in the derivation and computation of the Jacobian matrix $\boldsymbol{J}(\boldsymbol{\alpha},\boldsymbol{f})$. $\boldsymbol{J}(\boldsymbol{\alpha},\boldsymbol{f})$ is the Jacobian of $\vec{\boldsymbol{W}}(\boldsymbol{\alpha})\boldsymbol{f}$ with respect to $\boldsymbol{\alpha}$. To minimize the cost function (4.4), $\boldsymbol{J}(\boldsymbol{\alpha},\boldsymbol{f})$ has to be constructed explicitly. It should be noted that due to the inclusion of rotational motion, the motion operator $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)$ is no longer

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Toeplitz. Thus, $\vec{\boldsymbol{W}}(\boldsymbol{\alpha})\boldsymbol{f}$ is now nonlinear with respect to $\boldsymbol{\alpha}$. In other words, we cannot find an equivalent matrix such that $\vec{\boldsymbol{W}}(\boldsymbol{\alpha})\boldsymbol{f} = \boldsymbol{X}\boldsymbol{a}$. Therefore, we propose the following technique to solve the problem.

As each LR observation is independent, $\boldsymbol{J}(\boldsymbol{\alpha}, \boldsymbol{f})$ can be written as:

$$\boldsymbol{J}(\boldsymbol{\alpha}, \boldsymbol{f}) = \begin{bmatrix} \boldsymbol{J}_2(\boldsymbol{\alpha}_2, \boldsymbol{f}) & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{J}_N(\boldsymbol{\alpha}_N, \boldsymbol{f}) \end{bmatrix} \tag{4.7}$$

where $\boldsymbol{J}_k(\boldsymbol{\alpha}_k, \boldsymbol{f})$ is the Jacobian of $\boldsymbol{D}\boldsymbol{H}_k\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}$ with respect to $\boldsymbol{\alpha}_k$ $(2 \le k \le N)$. As the first LR image is used as the reference, we only need to estimate the unknown motion parameters for $N-1$ images. Using the chain rule, $\boldsymbol{J}_k(\boldsymbol{\alpha}_k, \boldsymbol{f})$ can be expressed as follows:

$$\begin{aligned} \boldsymbol{J}_k(\boldsymbol{\alpha}_k, \boldsymbol{f}) &= \frac{\partial\left(\boldsymbol{D}\boldsymbol{H}_k\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}\right)}{\partial \boldsymbol{\alpha}_k} \\ &= \frac{\partial(\boldsymbol{D}\boldsymbol{H}_k\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f})}{\partial\left(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}\right)} \frac{\partial\left(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}\right)}{\partial \boldsymbol{\alpha}_k} \\ &= \boldsymbol{D}\boldsymbol{H}_k \frac{\partial\left(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}\right)}{\partial \boldsymbol{\alpha}_k} \end{aligned} \tag{4.8}$$

It is not easy to obtain $\partial\left(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}\right)/\partial\boldsymbol{\alpha}_k$ directly since $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}$ is nonlinear with respect to $\boldsymbol{\alpha}_k$. To solve the problem, we propose to use bilinear interpolation to derive the relationship between $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}$ and $\boldsymbol{\alpha}_k$.

We first introduce the formation of relative shifted HR images $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}$. Fig. 4.1 shows the four possible positions between the pixels in the shifted HR grid and the pixels in the reference HR grid. The images with HR grid (HR images) correspond to the observed LR images. We define (x_1, y_1) as the coordinates of the reference HR grid and (x_k, y_k) as the coordinates of the k-th shifted HR grid. $(\Delta x_k, \Delta y_k)$, the coordinate change, is defined as:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\begin{aligned}\Delta x_k &= x_k - x_1 = x_1\cos\theta_k - y_1\sin\theta_k + s_{xk} - x_1;\\ \Delta y_k &= y_k - y_1 = x_1\sin\theta_k + y_1\cos\theta_k + s_{yk} - y_1\end{aligned} \tag{4.9}$$

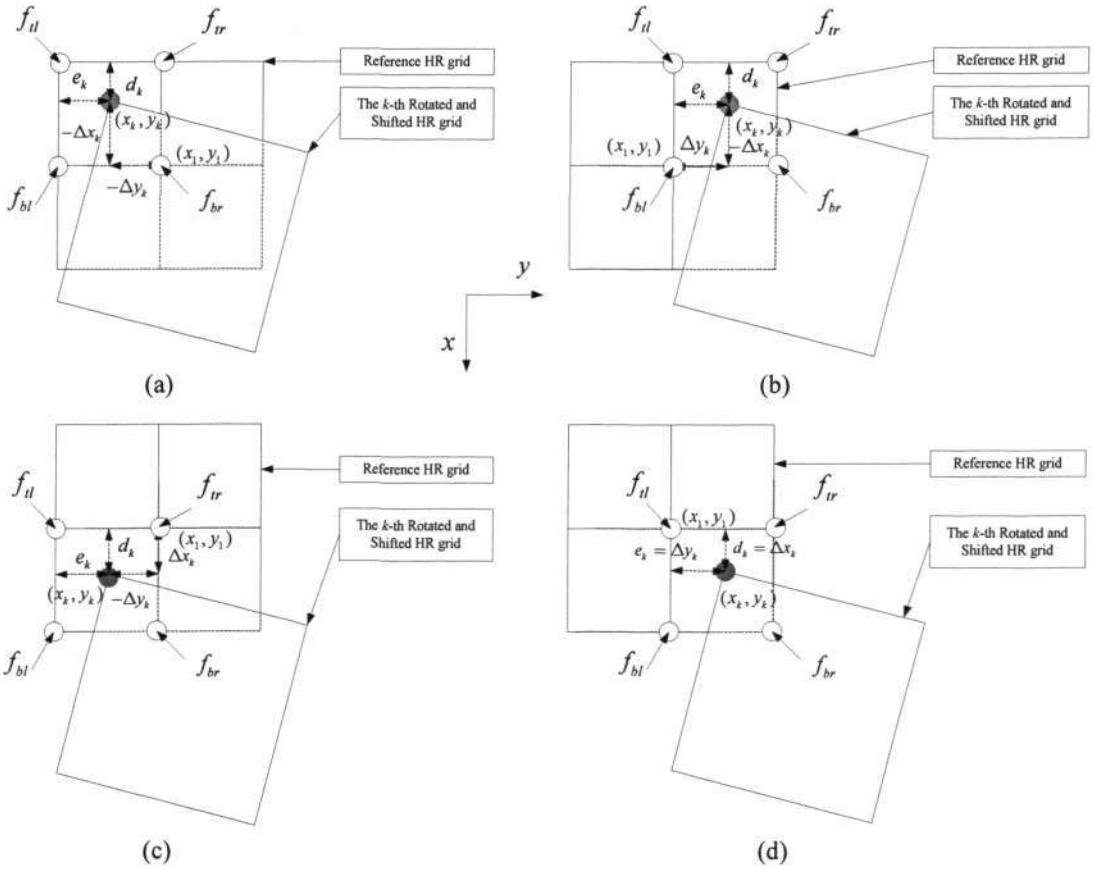


Fig. 4.1: Four possible relative positions between (x_k, y_k) and (x_1, y_1): (a)-(d) denote the pixel (x_k, y_k) in the k-th HR grid with respect to (x_1, y_1) in the reference HR grid.

From Fig. 4.1, it can be seen that there are four possible relative coordinate positions between (x_k, y_k) and (x_1, y_1). The shifted pixels (x_k, y_k) in the k-th HR grid with respect to the top-left, top-right, bottom-left and bottom-right pixels in the reference HR grid are shown. It is noted that $(\Delta x_k, \Delta y_k)$ may consist of both pixel-level and subpixel-level shifts. For simplicity of demonstration, the HR grids in Fig. 4.1 show the subpixel shift only. The pixel-level shift can be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

taken into consideration by the floor(.) operator in (4.12). Using the vector-matrix form, we can express (4.9) as:

$$\begin{bmatrix}\Delta \boldsymbol{x}_k \\ \Delta \boldsymbol{y}_k\end{bmatrix} = \begin{bmatrix}\boldsymbol{x}_k \\ \boldsymbol{y}_k\end{bmatrix} - \begin{bmatrix}\boldsymbol{x}_1 \\ \boldsymbol{y}_1\end{bmatrix} = \boldsymbol{C}\boldsymbol{a}_k - \begin{bmatrix}\boldsymbol{x}_1 \\ \boldsymbol{y}_1\end{bmatrix}$$
$$\text{where } \boldsymbol{C} = \begin{bmatrix}\boldsymbol{x}_1 & -\boldsymbol{y}_1 & \boldsymbol{1} & \boldsymbol{0} \\ \boldsymbol{y}_1 & \boldsymbol{x}_1 & \boldsymbol{0} & \boldsymbol{1}\end{bmatrix} \tag{4.10}$$

where $\Delta \boldsymbol{x}_k$, $\Delta \boldsymbol{y}_k$, $\boldsymbol{x}_k$, $\boldsymbol{y}_k$, $\boldsymbol{x}_1$ and $\boldsymbol{y}_1$ are the vectors representing the discrete and lexicographically ordered displacement vectors Δx_k, Δy_k, x_k, y_k, x_1 and y_1, respectively. $\boldsymbol{1}$ and $\boldsymbol{0}$ are vectors of all ones and zeros, respectively. From (4.10), it is clear that $\left[\Delta \boldsymbol{x}_k^T, \Delta \boldsymbol{y}_k^T\right]^T$ is linear with respect to $\boldsymbol{a}_k$. $\partial\left(\left[\Delta \boldsymbol{x}_k^T, \Delta \boldsymbol{y}_k^T\right]^T\right) \Big/ \partial \boldsymbol{a}_k$ can be obtained as:

$$\frac{\partial\left(\left[\Delta \boldsymbol{x}_k^T, \Delta \boldsymbol{y}_k^T\right]^T\right)}{\partial \boldsymbol{a}_k} = \frac{\partial(\boldsymbol{C}\boldsymbol{a}_k)}{\partial \boldsymbol{a}_k} = \boldsymbol{C} \tag{4.11}$$

Next, we use bilinear interpolation to describe the dependency between $\boldsymbol{S}_k(\boldsymbol{a}_k)\boldsymbol{f}$ and $\left[\Delta \boldsymbol{x}_k^T, \Delta \boldsymbol{y}_k^T\right]^T$. The aim is to determine $\partial\left(\boldsymbol{S}_k(\boldsymbol{a}_k)\boldsymbol{f}\right) \Big/ \partial\left(\left[\Delta \boldsymbol{x}_k^T, \Delta \boldsymbol{y}_k^T\right]^T\right)$. The pixel (x_k, y_k) in the k-th HR grid is determined by four neighboring pixel values f_{tl}, f_{tr}, f_{bl} and f_{br} as shown in Fig. 4.1. These four points are the reference HR grid points surrounding the pixel (x_k, y_k) at the top-left, top-right, bottom-left and bottom-right locations, respectively. We denote $[d_k, e_k]^T$ as the distance vector between the pixel (x_k, y_k) and the pixel at the top-left position in the reference HR grid. When examining the four possible coordinate positions in Figs. 4.1 (a)-(d), d_k and e_k can be seen to satisfy the following relationships:

$$(a)-(d): \begin{cases} d_k = \Delta x_k - \text{floor}(\Delta x_k); \\ e_k = \Delta y_k - \text{floor}(\Delta y_k) \end{cases} \tag{4.12}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

where floor(.) denotes the operator rounding the number to the nearest integer less than or equal to itself. Using bilinear interpolation, the shifted and rotated $\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}$ can be obtained by:

$$\begin{aligned}\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f} = {} & \boldsymbol{d}_k \odot (\boldsymbol{1}-\boldsymbol{e}_k)\odot \boldsymbol{f}_{bl} + \boldsymbol{d}_k \odot \boldsymbol{e}_k \odot \boldsymbol{f}_{br} + \\ & (\boldsymbol{1}-\boldsymbol{d}_k)\odot(\boldsymbol{1}-\boldsymbol{e}_k)\odot \boldsymbol{f}_{tl} + (\boldsymbol{1}-\boldsymbol{d}_k)\odot \boldsymbol{e}_k \odot \boldsymbol{f}_{tr}\end{aligned} \tag{4.13}$$

where $\odot$ is an entry-by-entry multiplication operator. $\boldsymbol{f}_{tl}, \boldsymbol{f}_{tr}, \boldsymbol{f}_{bl}, \boldsymbol{f}_{br}, \boldsymbol{d}_k$ and $\boldsymbol{e}_k$ are the vectors representing the lexicographically ordered $f_{tl}, f_{tr}, f_{bl}, f_{br}, d_k$ and e_k, respectively. For simplicity, we will use $\boldsymbol{B}_k$ to denote $\partial(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f}) \big/ \partial\left(\left[\Delta\boldsymbol{x}_k^T, \Delta\boldsymbol{y}_k^T\right]^T\right)$ for the rest of the derivation. Using (4.12) and (4.13), $\boldsymbol{B}_k$ can then be written as:

$$\begin{aligned}\boldsymbol{B}_k &= \frac{\partial(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f})}{\partial\left(\left[\boldsymbol{d}_k^T, \boldsymbol{e}_k^T\right]^T\right)} \frac{\partial\left(\left[\boldsymbol{d}_k^T, \boldsymbol{e}_k^T\right]^T\right)}{\partial\left(\left[\Delta\boldsymbol{x}_k^T, \Delta\boldsymbol{y}_k^T\right]^T\right)} \\ &= \left[diag\left\{(\boldsymbol{1}-\boldsymbol{e}_k)\odot(\boldsymbol{f}_{bl}-\boldsymbol{f}_{tl}) + \boldsymbol{e}_k \odot (\boldsymbol{f}_{br}-\boldsymbol{f}_{tr})\right\},\right. \\ &\quad \left. diag\left\{(\boldsymbol{1}-\boldsymbol{d}_k)\odot(\boldsymbol{f}_{tr}-\boldsymbol{f}_{tl}) + \boldsymbol{d}_k \odot (\boldsymbol{f}_{br}-\boldsymbol{f}_{bl})\right\}\right]\end{aligned} \tag{4.14}$$

as $\partial\left(\left[\boldsymbol{d}_k^T, \boldsymbol{e}_k^T\right]^T\right) \big/ \partial\left(\left[\Delta\boldsymbol{x}_k^T, \Delta\boldsymbol{y}_k^T\right]^T\right)$ equals to the identity matrix due to (4.12). It can be shown that the same equation (4.14) can be obtained for all the cases in Figs. 4.1 (a)-(d). Combining (4.11) and (4.14), $\partial(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f})/\partial\boldsymbol{\alpha}_k$ can then be obtained as:

$$\frac{\partial(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f})}{\partial\boldsymbol{\alpha}_k} = \frac{\partial(\boldsymbol{S}_k(\boldsymbol{\alpha}_k)\boldsymbol{f})}{\partial\left(\left[\Delta\boldsymbol{x}_k^T, \Delta\boldsymbol{y}_k^T\right]^T\right)} \frac{\partial\left(\left[\Delta\boldsymbol{x}_k^T, \Delta\boldsymbol{y}_k^T\right]^T\right)}{\partial\boldsymbol{\alpha}_k} = \boldsymbol{B}_k\boldsymbol{C} \tag{4.15}$$

Substituting (4.15) into (4.8), we can obtain a simple expression for $\boldsymbol{J}_k(\boldsymbol{\alpha}_k, \boldsymbol{f})$ which is $\boldsymbol{D}\boldsymbol{H}_k\boldsymbol{B}_k\boldsymbol{C}$. The Jacobian matrix in (4.7) can be finally integrated into the scheme iteratively as highlighted in Table 4.1.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\boldsymbol{A}\begin{pmatrix}\Delta\boldsymbol{\alpha}\\ \Delta\boldsymbol{f}\end{pmatrix}=\boldsymbol{b} \tag{4.16}$$

where

$$\boldsymbol{A}=\boldsymbol{P}^T\boldsymbol{P}=\begin{pmatrix}\boldsymbol{J}^T\boldsymbol{J}+\boldsymbol{R}^2 & \boldsymbol{J}^T\vec{\boldsymbol{W}}\\ \vec{\boldsymbol{W}}^T\boldsymbol{J} & \boldsymbol{W}_1^T\boldsymbol{W}_1+\vec{\boldsymbol{W}}^T\vec{\boldsymbol{W}}+\lambda\boldsymbol{T}\end{pmatrix}$$

$$=\begin{pmatrix}\boldsymbol{C}^T\boldsymbol{B}_2^T\boldsymbol{H}_2^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_2\boldsymbol{B}_2\boldsymbol{C}+\beta\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}^T\boldsymbol{B}_2^T\boldsymbol{H}_2^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_2\boldsymbol{S}_2\\ \boldsymbol{0} & \ddots & \boldsymbol{0} & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}^T\boldsymbol{B}_N^T\boldsymbol{H}_N^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_N\boldsymbol{B}_N\boldsymbol{C}+\beta\boldsymbol{I} & \boldsymbol{C}^T\boldsymbol{B}_N^T\boldsymbol{H}_N^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_N\boldsymbol{S}_N\\ \boldsymbol{S}_2^T\boldsymbol{H}_2^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_2\boldsymbol{B}_2\boldsymbol{C} & \cdots & \boldsymbol{S}_N^T\boldsymbol{H}_N^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_N\boldsymbol{B}_N\boldsymbol{C} & \boldsymbol{H}_1^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_1+\sum_{k=2}^{N}\boldsymbol{S}_k^T\boldsymbol{H}_k^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{H}_k\boldsymbol{S}_k+\lambda\boldsymbol{T}\end{pmatrix}$$

$$\boldsymbol{b}=\begin{pmatrix}\boldsymbol{J}^T\vec{\boldsymbol{r}}\\ \boldsymbol{W}_1^T\boldsymbol{r}_1+\vec{\boldsymbol{W}}^T\vec{\boldsymbol{r}}-\lambda\boldsymbol{T}\boldsymbol{f}\end{pmatrix}=\begin{pmatrix}\boldsymbol{C}^T\boldsymbol{B}_2^T\boldsymbol{H}_2^T\boldsymbol{D}^T(\boldsymbol{g}_2-\boldsymbol{D}\boldsymbol{H}_2\boldsymbol{S}_2\boldsymbol{f})\\ \vdots\\ \boldsymbol{C}^T\boldsymbol{B}_N^T\boldsymbol{H}_N^T\boldsymbol{D}^T(\boldsymbol{g}_N-\boldsymbol{D}\boldsymbol{H}_N\boldsymbol{S}_N\boldsymbol{f})\\ \boldsymbol{H}_1^T\boldsymbol{D}^T(\boldsymbol{g}_1-\boldsymbol{D}\boldsymbol{H}_1\boldsymbol{f})+\sum_{k=2}^{N}\boldsymbol{S}_k^T\boldsymbol{H}_k^T\boldsymbol{D}^T(\boldsymbol{g}_k-\boldsymbol{D}\boldsymbol{H}_k\boldsymbol{S}_k\boldsymbol{f})-\lambda\boldsymbol{T}\boldsymbol{f}\end{pmatrix}$$

Finally, we can formulate equation (4.6) in Table 4.1 by using the explicit equations in (4.16) to compute the increment $(\Delta\boldsymbol{\alpha},\Delta\boldsymbol{f})$. It should be noted that the matrix $\boldsymbol{A}=\boldsymbol{P}^T\boldsymbol{P}$ in (4.16) is a $\left[(4(N-1)+M_f\times N_f)\times(4(N-1)+M_f\times N_f)\right]$ sparse matrix and the vector $\boldsymbol{b}$ in (4.16) is $[(4(N-1)+M_f\times N_f)\times 1]$. To compute the closed form solution for this equation, we need to invert the matrix $\boldsymbol{A}$, which is computationally intensive. To address this problem, we employ the CG method as it can achieve fast convergence when compared with the other methods. After solving (4.6) in step 3 of Table 4.1, the estimates $\boldsymbol{\alpha}$ and $\boldsymbol{f}$ can then be updated using the new increment $\Delta\boldsymbol{\alpha}$ and $\Delta\boldsymbol{f}$. The algorithm will terminate until convergence or a maximum number of iterations is reached.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.4 Experiments

In this section, we study the performance of the proposed method and compare it with other methods. We conducted various Monte-Carlo simulations to compare the results obtained using the proposed method with those of the two other SR algorithms, namely two-stage disjoint SR method similar to [4] and AM SR method similar to [31]. Finally, experiments using real-life images are presented to illustrate the effectiveness of the proposed method. To evaluate the performance, we used the following well-known metrics in this work: NMSE for the identified motion parametric vector and the PSNR for the reconstructed HR image. The definition of the NMSE for the identified motion parametric vector is given as follows:

$$NMSE(\hat{\boldsymbol{\alpha}}) = 100\frac{\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|^2}{\|\boldsymbol{\alpha}\|^2} \tag{4.17}$$

where '^' denotes the currently computed estimate. Generally, a good algorithm is reflected by low NMSE and high PSNR. Nevertheless, the best performance measure remains human inspection of the reconstructed HR images.

4.4.1 Convergence Study

In this section, we conducted various experiments to study the effect of the initial choices of motion parameters on the convergence of the proposed method. We conducted the experiments using the 90×90 "Building" image, shown in Fig. 4.2. To generate N=5 LR images for each HR image, the HR image was rotated by different randomly chosen angles of (0, 1.67, 9.42, −6.59, −2.74) degrees, shifted by randomly chosen translational shifts of (0, 0), (10.75, 5.86), (−14.42, −5.78), (1.27, 7.88), (−8.27, −11.25) pixels, and blurred by 2×2 uniform kernel before subsampling by a decimation factor of 2. The images were further degraded by AWGN to produce a signal-to-noise ratio (SNR) at 35 dB. The convergence performance was evaluated

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

using different initial motion parameters. We denote $\theta_{initial}$ and θ_{true} as the initial estimate and true value for the rotation angle θ, respectively. $s_{initial}$ and s_{true} represent the initial estimate and true value for the shift s. s^p_{true} and s^s_{true} are denoted as the pixel- and subpixel-level of s_{true}, respectively. Then the motion parameters are initialized using the following:

$$\begin{aligned} \theta_{initial} &= \theta_{true} \pm \theta_{error}; \\ s_{initial} &= s^p_{true} + s^s_{true} \times (1 \pm e_s) \end{aligned} \tag{4.18}$$

where θ_{error} denotes the estimated angle error, and e_s represents the percentage of the error in s^s_{true}. Using the existing registration algorithms such as [20], it is observed that θ_{error} is usually small and falls within the range of $-2 \le \theta_{error} \le 2$ degrees. Therefore, we considered the worst-case scenario and fixed $\theta_{error} = 2$ degrees in these experiments. Previous studies [20], [149], [150] have shown that existing registration methods can achieve accurate pixel-level registration. Therefore, the registration error is considered mainly coming from the inaccurate subpixel registration. In this study, four values of $e_s = 50\%, 150\%, 250\%, 350\%$ were used as the errors in the initial estimated s^s_{true}. It is noted that we randomly chose '+' or '−' in (4.18) for each LR image. In the experiment, we considered that the algorithm had converged if the following convergence criterion was satisfied:

$$\frac{\| \boldsymbol{f}^i - \boldsymbol{f}^{i-1} \|^2}{\| \boldsymbol{f}^{i-1} \|^2} < 10^{-6} \tag{4.19}$$

We plotted the PSNR of the reconstructed HR images for our algorithm against the number of iterations in Fig. 4.3. It is observed that the experiments with these initial conditions all converge to the same solution. When the initial estimates are further away from the true values, the algorithm requires more iterations to converge. This provides empirical evidence on the convergence of the proposed method and an indication on the tightness of the initial estimates

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

that are required to achieve convergence. In order to further illustrate the registration performance, the NMSE of the estimated motion parametric vector is plotted in Fig. 4.4. From the figures, it is observed that the motion parameters with different initial conditions converge to the true value. These results further reconfirm the PSNR of the reconstructed HR images in Fig. 4.3.

Fig. 4.2: Original "Building" HR image.

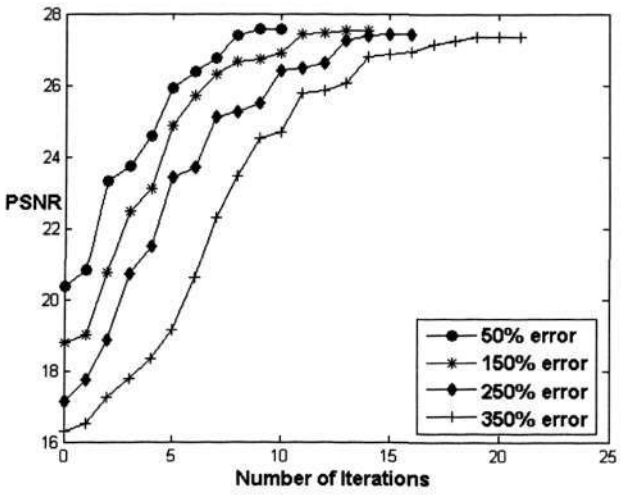


Fig. 4.3: PSNR of the reconstructed HR images with different initial conditions.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

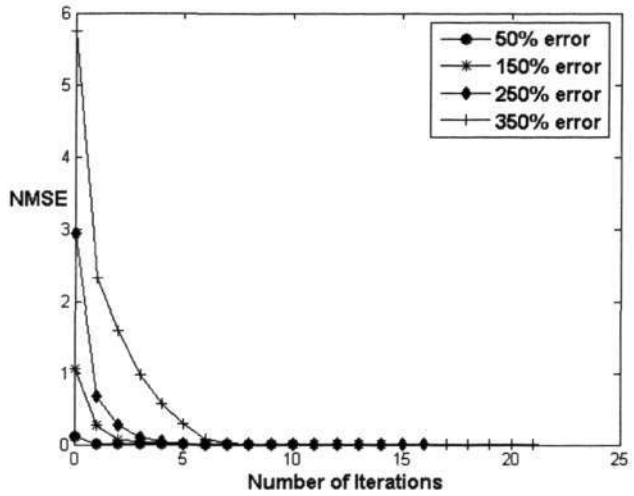


Fig. 4.4: NMSE of the estimated motion parametric vectors with different initial conditions.

4.4.2 Super-resolution on Simulated Images with Additive Noise

In this section, we conducted various Monte-Carlo simulations to perform image SR based on multiple noisy LR images at different noise levels. The number of Monte-Carlo simulations was set to 10 at different noise levels. A section of "Building" image in Fig. 4.7 (a) was selected as the test image. To generate 5 different LR images in each simulation, the HR image was rotated by randomly selected angles from a uniform distribution over [-20, 20] degrees, shifted by randomly selected subpixel translations from a uniform distribution over [0, 1], and blurred by 2×2 uniform kernel before subsampling by a decimation factor of 2. The LR images were further degraded by AWGN to produce different levels of SNR (45, 35, 25, 15 dB). Two other SR algorithms, namely two-stage disjoint SR method similar to [4] and AM SR method similar to [31], together with the proposed method were applied to the observed LR images to perform image SR. We initialized the motion parameters by using the registration method in [20] for all three methods. The two-stage disjoint SR was implemented by minimizing the cost function in (4.2a) with respect to the HR image. The AM method was implemented by projecting and

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

minimizing the cost function in (4.2a) with respect to the HR image and the motion parametric vector iteratively. The state-of-the-art AM method that is based on CG optimization is used for comparison in this study. The proposed method was run based on the procedure outlined in Table 4.1.

The PSNR of the reconstructed HR images for different SNR noise levels is given in Fig. 4.5. The PSNR obtained are based on 10 Monte-Carlo simulations. From the figure, it is observed that the proposed method produces the best results among the three methods. The two-stage disjoint SR method has the lowest PSNR, as it does not take the impact of inaccurate initial registration into consideration. The proposed method, on the other hand, offers better results than the AM method as the HR image and the motion parameters are estimated simultaneously in the proposed method. This is as opposed to the AM method where the cost function is projected onto the HR image and motion parameter domains, and minimized iteratively. To further compare the registration performance of the three methods, the NMSE of the estimated motion parametric vector for different SNR levels is plotted in Fig. 4.6. From the figure, it is observed that the proposed method achieves the best estimated motion parameters. This observation further reconfirms the PSNR results given in Fig. 4.5.

We also compared the reconstructed HR images of an experiment at 35dB noise level using the proposed and AM methods. The results are given in Fig. 4.7. From the figures, it can be seen that the reconstructed HR images using the AM method contain more artifacts than those of the proposed method. Further, Figs. 4.8 and 4.9 show that the PSNR of the reconstructed HR image and the NMSE of the estimated motion parametric vector using the proposed method are superior to the AM method.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

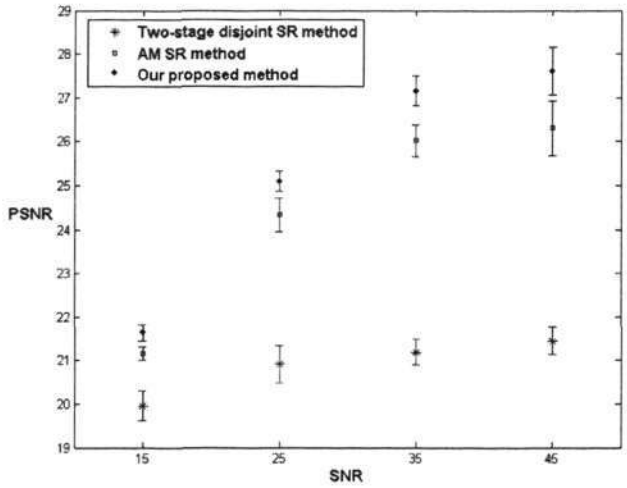


Fig. 4.5: PSNR of the reconstructed HR image against the SNR of the LR images.

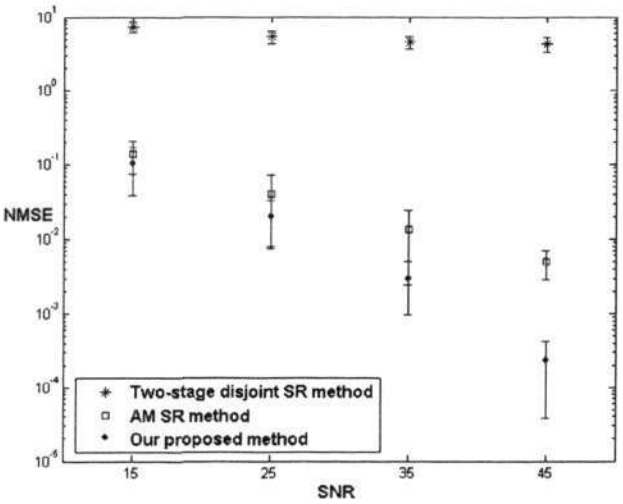


Fig. 4.6: NMSE of the estimated motion parametric vector against the SNR of the LR images.

Finally, we will provide an analysis on the computational complexity of the proposed method. As the exact number of arithmetic operations required is difficult to determine for the proposed method, we will use computational time to show the complexity of the proposed method. Further, we have used the AM SR method as the benchmark for comparison. In order

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

to provide a good indication of the computational time, the above Monte-Carlo simulation experiment at SNR=35dB is used. The simulation environments are given as follows: Windows XP, MATLAB 7.1, CPU P4-3.4GHz, and 1G RAM. The total computational time, averaged over 10 runs, is 37.7s for the proposed method, and 45.9s for the AM method. This shows that the proposed method has better computational time. It is noted that although an average iteration in our algorithm requires more computation than that of the AM method, the total number of iterations required for the proposed method is much less than that of the AM method. This observation can be shown in Figs. 4.8 and 4.9 where the proposed method requires much smaller number of iterations to achieve convergence. It should also be noted that the proposed method is implemented in MATLAB for these Monte-Carlo simulations. MATLAB is an interpreter-programming platform with lower processing speed. The computational time of the experiments can be improved if the algorithm is implemented in a compiler language such as C.

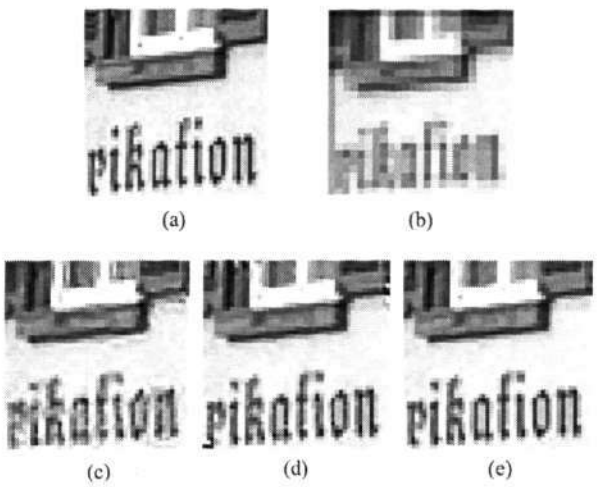


Fig. 4.7: SR on LR images ("Building") degraded by AWGN (a) Original HR image, (b) A sample of the scaled-up LR images, (c) Reconstructed image using the two-stage SR algorithm [4], (d) Reconstructed image using the AM method [31], (e) Reconstructed image using the proposed algorithm.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

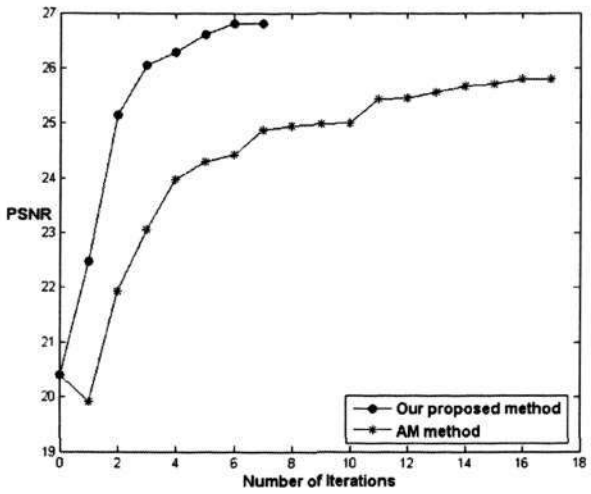


Fig. 4.8: PSNR of the reconstructed HR image.

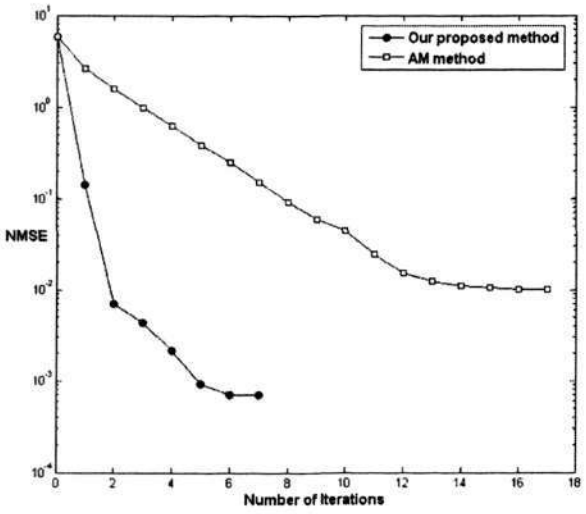
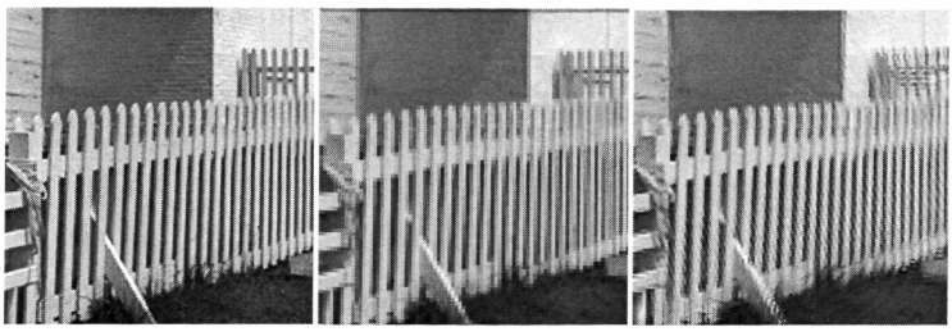


Fig. 4.9: NMSE of the estimated motion parametric vector.

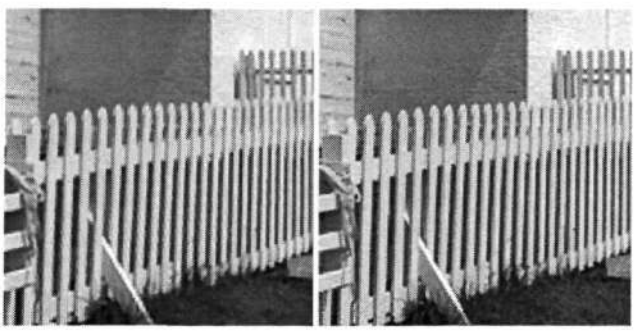
To further show the performance of the proposed method, we used a section of "lighthouse" image as the test image. The experimental setting is similar to that in previous one. To generate N=5 LR images, the HR image in Fig. 4.10 (a) was rotated by different randomly chosen angles

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

of (0, −17.23, −0.03, −15.72, 2.87) degrees, shifted by randomly chosen translational shifts of (0, 0), (0.83, 0.18), (0.14, 0.68), (0.21, 0.46), (0.38, 0.97) pixels, and blurred by 2×2 uniform kernel before subsampling by a decimation factor of ρ=2. The images were further degraded by AWGN to produce a SNR of 35 dB. The two-stage disjoint SR method algorithm [4], the AM method [31] and the proposed method were run to perform SR image reconstruction. The reconstructed HR images using the three methods are given in Figs. 4.10 (c)-(e), respectively. It can be seen that the result by the proposed method has better quality and sharper edges e.g. near the fence, when compared with those of the other two methods.



(a) (b) (c)



(d) (e)

Fig. 4.10: SR on LR images ("Lighthouse") degraded by AWGN (a) Original HR image, (b) A sample of the scaled-up LR images, (c) Reconstructed image using the two-stage SR algorithm [4], (d) Reconstructed image using the AM method [31], (e) Reconstructed image using the proposed algorithm.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.4.3 Super-resolution on Real-life Images

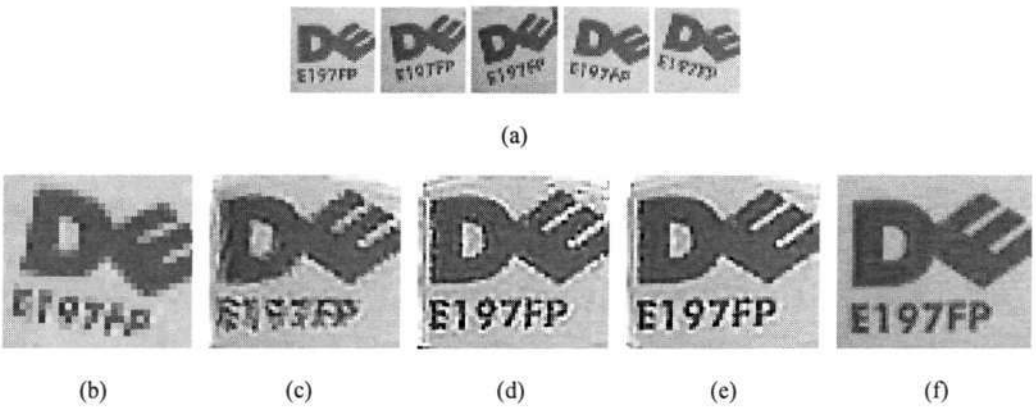
The real-life experiment was conducted by capturing five "trademark" images using a web camera with relative translations and rotations. The LR images are shown in Fig. 4.11(a) and a sample of the scaled-up LR images is shown in Fig. 4.11(b). We chose a decimation factor of 2 so that the estimated HR image will have twice the resolution of the LR images. The registration algorithm in [20] was again used to estimate the initial shifts and rotations among the LR images. Next the two-stage SR algorithm, the AM method and the proposed method were run to perform SR image reconstruction. The reconstructed HR images using the three methods are given in Figs. 4.11(c)-(e), respectively. The estimated translational motions by using the proposed algorithm are given as in Table 4.2. To provide a fair comparison, an image with the resolution of the HR image was captured as the ground truth in Fig. 4.11(f). From Figs. 4.11(c)-(e), it is observed that the considerable clarity of images has been recovered by both the AM and the proposed methods. Further, it can be seen that the result by the proposed method has less artifacts than that of the AM method, in particular near the edge of the words. Comparison reveals that our approach is superior in handling real-life image SR, as it is able to estimate motion parameters accurately, leading to superior HR image reconstruction.

4.5 Summary

This chapter presents a new algorithm to integrate image registration into image SR. As opposed to the methods that treat image registration and HR reconstruction as disjoint processes, the new framework enables image registration and HR reconstruction to be estimated and improved progressively. Further, unlike most algorithms that focus on translational motion model, the proposed framework adopts a more generic motion model that includes both translation as well as rotation. An iterative scheme is developed to solve the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

arising nonlinear least squares problem in order to perform simultaneous image registration and HR reconstruction. Experimental results show that the proposed method is effective in performing image registration and SR for simulated as well as real-life images.



(a)

(b) (c) (d) (e) (f)

Fig. 4.11: SR on real-life images. (a) Five LR images, (b) A sample of the scaled-up LR images, (c) Reconstructed image using the two-stage SR algorithm [4]. (d) Reconstructed image using the AM algorithm [31]. (e) Reconstructed image using the proposed algorithm. (f) Ground truth.

Appendix A: Justification of the Algorithm

In this Appendix, a brief theoretical justification for the algorithm is presented. As discussed in Section 4.3, our joint framework for simultaneous image registration and SR can be stated in terms of minimizing the following differentiable function:

$$A(\boldsymbol{\alpha}, \boldsymbol{f}) = \frac{1}{2}\|\boldsymbol{r}_1(\boldsymbol{f})\|^2 + \frac{1}{2}\|\bar{\boldsymbol{r}}(\boldsymbol{\alpha}, \boldsymbol{f})\|^2 + \frac{1}{2}\lambda \boldsymbol{f}^T \boldsymbol{T}(\varepsilon)\boldsymbol{f} \tag{4A.1}$$

The solution of the minimization problem can be determined by solving $(\partial A/\partial \boldsymbol{\alpha}) = (\partial A/\partial \boldsymbol{f}) = 0$. This condition can be expressed as:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\begin{aligned} \partial A/\partial \boldsymbol{a} &= -\boldsymbol{J}^T(\boldsymbol{a},\boldsymbol{f})\vec{\boldsymbol{r}}(\boldsymbol{a},\boldsymbol{f}) = 0 \\ \partial A/\partial \boldsymbol{f} &= -\boldsymbol{W}_1^T\boldsymbol{r}_1(\boldsymbol{f}) - \vec{\boldsymbol{W}}^T(\boldsymbol{a})\vec{\boldsymbol{r}}(\boldsymbol{a},\boldsymbol{f}) + \lambda\boldsymbol{T}(\varepsilon)\boldsymbol{f} = 0 \end{aligned} \tag{4A.2}$$

The aim is to show that the proposed method can satisfy this condition when the convergence is reached.

In Section 4.3, the original problem for direct estimation of $\boldsymbol{f}$ and $\boldsymbol{a}$ has been transformed into finding the increment $\Delta\boldsymbol{f}$ and $\Delta\boldsymbol{a}$ by minimizing (4.4). We first show that the minimization problem is equivalent to solving the equation in (4.6). It should be noted that the sufficient condition for this equivalence is that the Hessian matrix of the minimized cost function is positive definite [6]. As the Hessian matrix is $\boldsymbol{A} = \boldsymbol{P}^T\boldsymbol{P}$, it can be shown that $\boldsymbol{A}$ is positive definite. Thus, we can obtain the minimum of (4.4) by solving (4.6) in our method. Further, when the convergence of the algorithm is reached, the vector $\left[\Delta\boldsymbol{a}^T, \Delta\boldsymbol{f}^T\right]$ will be zero, if and only if, $\boldsymbol{b}$ in (4.16) vanishes. Combining the expression in (4.16) and (4A.2), the following can be obtained at the convergence of the algorithm:

$$b = -\begin{pmatrix} \partial A/\partial \boldsymbol{a} \\ \partial A/\partial \boldsymbol{f} \end{pmatrix} = \boldsymbol{0} \tag{4A.3}$$

This means that the solution satisfying (4A.2) is obtained when the iterative algorithm reaches convergence. Experimental results in Section 4.4 support this theoretical analysis.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 5

Learning-based Single-frame Face Super-resolution

5.1 Introduction

Chapters 3 and 4 have presented the techniques for multi-frame SR, which is the fusion of a number of LR images to produce a HR image. In this chapter, we will investigate the learning-based single-frame SR algorithm that estimates missing HR details from a single observed LR image by learning from a set of training examples. It is worth mentioning that the multi-frame SR recovers high frequencies in the reconstructed HR image, as there is new information available. The new information comes from multiple observed LR images, each exhibiting a different aliasing effect. However, for the single-frame SR problem, only one image is given. Hence, an alternative source of new information is necessary, so as to compensate for the lack of observed data. In this case, a large training set of HR images could be employed as such source of new information. Then the missing HR details of the observed image are learned from the training set. This leads to the concept of learning-based single-frame SR.

It is noted that the use of training examples becomes much more effective when handling domain-specific images [151], such as text or face images. This is because the target image may share more characteristics with the training examples that will be helpful for SR reconstruction.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

In this thesis, we focus our attention on the SR of human face images. In many imaging applications, of particular interest is to render a HR face image from limited LR observations. This can be especially useful in a surveillance system where the resolution of a face image may be low in poor video, and insufficient useful observation, in the extreme case, only a single face observation is available. Hence, this motivates the study of learning-based single-frame face SR, which aims to estimate the missing HR facial features from a single LR face image by using a database of HR face images.

This chapter proposes a new learning-based single-frame face SR method which uses the techniques: principle component analysis (PCA) and locally linear embedding (LLE) [81]. A common issue in conventional methods is that, when the given LR face is a new face significantly different from those in the training set, the obtained result is less than satisfactory. To alleviate this difficulty, we develop an improved PCA-based method to render a global HR face by incorporating an estimated initial HR face in the training set. The term "*global*" is used here as it denotes that the reconstructed HR face is able to maintain global characteristics of a human face. In other words, the global face is in agreement with ordinary facial features, e.g. eyes, mouth and nose. To further enhance the detailed local information, we develop a patch-based residue prediction approach by using a manifold learning method named LLE. Experimental results in later sections show that the proposed method is effective in performing single-frame face SR.

The rest of this chapter is organized as follows. The problem formulation of single-frame face SR is introduced in Section 5.2. A new learning-based algorithm using the techniques of PCA and LLE is developed in Section 5.3. Experimental results on simulated and real-life images are presented and discussed in Section 5.4. A brief summary is given in Section 5.5.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.2 Problem Formulation

We formulate single-frame face SR problem as follows:

$$\boldsymbol{g} = \boldsymbol{W}\boldsymbol{f} + \boldsymbol{n}, \text{ where } \boldsymbol{W} = \boldsymbol{D}\boldsymbol{H} \tag{5.1}$$

where $\boldsymbol{g}, \boldsymbol{f}$ and $\boldsymbol{n}$ are the vectors representing the discrete and lexicographically ordered acquired LR face image g, original HR face image f and additive noise n, respectively. $\boldsymbol{D}$ and $\boldsymbol{H}$ denote the downsampling and blurring operators, respectively. The problem of solving for $\boldsymbol{f}$ in (5.1) is that the system is underdetermined, and hence has infinitely many solutions. To tackle this problem, we should constrain the solution space according to some prior knowledge. Based on the idea of the learning-based SR methods, the prior knowledge can be learned from the training examples for the SR reconstruction. It is worth mentioning that the face SR has its unique challenges when compared with the SR for generic images, because people are familiar with facial features. A small error in a reconstructed HR face might be unacceptable to human perception, such as asymmetry of the eyes or noses, whereas the errors in textured regions of a generic reconstructed HR image usually go unnoticed. Therefore, the single-frame face SR problem is challenging, as there is less observed data available and we have to take into account the specific features of human faces.

5.3 New Learning-based Face Super-resolution Method

It is worth pointing out that the global characteristics of a human face is essential for face SR [97], such as explicit contour, coherent illumination, symmetry of the eyes or noses. In view of this, the schematic overview of the proposed method is given in Fig. 5.1. Given a single LR image g, a global HR image f_g is first reconstructed by using a PCA-based method. PCA

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

technique is employed to represent the LR image g as a linear combination of the training LR face images. Then the combination weights can be used to render the global HR face as a linear combination of the HR face images in the training set. However, the obtained HR result often looks smooth due to lack of detailed local face features. Therefore, at the second stage, a patch-based method using the LLE technique is developed to estimate the *residue image* $f_r = f - f_g$, which represents detailed local features missing in the global HR image f_g. In this patch-based method, we divide the images into overlapped patches. Then the residual image can be reconstructed by putting together all the estimated residual patches. Inspired by the LLE technique, we assume that the patches in the high-pass filtered global images and residual images form manifolds with approximately the same local structure. Then each residual patch can be estimated from its neighbors in the training set. The local geometry is obtained by the corresponding patches in the high-pass filtered global images. Finally, the estimated HR image $\hat{f}$ can be reconstructed by adding the estimated residue image to the reconstructed global image. We will explain more on the estimation of f_g and f_r in the following subsections.

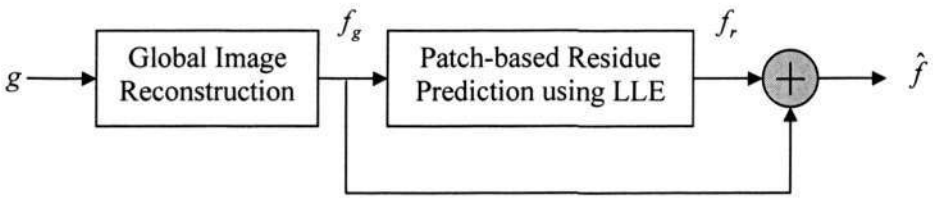


Fig. 5.1: Schematic diagram of the proposed method.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.3.1 PCA-based Global Face Reconstruction

In the Wang's method [98], the reconstructed SR face is synthesized by the linear combination of HR images in the database. A main issue arising from the method is that, when the given LR face is a new face significantly different from those in the face database, the obtained result can be less than satisfactory. In this Chapter, we name this issue as the "*new data*" problem. To alleviate this difficulty, we develop an improved PCA-based method to render a global HR face f_g by incorporating an estimated initial HR face f_c in the training set.

We define N as the number of the images in our face database. Different from the Wang's algorithm [98], our face training set for eigentransformation contains $N+1$ pairs of LR and HR face images as shown in Fig. 5.2. l_j and z_j $(1 \le j \le N)$ represent the LR faces and the corresponding HR faces in the face image database. f_c denotes the initial estimate of the HR face that can be obtained by using the restoration-based interpolation method given by:

$$\boldsymbol{f}_c = (\boldsymbol{W}^T\boldsymbol{W} + \lambda \boldsymbol{R}^T\boldsymbol{R})^{-1}\boldsymbol{W}^T\boldsymbol{g} \tag{5.2}$$

where $\boldsymbol{f}_c$ denotes the lexicographically ordered f_c. $\boldsymbol{R}$ is a conventional 2D Laplacian filter used as the regularization operator, and λ is the regularization parameter moderating the weight of the regularization component.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

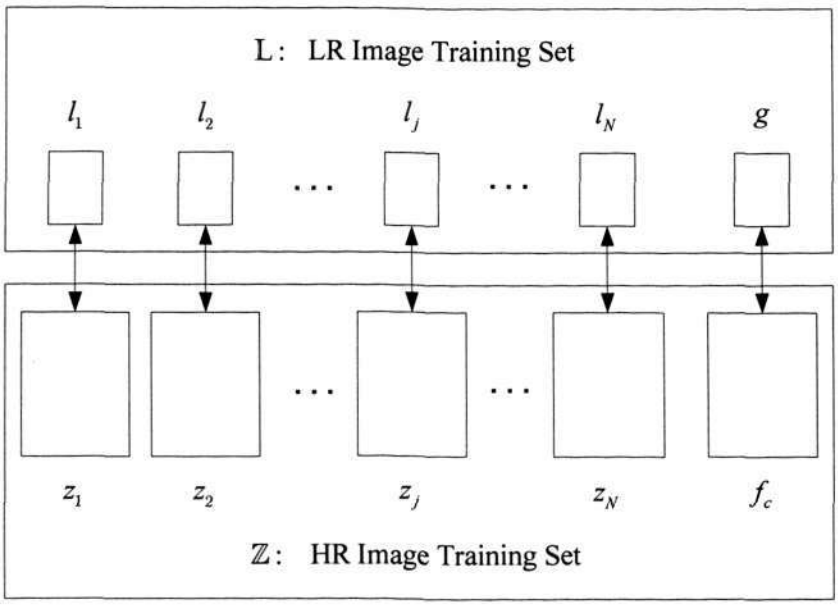


Fig. 5.2: Structure of the face training set.

The schematic diagram for estimating the global face f_g is given in Fig. 5.3. In this scheme, PCA technique is used to decompose the LR image g as a linear combination of the training LR face images. We define the combination weights as $w_j (1 \le j \le N+1)$, which represent the importance each LR image in the training set contributes to the reconstruction of the LR image. Previous studies in [98] have shown that it is reasonable to assume that the combination weights in both the LR and HR domains are similar. Then the HR counterparts in the training set can be linearly combined to render the global face f_g with the same combination weights. The details are discussed as follows.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

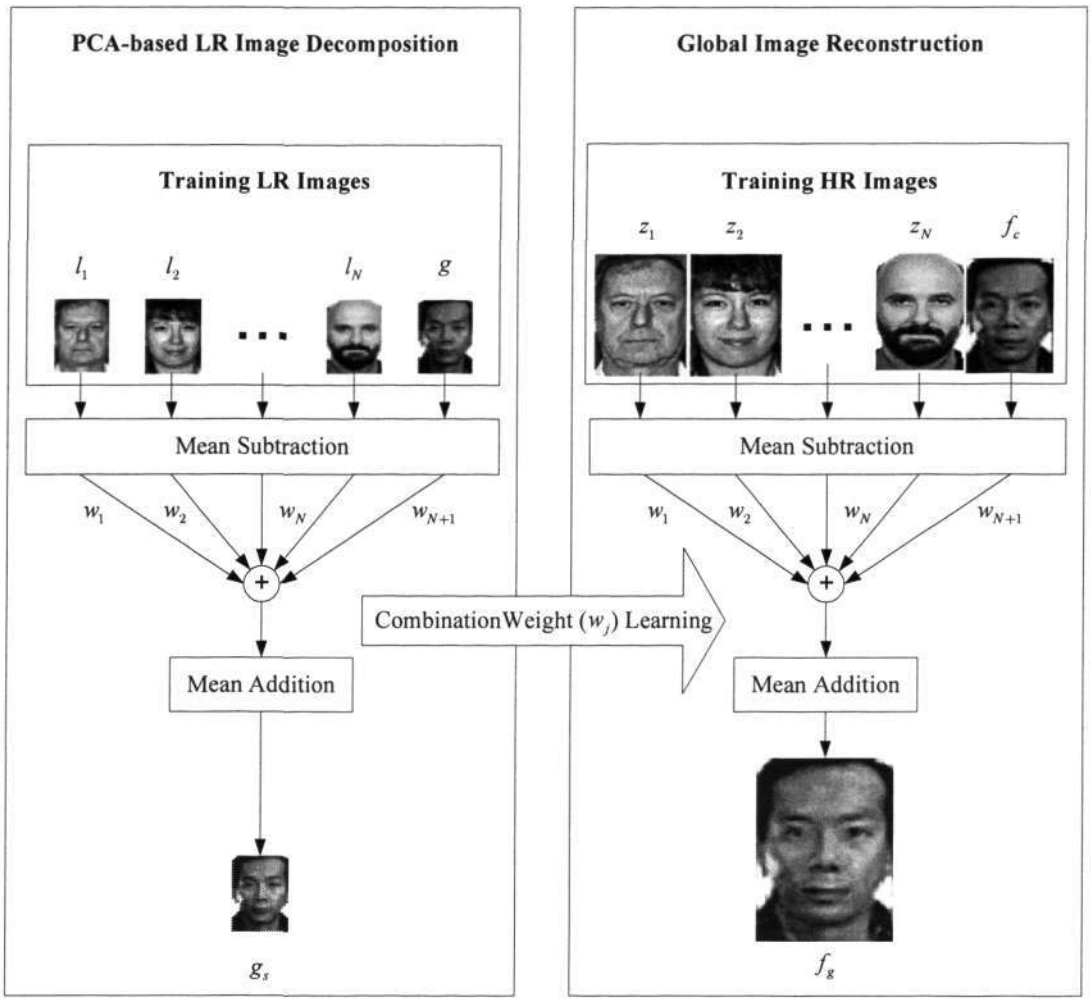


Fig. 5.3: Schematic diagram for estimating the global face f_g.

In order to estimate the combination weights, we first decompose the LR image g from eigenfaces, which can be thought of as a set of features that together characterize the variation between face images. We denote $\boldsymbol{m}_l$ and $\boldsymbol{m}_z$ as the lexicographically ordered mean faces for the LR and HR image training sets, respectively. We construct a matrix $\boldsymbol{L}$ in which each column represents each lexicographically ordered zero centered LR image in the training set. Then $\boldsymbol{L}$ can be written as:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\boldsymbol{L} = [\tilde{\boldsymbol{l}}_1 \ \cdots \ \tilde{\boldsymbol{l}}_N \ \tilde{\boldsymbol{g}}], \text{ where } \tilde{\boldsymbol{l}}_j = \boldsymbol{l}_j - \boldsymbol{m}_l, \text{ and } \tilde{\boldsymbol{g}} = \boldsymbol{g} - \boldsymbol{m}_l \tag{5.3}$$

Similarly, we define a matrix $\boldsymbol{Z}$ for the HR image training set as follows:

$$\boldsymbol{Z} = [\tilde{\boldsymbol{z}}_1 \ \cdots \ \tilde{\boldsymbol{z}}_N \ \tilde{\boldsymbol{f}}_c], \text{ where } \tilde{\boldsymbol{z}}_j = \boldsymbol{z}_j - \boldsymbol{m}_z, \text{ and } \tilde{\boldsymbol{f}}_c = \boldsymbol{f}_c - \boldsymbol{m}_z \tag{5.4}$$

To represent the observed LR image $\boldsymbol{g}$ by eigenfaces, we need to calculate the eigenvectors of the covariance $\boldsymbol{C} = \boldsymbol{L}\boldsymbol{L}^T$. Due to the large size of $\boldsymbol{C}$, determining the eigenvectors is an intractable task for medium-large image size. A computationally feasible method is to compute the eigenvector matrix $\boldsymbol{V}$ for a smaller matrix $\boldsymbol{L}^T\boldsymbol{L}$ [152]. $\boldsymbol{V}$ is composed of q eigenvectors corresponding to the first q eigenvalues in decreasing order. Based on the PCA properties, the eigenvectors corresponding to the relatively very small eigenvalues have less contribution to the PCA synthesis in the sense of mean-square errors. To determine the number of required eigenvectors q, we choose that the sum of the first q eigenvalues is above a certain threshold. The impact of choosing different q is discussed in Section 5.4.1. The eigenvector matrix $\boldsymbol{E}$ of $\boldsymbol{C}$ can then be computed by $\boldsymbol{E} = \boldsymbol{L}\boldsymbol{V}$.

Following the PCA theory, the observed LR face $\boldsymbol{g}$ can be approximated by a linear combination of the eigenvectors. The combination weight vector $\boldsymbol{b}$ can be obtained by projecting the observed LR face $\boldsymbol{g}$ into the eigenface space as:

$$\boldsymbol{b} = \boldsymbol{E}^T(\boldsymbol{g} - \boldsymbol{m}_l) \tag{5.5}$$

Based on the combination weight vector $\boldsymbol{b}$, the synthesized LR face image $\boldsymbol{g}_s$ is computed by:

$$\begin{aligned} \boldsymbol{g}_s &= \boldsymbol{E}\boldsymbol{b} + \boldsymbol{m}_l \\ &= \boldsymbol{L}\boldsymbol{V}\boldsymbol{b} + \boldsymbol{m}_l \\ &= \boldsymbol{L}\boldsymbol{w} + \boldsymbol{m}_l \end{aligned} \tag{5.6}$$

where the coefficients in $\boldsymbol{w}$ are the combination weights w_j and the vector $\boldsymbol{w}$ can be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

computed by using $\boldsymbol{w} = \boldsymbol{L}\boldsymbol{b}$. Previous studies in [98] have shown that it is reasonable to assume that the combination weights in both the LR and HR domains are similar. Therefore, the reconstructed HR global face f_g can be computed as:

$$\boldsymbol{f}_g = \boldsymbol{Z}\boldsymbol{w} + \boldsymbol{m}_z \tag{5.7}$$

where $\boldsymbol{f}_g$ denotes the lexicographically ordered global face f_g. From (5.7), it is observed that the global face is a linear combination of the HR faces in the training set. It can be considered that the estimated f_g is an approximation of the true HR face. Experimental results in later sections have shown that the reconstructed global face is able to generate a better HR image that bears closer resemblance to the original face, while preserving common characteristics of a human face.

5.3.2 Patch-based Residue Prediction by LLE

An issue arising from the global image reconstruction is that the obtained global HR face f_g looks smooth, since the initial estimate of the HR face f_c in the training set has lost much of the detailed information. In this section, we aim to predict the residual image $f_r = f - f_g$, which represents the detailed local features missing in the global image. The residual image can be divided into overlapped residual patches. We will try to find each residual patch f_{rp} and put them together to render the residual image. In this section, a patch-based residue prediction method is developed by employing the LLE-based neighbor reconstruction technique, which estimates the residual patches as a linear combination of its k-nearest neighbors (k-NN) in the training set.

The schematic diagram for estimating the residual face f_r is given in Fig. 5.4. From the figure, we summarize the proposed algorithm as follows.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(1) Generate the global image set G corresponding to the LR images in the database by using the method presented in the previous Section.

(2) Generate the residual image set R from $\mathbb{Z}$ by using $\mathrm{R} = \mathbb{Z} - \mathrm{G}$.

(3) Partition and extract the patch set $\mathrm{G_P}$ from the high-pass filtered G.

(4) Partition and extract the patch set $\mathrm{R_P}$ from the residual images R.

(5) Partition and extract each patch f_{gp} in raster-scan order from the high-pass filtered global image f_g.

(6) Search the k-NN f_{gp}^k of f_{gp} from $\mathrm{G_P}$, and compute the linear combination weights.

(7) Synthesize the residual patch f_{rp} through its k-NN f_{rp}^k by employing the LLE algorithm.

(8) Reconstruct the residue image f_r by putting together all the residual patches.

In steps (1-4), we prepare our training data by extracting the patches from training images. Similar to [2] and [80], we believe that the high-frequency components of the reconstructed global face are important in predicting the residue image. Thus, the training data consists of two groups. One group contains the patches $\mathrm{G_P}$ extracted from the high-pass filtered G. The other group includes the patches $\mathrm{R_P}$ extracted from the residual images R. Then the patches $\mathrm{G_P}$ and $\mathrm{R_P}$ are considered as training pairs according to their locations. Now the problem for the patch-based residual image estimation can be stated as: given the patch f_{gp} in the estimated global image f_g, find the corresponding residual patch f_{rp} with the help of the training pairs $\mathrm{G_P}$ and $\mathrm{R_P}$.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

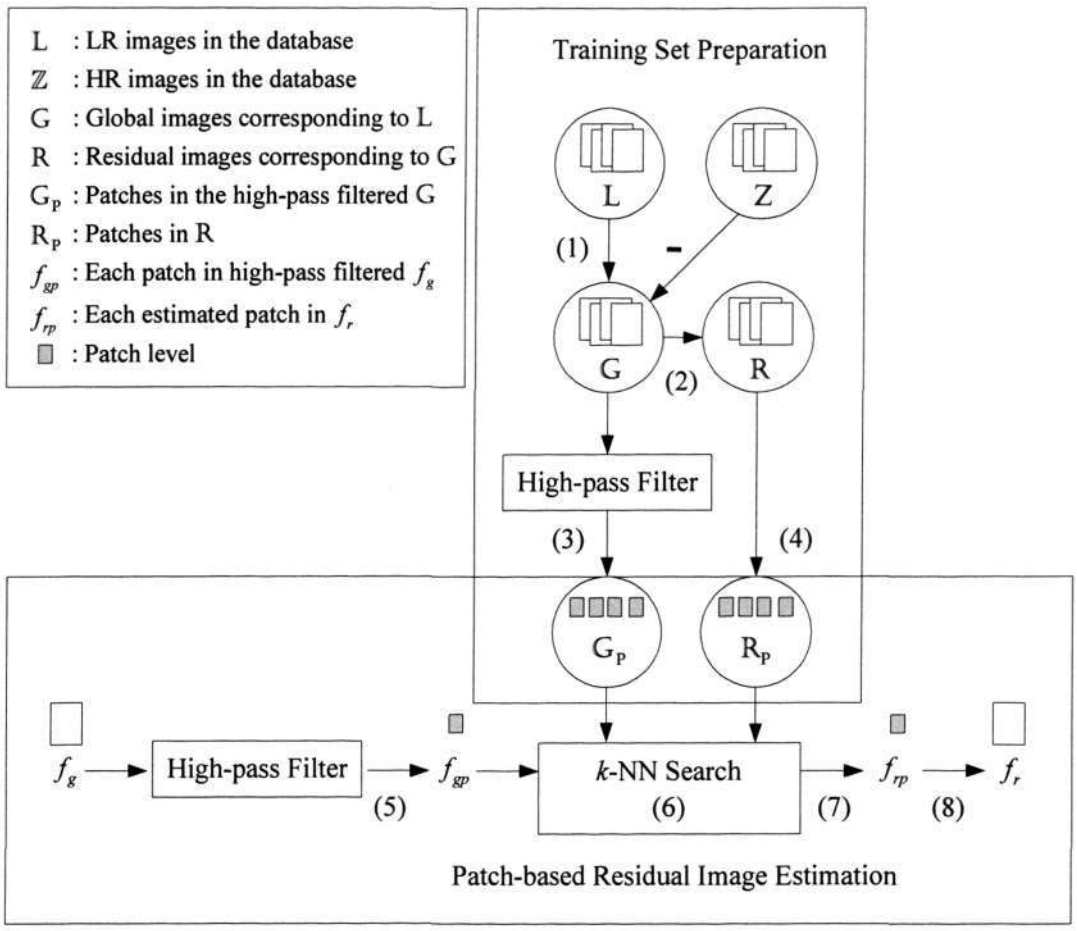


Fig. 5.4: Schematic diagram for estimating the residual face f_r.

One direct method is to find the best match with respect to f_{gp} from the training patches G_P. The corresponding residual patch in R_P could then be used for the estimation of f_{rp} [2]. However, if no suitable patch is available from G_P, a weighted average of several best matches can be used to produce an improved residue synthesis. In view of this, we employ the LLE-based neighbor reconstruction technique [81], [153] for the residual patch prediction. The LLE algorithm has been recently proposed for the purpose of manifold learning or nonlinear dimensionality reduction. It characterizes the local geometry of high-dimensional data by linear coefficients that reconstruct each data point from its neighbors. It is assumed that the high-and

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

low-dimensional data share the same local geometry. Then the low-dimensional points can be computed by employing neighborhood preserve mapping techniques. Inspired by the LLE technique, we assume that the patches in the high-pass filtered global images G_p and residual images R_p form manifolds with approximately the same local structure. For better understanding, Fig. 5.5 illustrates the basic idea for estimating the patch f_{rp} corresponding the patch f_{gp}. The patch f_{gp} could be approximated by a weighted linear combination of its *k*-NN f_{gp}^k, the *k* best matched patches in G_p. The linear reconstruction weights ξ_k can be found by minimizing the local reconstruction errors. This will be explained later. We denote the counterpart of f_{gp}^k in training pairs as f_{rp}^k. Based on the LLE property, the residual patch f_{rp} can be estimated from f_{rp}^k by preserving local geometry formed by f_{gp} and its neighbors f_{gp}^k. More specifically, the patch f_{rp} is linearly combined by its neighbors f_{rp}^k, while retaining the same weights ξ_k.

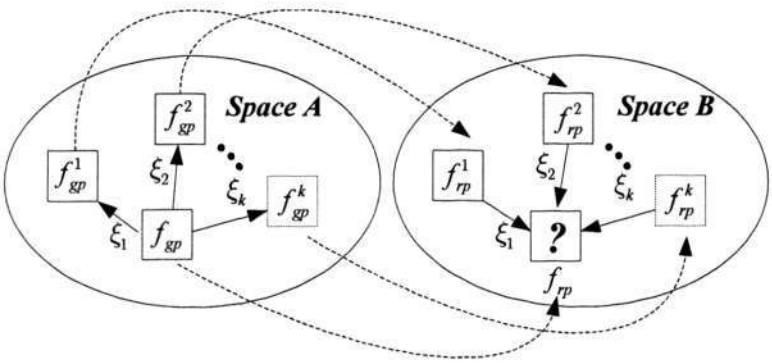


Fig. 5.5: Local structure of the patches in different spaces. (***Space A***: the space for the patches in the high-pass filtered global images G_p, ***Space B***: the space for the patches in residual images R_p).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The implementation of this LLE-based neighbor reconstruction algorithm is explained as follows. It is noted that our neighbor-searching algorithm takes local compatibility and smoothness into consideration. We decide that the predicted residue patches at their border are overlapped. Similar to Freeman's method [2], we construct a search vector to seek the *k*-NN of f_{gp}. The search vector is consisted of the pixel values in the patch f_{gp} and previous estimated residues at the border. Then we can use Euclidean distance to determine the *k*-NN f_{gp}^k. After finding the neighbors, the optimal weights ξ_k can then be obtained by minimizing the local reconstruction errors. We denote $\boldsymbol{f}_{gp}$ and $\boldsymbol{f}_{gp}^k$ as the vectors formed by f_{gp} and f_{gp}^k, respectively. The minimization problem subject to the constraints is solved as follows:

$$\xi_k = \arg\min_{\xi_k} \frac{1}{2}\left\| \boldsymbol{f}_{gp} - \sum_k \xi_k \boldsymbol{f}_{gp}^k \right\|^2 \quad \text{subject to: } \sum_k \xi_k = 1 \tag{5.8}$$

The problem is equivalent to solving a constrained least squares problem. The computation of the reconstruction of the weights ξ_k is discussed in Appendix A of this chapter. Finally, the residual patch f_{rp} can be computed based on ξ_k:

$$f_{rp} = \sum_k \xi_k f_{rp}^k \tag{5.9}$$

After putting together all the residual patches, the final estimated HR image could be obtained by adding the predicted residue image f_r to the reconstructed global image f_g as shown in Fig. 5.1. However, some artifacts may appear in the HR face image. We use bilateral filtering in [154] to remove these artifacts.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.4 Experiments

In this section, we study the performance of the proposed method and compare it with other methods. The subset[3] of the FERET data set [155] was used as the face database, which consists of 454 face images. All the images were roughly aligned by affine transform based on the location of the eyes and the nose [96]. We conducted experiments on both simulated and real-life images. We compared the results obtained using the proposed method with those using bicubic interpolation and the Wang's method [98]. However, the best performance measure is still human inspection of the reconstructed HR images.

5.4.1 Super-resolution on Simulated Images with Additive Noise

The first experiment is based on the simulated images. Following the works in [96] and [97], we randomly extracted images in the face database as test images, while using the remaining for training. To generate the corresponding test LR images, the HR images were blurred by uniform blurs with support of 4×4, downsampled by factor 4 and further degraded by AWGN to produce a signal-to-noise ratio (SNR) of 40dB. Some samples of the scaled-up LR faces are shown in Fig. 5.6(a).

We first study the impact of the number of eigenvectors q for the reconstruction of the global face f_g. It is noted that we choose q such that the sum of the first q eigenvalues is above a certain percentage of total of eigenvalues. The percentage varies from 96%, 97%, 98% and 99%, and the corresponding results are shown in Figs. 5.6 (b)-(e). It is observed that with higher number of eigenvectors, the details of the facial features in the reconstructed global face increases. The difference of the results with respect to the percentage 98% and 99% are difficult

[3] The image data are provided by Simon Baker from Carnegie Mellon University.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

to perceive. Therefore, we selected 98% as the threshold during the reconstruction of the global face.

To further evaluate our method, we performed the patch-based algorithm discussed in Section 5.3.2 in order to predict the residue images. We chose the patch size as 5×5. The number of the nearest neighbors was set to 3. The final results are displayed in Fig. 5.7(e). Compared with the global faces shown in Fig. 5.7(d), the final results generate clearer faces with strong facial features, in particular near the eyes and nose. When compared with the results by bicubic interpolation in Fig. 5.7(b), it is seen that the interpolated results are able to maintain the general characteristics of human faces. However, they look too smooth and much facial details are lost. We also compared our results with that by the Wang's method [98] in Fig. 5.7(c), it is observed that the proposed method can generate with satisfactory global appearance and detailed local facial features.

For objective performance evaluation, the PSNR of the reconstructed HR face image is used. From the Table 5.1, it is observed that the proposed method is able to provide the highest PSNR values among all three methods. We also conducted a survey on ten persons to evaluate the performance of the proposed method. A point scoring scheme (5, 4, 3, 2, 1, 0) that corresponds to the quality of the reconstructed face image (very good, good, average, not good, poor) is used. The average results are calculated and shown in Table 5.1. Female faces 1, 2 and male faces 1, 2 correspond to the ones in Fig. 5.7 from the left to the right. The subjective evaluation by survey further reconfirms the objective performance measure. Comparison reveals that the proposed method is able to reconstruct better HR face images when compared with the other two methods.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table 5.1: Performance of different methods for single-frame face SR

Method	Female 1		Female 2		Male 1		Male 2	
	PSNR	Subjective score	PSNR	Subjective score	PSNR	Subjective score	PSNR	Subjective score
Bicubic interpolation	20.74	2.2	23.68	1.9	20.91	2	19.86	1.9
Wang's method	21.87	2.3	23.23	2	22.34	1.7	22.61	2.5
Proposed method	23.54	3.5	25.47	3.2	24.2	3	23.4	3.8

The proposed method is able to handle new face, which is interpreted as faces with characteristics significantly different from the rest of faces in the database. Under previous experimental setting, it is noted that all the faces in the database do not wear glasses. The test images, on the other hands, consist of faces wearing glasses, which could be considered as new faces. Some samples of the scaled-up LR faces are shown in Fig. 5.8(a). We compared the results using the proposed method with those obtained by the Wang's method [98] in Fig. 5.8(c). It is observed that the proposed method can generate better HR images that bear closer resemblance to the original faces. The new experiment demonstrates that when the test image is a new face with characteristic significantly different from the rest of the faces in the database, the Wang's method is unable to reconstruct a satisfactory HR face. Experimental results show that the proposed approach is able to generate better quality HR images.

We will use the Wang's method as the benchmark for comparison of computational complexity. At the stage of global image reconstruction, both methods use a linear parametric model, resulting in similar run time. However, the proposed scheme further uses the LLE algorithm for residue prediction, which is time consuming due to the k-NN search. Therefore, the computational complexity of our algorithm is more than that of the Wang's method. We use experimental simulation time to demonstrate this. The simulation environments are as follows: Windows XP, MATLAB 7.1, CPU P4-3.4 GHz, and 1G RAM. The average computational time for our method with a HR test image of dimension 96×128 is 127s. The average computational

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

time of the Wang's method is 33s. Nevertheless, it should be pointed out that the proposed method offers superior reconstructed face images when compared with the Wang's method, as shown in the experiments.

5.4.2 Super-resolution on Real-life image

We conduct real-life experiment to further illustrate the efficiency of the proposed method. Fig. 5.9(a) is a LR picture taken by a web-camera in a lab. The LR face was extracted, scaled-up and shown in Fig. 5.9(b). We chose a decimation factor of 4 so that the estimated HR face will have four times the resolution of the LR face. For training, we also used the image database in the previous experiments. The reconstructed HR faces using bicubic interpolation, the Wang's method and the proposed method are shown in Figs. 5.9 (c)-(e), respectively. It is observed that the result obtained by bicubic interpolation is too blurry and loses much of the facial details. Furthermore, it can be seen that the result by the Wang's method experiences significant degradation and can hardly preserve the facial characteristics. It is shown that the proposed method is able to generate better result, as it keeps the global as well as local appearance, in particular near the eyes and nose. However, it is observed that there are still some artifacts around the face contour and the margin of the mouth. We could employ some adaptive bilateral filtering works dependent on image locations to improve the final result. This can be pursued in future work.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

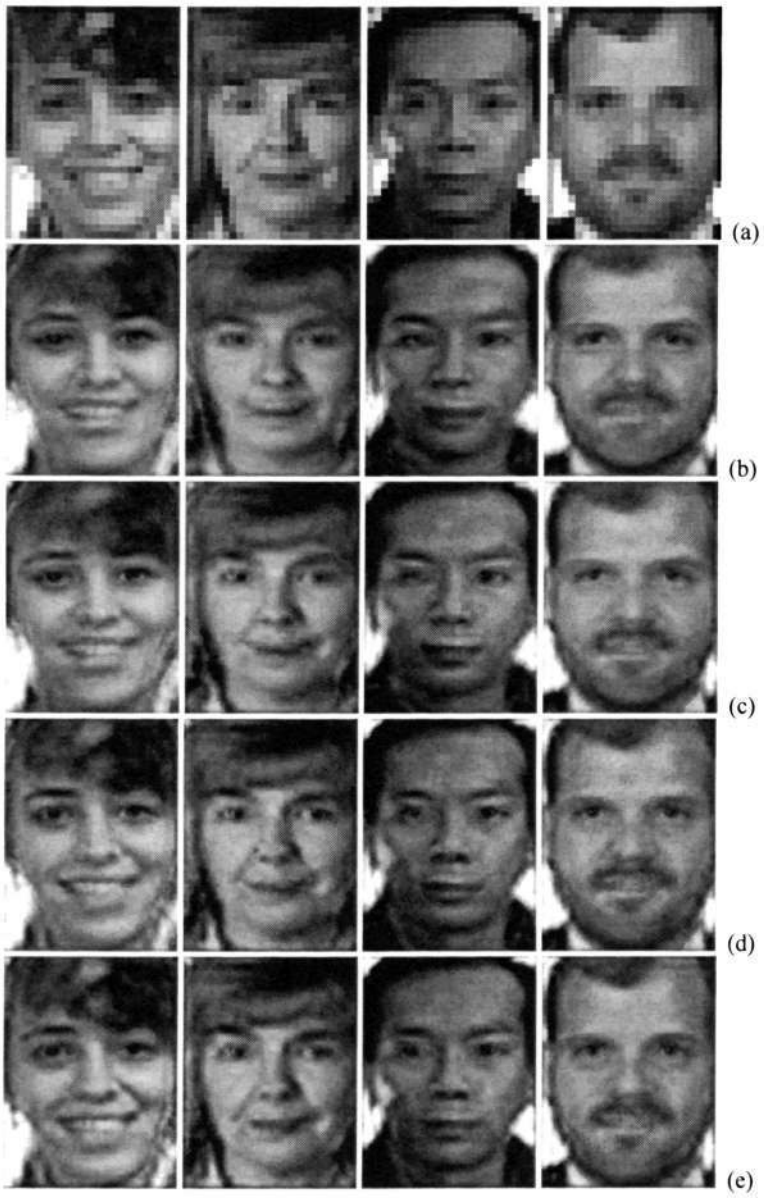


Fig. 5.6: Reconstruction of global faces (a) Scaled-up LR faces, (b)-(e) Reconstructed global faces by using different numbers of eigenvectors corresponding to variant percentage of total of eigenvalues, 96%, 97%, 98% and 99%, respectively.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

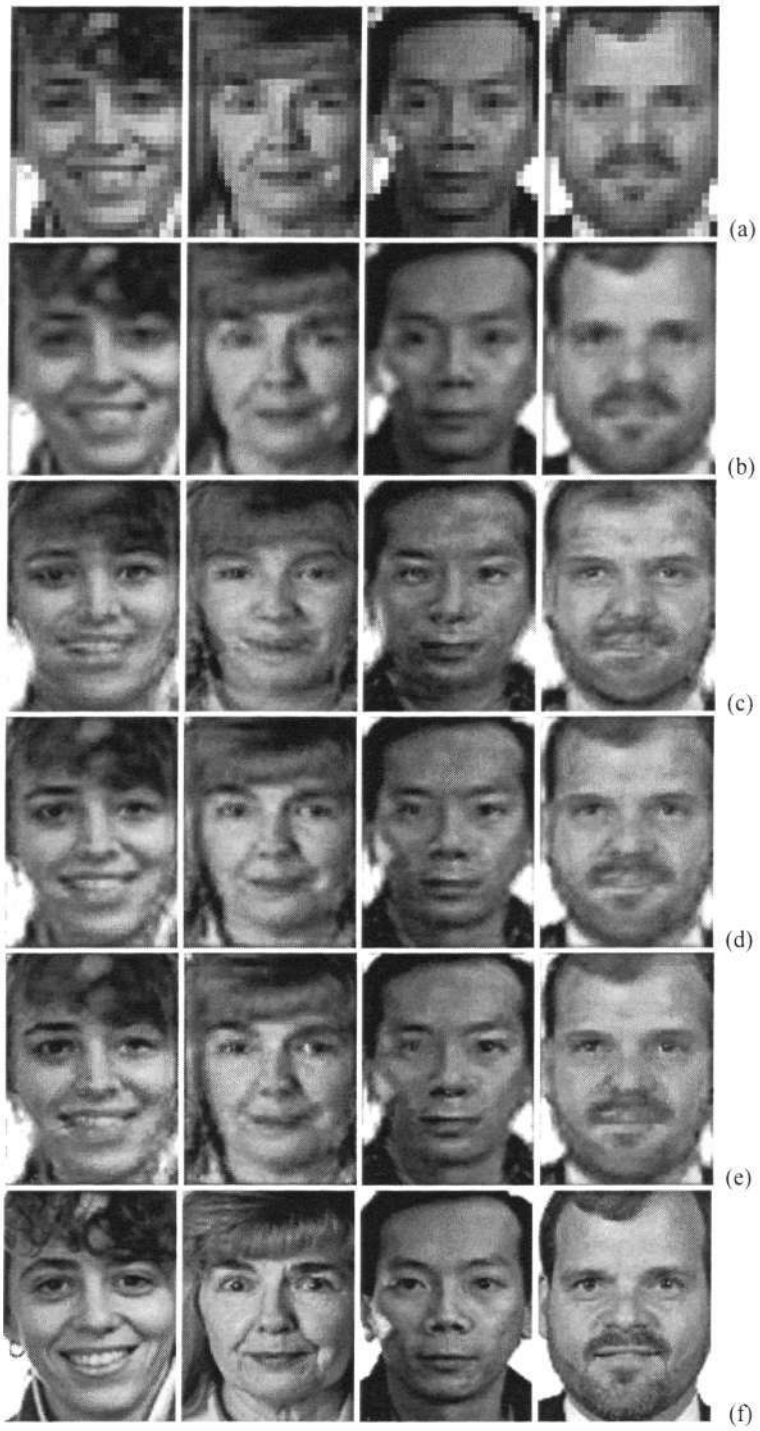
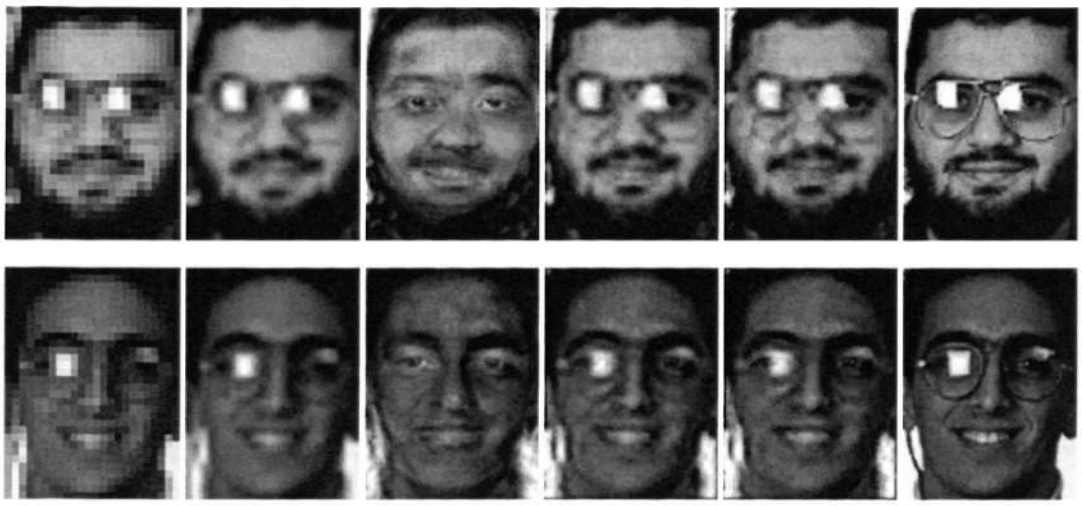


Fig. 5.7: SR on LR faces degraded by AWGN (a) Scaled-up LR faces, (b) Reconstructed faces by bicubic interpolation, (c) Reconstructed faces by the Wang's method [98], (d) Reconstructed global faces by the proposed method, (e) Final results by the proposed method, (f) Original faces.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) (b) (c) (d) (e) (f)

Fig. 5.8: SR on new LR faces degraded by AWGN (a) Scaled-up LR faces, (b) Reconstructed faces by bicubic interpolation, (c) Reconstructed faces by the Wang's method [98], (d) Reconstructed global faces by the proposed method, (e) Final results by the proposed method, (f) Original faces.

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

Fig. 5.9: SR on a real-life face. (a) A LR image, (b) Extracted and scaled-up LR face, (c) Reconstructed face by bicubic interpolation, (d) Reconstructed face by the Wang's method [98], (e) Reconstructed face by the proposed method.

5.5 Summary

This chapter presents a new algorithm to address single-frame face SR. An improved PCA-based method is proposed to render a global HR face, which is able to maintain common

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

characteristics of a human face. To compensate the global face for detailed features, we propose a patch-based residue prediction approach by using a manifold learning method named locally linear embedding. Unlike the method in [98], the proposed method is more effective in handling cases when the given LR image is significantly different from those in the face database.

Appendix A: Computation of Reconstruction Weights ξ_k

In this Appendix, we will briefly present the computation of the reconstruction weights ξ_k. The minimization problem subject to constraint in (5.8) is equivalent to a constrained least squares problem. The problem can be solved by using the Lagrange multipliers method [82], [156]. We construct a matrix $\boldsymbol{N}$ in which each column is used by $\boldsymbol{f}_{gp}^{k}$ and a vector $\boldsymbol{\xi}$ in which each coefficient is ξ_k, respectively. Then the minimized cost function is modified as:

$$\begin{aligned} A(\boldsymbol{\xi}) &= \frac{1}{2}\left\| \boldsymbol{f}_{gp} - \sum_k \xi_k \boldsymbol{f}_{gp}^{k} \right\|^2 + \beta\left(\sum_k \xi_k - 1\right) \\ &= \frac{1}{2}\left\| \boldsymbol{f}_{gp}\boldsymbol{I}^T\boldsymbol{\xi} - \boldsymbol{N}\boldsymbol{\xi} \right\|^2 + \beta(\boldsymbol{I}^T\boldsymbol{\xi} - 1) \\ &= \frac{1}{2}\left\| (\boldsymbol{M} - \boldsymbol{N})\boldsymbol{\xi} \right\|^2 + \beta(\boldsymbol{I}^T\boldsymbol{\xi} - 1), \qquad \text{where } \boldsymbol{M} = \boldsymbol{f}_{gp}\boldsymbol{I}^T \end{aligned} \tag{5A.1}$$

where β is the Lagrange multiplier or regularization parameter and $\boldsymbol{I}$ is a vector of all ones. The solution of the minimization problem can be determined by solving $\partial(A(\boldsymbol{\xi}))/\partial(\boldsymbol{\xi}) = 0$ as follows:

$$\frac{\partial(A(\boldsymbol{\xi}))}{\partial(\boldsymbol{\xi})} = (\boldsymbol{M} - \boldsymbol{N})^T(\boldsymbol{M} - \boldsymbol{N})\boldsymbol{\xi} + \beta\boldsymbol{I} = 0 \tag{5A.2}$$

Then $\boldsymbol{\xi}$ can be obtained by

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\boldsymbol{\xi} = -\beta\left((\boldsymbol{M}-\boldsymbol{N})^T(\boldsymbol{M}-\boldsymbol{N})\right)^{-1}\boldsymbol{1} \tag{5A.3}$$

Substituting for $\boldsymbol{\xi}$ from (5A.3) in the constraint $\boldsymbol{1}^T\boldsymbol{\xi} = 1$, we get:

$$\beta = -\frac{1}{\boldsymbol{1}^T\left((\boldsymbol{M}-\boldsymbol{N})^T(\boldsymbol{M}-\boldsymbol{N})\right)^{-1}\boldsymbol{1}} \tag{5A.4}$$

Then the constrained least squared problem has the following closed-form solution as:

$$\boldsymbol{\xi} = \frac{\left((\boldsymbol{M}-\boldsymbol{N})^T(\boldsymbol{M}-\boldsymbol{N})\right)^{-1}\boldsymbol{1}}{\boldsymbol{1}^T\left((\boldsymbol{M}-\boldsymbol{N})^T(\boldsymbol{M}-\boldsymbol{N})\right)^{-1}\boldsymbol{1}} \tag{5A.5}$$

Instead of calculating the denominator, a more efficient way is to solve the linear system of equation $\left((\boldsymbol{M}-\boldsymbol{N})^T(\boldsymbol{M}-\boldsymbol{N})\right)\boldsymbol{\xi} = \boldsymbol{1}$, and then normalize the weights so that the summation of ξ_k equals to one. Further information on the computation of ξ_k can be found in [156].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 6

Blind Color Image Deconvolution

6.1 Introduction

This chapter focuses on the blind color image deconvolution (restoration), which is an inverse problem that attempts to estimate original color image from single degraded blurred color image at the same pixel resolution scale, given limited or no prior knowledge of the blurring function [135]. It is noted that image deconvolution and image SR in the previous chapters are closely related problems. In image SR, HR images could be reconstructed in three stages, namely (i) image registration from LR images, (2) image fusion and followed by (3) image deconvolution. Thus, image deconvolution can also be considered as one stage of image SR. It is used in the areas of photography deblurring, astronomical imaging, remote sensing, and microscopy imaging, among others.

Most current algorithms on blind image restoration focus on restoration of grayscale images. They can be divided broadly into two categories: (i) blind SC grayscale image restoration [104], [157], and (ii) blind MC grayscale image restoration [122], [124]. The amount of works on blind color image restoration is still very limited [135]. Blind SC grayscale restoration attempts to estimate the original grayscale image from a single observation of blurred grayscale image. In contrast, blind MC grayscale restoration attempts to estimate the original grayscale image from multiple blurred grayscale images of the same imaging scene. Efforts to extend these two frameworks to blind color image restoration have achieved only

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

limited success so far. Works to extend blind SC restoration to blind color image restoration include [135]. The results are usually unsatisfactory because: (i) the method ignores interchannel blurring due to causes such as channel crosstalk, and (ii) the information of interchannel correlation has not been fully utilized. On the other hand, attempts to extend blind MC restoration to blind color image restoration encounter the following problems: (i) there is only a single observation of blurred color image, and (ii) the color image degradation system follows multi-input multi-output (MIMO) modeling, rather than the single-input multi-output (SIMO) modeling used by blind MC restoration algorithms.

One of the most challenging issues in addressing blind color image deconvolution is that the solution to the original MIMO model is intractable. In other words, we cannot obtain a direct estimate to the original color image using conventional MIMO algorithms since the RGB color channels are highly correlated. Further, the computational cost involved in extending the one-dimensional (1D) MIMO algorithms in communication theory to two-dimensional (2D) images is significant. In view of this, there is a real need and motivation to reformulate the original MIMO model into a new framework so that the solution to the new model will provide a reasonably good estimate to the original color image, while preserving the visual quality of the solution.

The contribution of this chapter, therefore, is to propose a new framework that offers a tractable solution to alleviate the intractable original MIMO deconvolution problem. It also addresses the issues encountered when extending the SC and MC grayscale blind image deconvolution into the color image domain. In the proposed framework, the blind color image restoration is performed using a hybrid of SISO and SIMO models. The new method has two characteristics: (i) in SISO model, we utilize the local spatial smoothness of low frequency component in each channel to perform image regularization, and (ii) in SIMO model, the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

technique exploits the spectral correlation between the high-frequency subbands of different channels to perform deconvolution. The method also takes intrachannel and interchannel blurring into consideration. Experimental results show that the new method is effective in restoring color images where there is limited information about the blurring functions.

The organization of the rest of this chapter is outlined as follows. The mathematical model of color image degradation is introduced in Section 6.2. A wavelet-based deconvolution scheme for color images and its justification are presented in Section 6.3. Section 6.4 discusses the approximate subband deconvolution using SISO modeling. In Section 6.5, the detailed subband deconvolution based on SIMO modeling is explained. In Section 6.6, the optimization procedure is discussed. Experimental results are presented in Section 6.7. Finally, a summary is given in Section 6.8.

6.2 Problem Formulation

The linear color image degradation processes are commonly modeled by [127], [135]

$$g_j = \sum_{i=R,G,B} h_{ij} \otimes f_i + n_j \ , \quad j = R,G,B \tag{6.1}$$

where g_j, f_i, and n_j are the j-th blurred color channel, i-th original color channel, and j-th channel noise, respectively. h_{ii} and h_{ij} $(i \neq j)$ are intrachannel and interchannel PSFs or blurs. The operator $\otimes$ denotes two-dimensional convolution operation. The intrachannel blurring usually dominates the interchannel blurring, or in other words, the coefficients of h_{ij} $(i \neq j)$ are usually much smaller than that of h_{ii}. The objective of blind color image deconvolution is to estimate f_i , $i = R,G,B$ given the three channels of the blurred image y_j , $j = R,G,B$. It is obvious that this is a challenging problem as the information about the blurring functions h_{ii}

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

and h_{ij} are not known and each color channel f_i is highly intercorrelated.

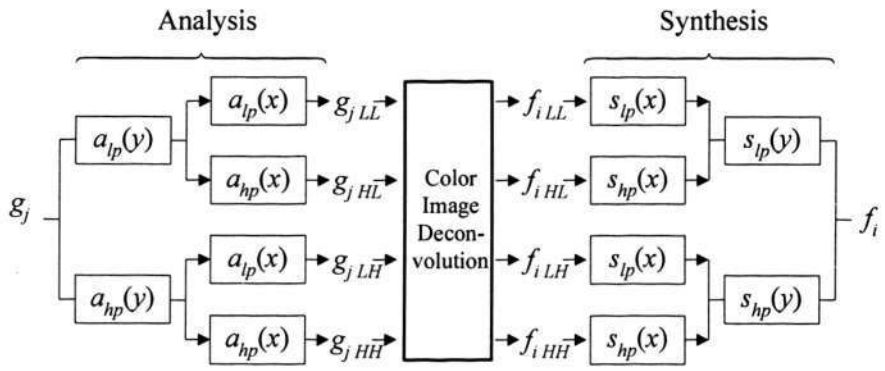


Fig. 6.1: Analysis and synthesis filter banks for color image deconvolution.

Before outlining the proposed restoration algorithm, we will first cover some preliminaries on the filter bank used in this chapter. Fig. 6.1 illustrates the analysis and synthesis stages of wavelet decomposition. The low-pass filters $a_{lp}(x)$ and $a_{lp}(y)$, and the high-pass filters $a_{hp}(x)$ and $a_{hp}(y)$ denote the analysis filter banks performing an undecimated wavelet transform in the vertical (x) and horizontal (y) directions, respectively. In contrast, the synthesis filters $s_{lp}(x)$, $s_{lp}(y)$, $s_{hp}(x)$, and $s_{hp}(y)$ reverse the process to obtain the original signal. As the redundant coefficients in the wavelet transform are useful for reconstruction in this case, we have chosen not to include the decimator in wavelet transformation. This operation is known as redundant discrete wavelet transform (RDWT). During the analysis stage, each decomposed channel of the blurred image is given as follows:

$$\begin{cases} g_{j\,LL} = a_{lp}(x) \otimes [a_{lp}(y) \otimes g_j(x,y)] \\ g_{j\,HL} = a_{hp}(x) \otimes [a_{lp}(y) \otimes g_j(x,y)] \\ g_{j\,LH} = a_{lp}(x) \otimes [a_{hp}(y) \otimes g_j(x,y)] \\ g_{j\,HH} = a_{hp}(x) \otimes [a_{hp}(y) \otimes g_j(x,y)] \end{cases} \tag{6.2}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

where $g_{j\,LL}$ is the approximate subband (LL), and $g_{j\,HL}, g_{j\,LH}, g_{j\,HH}$ are the horizontal (HL), vertical (LH), and diagonal (HH) detailed subbands of the j-th channel of the blurred image, respectively. Performing wavelet decomposition in similar fashion to (6.2) on both sides of (6.1), and ignoring the additive noise, we obtain the following:

$$\begin{cases} g_{j\,LL} = \sum_{i=R,G,B} h_{ij} \otimes f_{i\,LL}\,, & g_{j\,HL} = \sum_{i=R,G,B} h_{ij} \otimes f_{i\,HL} \\ g_{j\,LH} = \sum_{i=R,G,B} h_{ij} \otimes f_{i\,LH}\,, & g_{j\,HH} = \sum_{i=R,G,B} h_{ij} \otimes f_{i\,HH} \end{cases} \tag{6.3}$$

The formulation in (6.3) demonstrates that the original system in (6.1) can be transformed into an equivalent system consisting of subband decompositions of the original image f_i and the observed blurred image g_j. The reason for performing this transform is because it is observed that there exists different characteristics between spectral correlations for different subbands, and therefore they should be handled differently. The procedures and rationales of the proposed blind color image restoration algorithm in the subband domains will be explained in greater details in the following sections.

6.3 A Wavelet-based Deconvolution Scheme

6.3.1 Schematic Overview

Current blind color image deconvolution schemes strive to extend the framework of SISO deconvolution in grayscale images to color images [135], [158]. These approaches, however, give rise to artifacts near the edges and textured regions, as they do not fully utilize the correlation between different color channels. Further, they ignore the interchannel blurs, which will inevitably lead to a certain amount of restoration error. On the other hand, conventional

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

MIMO algorithms cannot be employed in this case, as the color channels are highly correlated. In view of this, we propose a scheme that is able to alleviate this dilemma.

The development of the new framework is centered on the following two observations:

(i). Studies show that images usually contain strong high-frequency correlation between three color (RGB) channels. This is reflected particularly in the edges and textured regions where changes in color profiles are strongly correlated [159].

(ii). The low-frequency components of each color channel of the images are often moderately correlated, as demonstrated in the smooth regions.

We develop a framework to exploit the properties above to address blind color image deconvolution. The schematic diagram of the new scheme is shown in Fig. 6.2. Based on the observations highlighted above, the proposed hybrid framework is divided into a two-stage process, namely subband deconvolution and subband synthesis. For approximate subband (LL) deconvolution, a SISO image deconvolution algorithm is used on the blurred RGB channels to produce the restored color channels. The restored LL subband of each channel can then be obtained by performing wavelet decomposition on their respective restored color channels. It is worth noting that we choose to apply SISO deconvolution on the blurred RGB channels instead of the blurred LL subbands as it is easier to introduce smoothness constraint in the original RGB domain by employing a 3D weighted Laplacian regularization function for image-domain. On the other hand, for detailed subbands (HL, LH, and HH) deconvolution, the proposed method first performs wavelet decomposition on the RGB blurred color channels to produce three sets of blurred detailed subbands. A SIMO image deconvolution algorithm is then employed to estimate the detailed subbands of the restored image. Finally, the restored color image is obtained by applying the synthesis filters on the restored LL, LH, HL and HH subbands obtained from the SISO and SIMO deconvolution.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

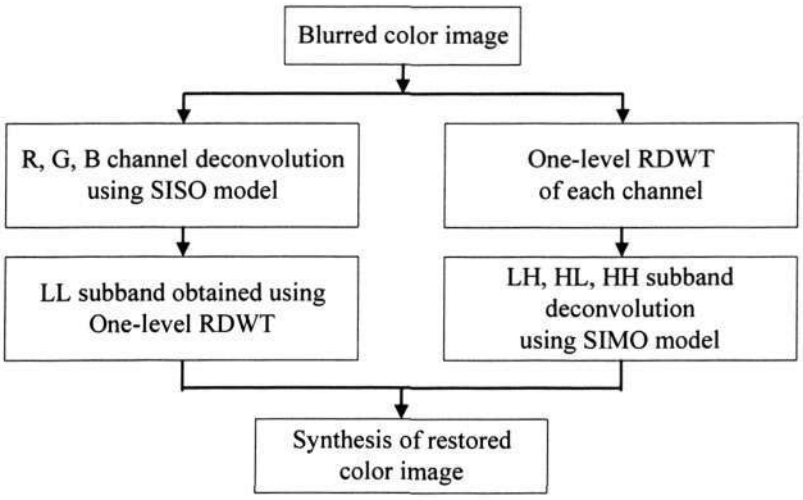


Fig. 6.2: Schematic diagram of the proposed algorithm.

6.3.2 Justifications of the Proposed Framework

In this section, we will attempt to provide justifications for the proposed framework. It should be noted that the proposed method does not attempt to solve the original MIMO structure directly, but rather approximate the system with a hybrid of SISO and SIMO models. We will provide holistic and experimental justifications to demonstrate the soundness of the proposed algorithm.

The property of color image highlighted in Section 6.3.1 has been widely reported and used in color filer array (CFA) demosaicking and color image denoising. Gunturk *et al.* [159] have shown that high-frequency wavelet coefficients of color components are strongly correlated. Further, Lian *et al.* [160] point that high-frequency coefficients of color components are not only correlated but also almost identical to each other. To illustrate the similarity between the high-frequency components of different color channels, we adopt a performance measure called spectral correlation coefficient in this work. The spectral correlation coefficient between two corresponding subbands of different color channels X and Y is given as [159]:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\rho_{X,Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{6.4}$$

where μ_X and μ_Y denote the means of X and Y, while σ_X and σ_Y represent the standard deviations of X and Y, respectively. Table 6.1 shows the spectral correlation coefficients computed based on the sample images given in Fig. 6.3. It is observed that the correlation coefficients in the high-frequency subbands $\{LH, HL, HH\}$ are very close to value of one, thereby demonstrating that there exists a very strong correlation between different color channels in the high-frequency subbands. Secondly, to further determine the similarity between corresponding subbands of different color channels, we show 3D scatter plots of high-frequency wavelet coefficients in Fig. 6.4. The scatter plots that we obtain are similar to those demonstrated in [160]. We use three color images "Boat", "Fence" and "Satellite" in Fig. 6.3 as the illustrative images. Different colors of magenta, cyan and black are used in the scatter plots to represent each of these color images, respectively. Each point in the plot represents the magnitudes of the R, G and B subband coefficients of the color images. It is observed that the points representing the detailed subband coefficients are lying predominantly along the line with vector $(1,\ 1,\ 1)^T$. This implies that the detailed subband coefficients are almost identical, an observation reported in many other color images [160]. Hence, we can employ SIMO deconvolution for restoration of the detailed subbands. In contrast, it is observed that the coefficients of the approximated LL subband are moderately correlated (values ranging from 0.54-0.95), and they do not lie along the line with vector $(1,\ 1,\ 1)^T$. Hence, the precondition for using SIMO deconvolution on LL approximate subband is not satisfied.

Both the above two performance measures show that there exists a strong correlation and similarity between the high-frequency subbands. Thus we assume that all channels are approximately equal at each high-frequency subband. Therefore, we utilize this property in the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

context of our algorithm. We use the SISO model for LL subband deconvolution, and the SIMO model for detailed subband deconvolution.

Fig. 6.3: Sample color images.

Table 6.1: Spectral correlation coefficients for different color images.

Images	Red (R) / Green (G)				Blue (B) / Green (G)			
	LL	LH	HL	HH	LL	LH	HL	HH
Boat	0.8317	0.9861	0.9875	0.9858	0.9864	0.9949	0.9945	0.9873
Cap	0.7119	0.9711	0.9695	0.9307	0.5425	0.964	0.9543	0.997
Woman	0.8987	0.9006	0.9751	0.9289	0.9401	0.9373	0.9772	0.9119
Satellite	0.8831	0.9966	0.9962	0.9987	0.9521	0.9957	0.995	0.9954
Car	0.7447	0.9554	0.9472	0.9797	0.851	0.978	0.9662	0.9885
Butterfly	0.766	0.9452	0.9289	0.9749	0.6744	0.9555	0.957	0.9766
Lakeside	0.8661	0.9542	0.9448	0.9352	0.8698	0.9693	0.9737	0.9598
Motor	0.8858	0.9865	0.9899	0.9822	0.8921	0.9766	0.9856	0.9682

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

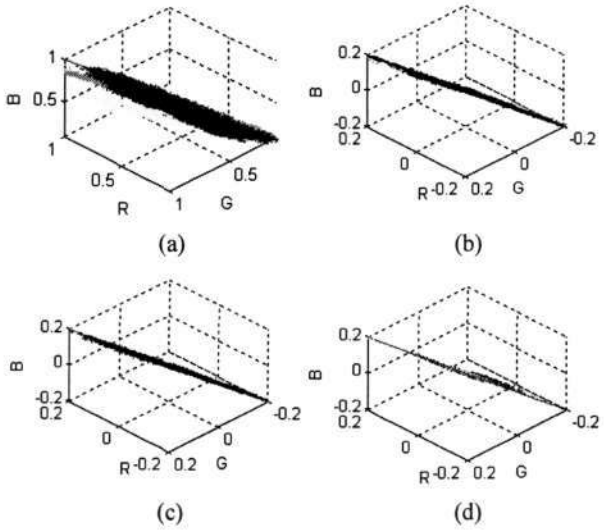


Fig. 6.4: Scatter plots of various subband coefficients (magenta represents the "Boat"; cyan represents the "Fence"; black represents the "Satellite". (a) LL subband, (b) HL subband, (c) LH subband, (d) HH subband.

6.4 Approximate Subband Deconvolution

The LL subband of each color channel is estimated using the SISO model given in Fig. 6.5. We will propose the development of a meaningful cost function in this subsection. The image formation process in SISO model, which can be represented by a vector-matrix form as:

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{n} = \boldsymbol{F}\boldsymbol{h} + \boldsymbol{n} \tag{6.5}$$

where

$$\boldsymbol{g} = \begin{bmatrix} \boldsymbol{g}_R \\ \boldsymbol{g}_G \\ \boldsymbol{g}_B \end{bmatrix}, \boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}_R \\ \boldsymbol{f}_G \\ \boldsymbol{f}_B \end{bmatrix}, \boldsymbol{h} = \begin{bmatrix} \boldsymbol{h}_{RR} \\ \boldsymbol{h}_{GG} \\ \boldsymbol{h}_{BB} \end{bmatrix}, \boldsymbol{n} = \begin{bmatrix} \boldsymbol{n}_R \\ \boldsymbol{n}_G \\ \boldsymbol{n}_B \end{bmatrix}, \boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_{RR} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{H}_{GG} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{H}_{BB} \end{pmatrix}, \boldsymbol{F} = \begin{pmatrix} \boldsymbol{F}_R & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{F}_G & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{F}_B \end{pmatrix} \tag{6.6}$$

$\boldsymbol{f}_i$, $\boldsymbol{y}_i$, $\boldsymbol{n}_i$, and $\boldsymbol{h}_{ii}$ $i \in \{R, G, B\}$ are the vector representation of the i-th channel of original image, blurred image, additive noise, and intrachannel PSF, respectively. $\boldsymbol{H}_{ii}$ and $\boldsymbol{F}_i$ are the corresponding matrices constructed from h_{ii} and f_i [137].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

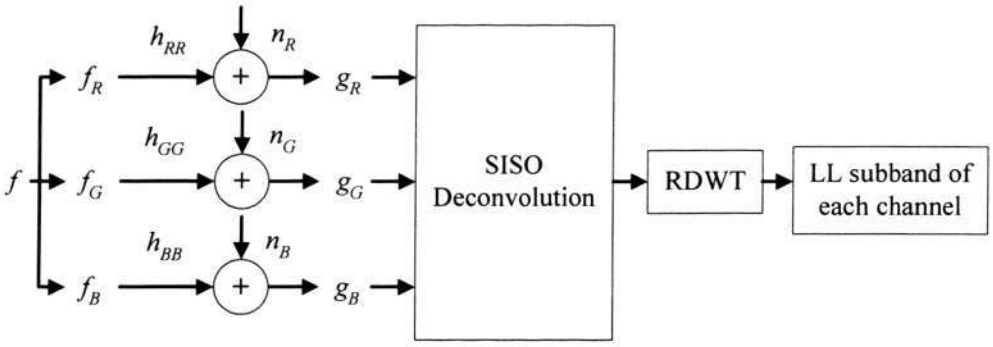


Fig. 6.5: The schematic diagram of SISO model for LL subband deconvolution.

The proposed cost function is given as follows:

$$A(\boldsymbol{f},\boldsymbol{h}) = \frac{1}{2}\|\boldsymbol{g}-\boldsymbol{H}\boldsymbol{f}\|_2^2 + U(\boldsymbol{f},\boldsymbol{h}) = \frac{1}{2}\|\boldsymbol{g}-\boldsymbol{F}\boldsymbol{h}\|_2^2 + U(\boldsymbol{f},\boldsymbol{h}) \tag{6.7}$$

The first term in (6.7) represents the data fidelity to the observed data. $U(\boldsymbol{f},\boldsymbol{h})$ is a regularization functional that introduces stability into the solution by imposing smoothness constraints on the image and PSFs. We formulate a spatially adaptive regularization scheme by introducing the function below:

$$U(\boldsymbol{f},\boldsymbol{h}) = \frac{1}{2}\left\|\sqrt{\boldsymbol{\Lambda}}\boldsymbol{Q}\boldsymbol{f}\right\|_2^2 + \frac{1}{2}\left\|\sqrt{\boldsymbol{\Psi}}\boldsymbol{E}\boldsymbol{h}\right\|_2^2 \tag{6.8}$$

where the matrices Λ and Ψ are space-adaptive weighting terms, $\boldsymbol{E}$ is the 2D Laplacian high-pass operator in the blur domain and $\boldsymbol{Q}$ is the 3D weighted Laplacian operator in the image domain as follows:

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{L} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{L} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{L} \end{pmatrix}, \quad \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{L}_3 & \boldsymbol{C}_{RG} & \boldsymbol{C}_{RB} \\ \boldsymbol{C}_{GR} & \boldsymbol{L}_3 & \boldsymbol{C}_{BG} \\ \boldsymbol{C}_{BR} & \boldsymbol{C}_{BG} & \boldsymbol{L}_3 \end{pmatrix} \tag{6.9}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

where $\boldsymbol{L}$ is the conventional 2D Laplacian matrix, $\boldsymbol{L}_3$ is the 2D Laplacian operator constructed from (6.10), and $\boldsymbol{C}_{ij}$ is given as $\boldsymbol{C}_{ij} = diag\{c_i(1,1,j), c_i(1,2,j), \cdots, c_i(M_f, N_f, j)\}$, where $c_i(x,y,j)$ is defined by (6.11) and $M_f \times N_f$ is the dimension of the color image. The Laplacian operator $\boldsymbol{E}$ imposes smoothing constraint on the PSFs. The operator $\boldsymbol{Q}$ is a 3D weighted Laplacian operator that is used in the regularization functional to introduce smoothness constraint in the image domain. The rationale for the choice of $\boldsymbol{Q}$ is to ensure smoothness within a small neighborhood: (i) of each color channel and (ii) across different channels. The operator $\boldsymbol{Q}$ is essentially the matrix representation constructed from the equivalent 3D filter mask c_i below, where the subscript i represents the mask filtering on the i-th channel:

$$\begin{aligned} &c_i(x,y,i) = 6, \\ &c_i(x-1,y,i) = c_i(x+1,y,i) = c_i(x,y-1,i) = c_i(x,y+1,i) = -1, \quad i \in \{R,G,B\} \end{aligned} \tag{6.10}$$

$$c_i(x,y,j) = -\frac{\|f_i(\Omega)\|_2}{\|f_j(\Omega)\|_2} \quad j \in \{R,G,B\} \ \ (j \neq i) \tag{6.11}$$

where (x,y,i) denote the coordinate (x,y) on the i-th channel and $c_i(\bullet)$ are the coefficients of the mask operating on the pixel at the position $(\bullet)$. Ω is the $p \times q$ neighborhood support. The filter coefficients $c_i(x,y,i)$, $c_i(x-1,y,i)$, $c_i(x+1,y,i)$, $c_i(x,y-1,i)$ and $c_i(x,y+1,i)$ will ensure smoothness within a small neighborhood of each color channel. On the other hand, the coefficients $c_i(x,y,j)$ try to impose smoothness between the i-th and j-th channels. The coefficient $c_i(x,y,j)$ is computed as the ratio of the average intensity between the i-th and j-th channels. This is to take into account that even for smooth regions, the average intensity for different color channel may be different.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

In addition, we have employed the matrices Λ and Ψ in the regularization term as

$$\Lambda = \begin{pmatrix} \lambda_R \boldsymbol{I} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \lambda_G \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \lambda_B \boldsymbol{I} \end{pmatrix}, \Psi = \begin{pmatrix} \beta_R \boldsymbol{I} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \beta_G \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \beta_B \boldsymbol{I} \end{pmatrix} \tag{6.12}$$

Here λ_i and β_i are the image- and blur-domain regularization parameters that control the trade-off between the data fidelity term and the image- and blur-domain regularization terms. We adopt the alternating minimization (AM) technique with conjugate gradient (CG) optimization to minimize the cost function in (6.7). The details of the optimization procedure will be further discussed in Section 6.6.

6.5 Detailed Subband Deconvolution

6.5.1 Blind Deconvolution under SIMO Modeling

In line with the study in Section 6.3.2, we assume that all channels are approximately equal at each high-frequency subband due to the high correlation, namely, $f_{DB} = f_{R\,(DB)} = f_{G\,(DB)} = f_{B\,(DB)}$ where $DB = \{LH, HL, HH\}$ are the detailed subbands. This assumption will simplify the original MIMO model of blurred color images into three SIMO models, so that multichannel blind image restoration algorithm can be employed.

The DB subbands of each color channel are estimated using the SIMO model as given in Fig. 6.6. Following this new model and considering (6.3), we can deduce that:

$$\begin{cases} h_{R\,DB} \approx h_{RR} + h_{GR} + h_{BR}; \\ h_{G\,DB} \approx h_{RG} + h_{GG} + h_{BG}; \\ h_{B\,DB} \approx h_{RB} + h_{GB} + h_{BB}; \end{cases} \tag{6.13}$$

By performing wavelet decomposition and exploiting the observation that the high-frequency

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

profiles of RGB channels are almost identical, we can reduce the MIMO model in (6.1) into three SIMO structures, corresponding to the RGB channels. This allows us to employ SIMO algorithms to estimate the high-frequency subbands of the color image.

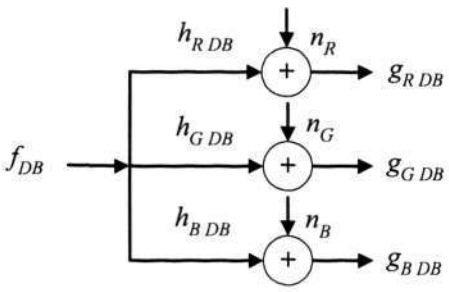


Fig. 6.6: The schematic diagram of SIMO models for HL, LH, and HH subbands estimation.

Mutichannel grayscale blind image restoration using the SIMO model has been studied in recent years. Some promising results have been achieved using eigenvector-based algorithm (EVAM) by Harikumar and Bresler [122]. However, their method is sensitive to noise. Thus, Sroubek and Flusser propose an algorithm that combines the TV technique with the EVAM in order to offer a solution that is more robust towards noise. However, it should be noted that, to the best of our knowledge, no previous works have applied these MC grayscale algorithms on color images due to the mismatch between the SIMO model of MC grayscale problem and the MIMO model of blind color deconvolution. In this work, we have proposed a framework to extend the MC grayscale method in [124] to address blind color image deconvolution. Based on the discussion above, the cost function using the SIMO model for detailed subband deconvolution is given as:

$$A(\boldsymbol{f}_{DB},\boldsymbol{h}_{DB})=\frac{1}{2}\|\boldsymbol{g}_{DB}-\boldsymbol{H}\boldsymbol{f}_{DB}\|_2^2+U_1(\boldsymbol{f}_{DB})+U_2(\boldsymbol{h}_{DB})=\frac{1}{2}\|\boldsymbol{g}_{DB}-\boldsymbol{F}\boldsymbol{h}_{DB}\|_2^2+U_1(\boldsymbol{f}_{DB})+U_2(\boldsymbol{h}_{DB}) \quad (6.14)$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

where

$$\boldsymbol{g}_{DB}=\begin{bmatrix}\boldsymbol{g}_{R\,DB}\\ \boldsymbol{g}_{G\,DB}\\ \boldsymbol{g}_{B\,DB}\end{bmatrix},\boldsymbol{h}_{DB}=\begin{bmatrix}\boldsymbol{h}_{R\,DB}\\ \boldsymbol{h}_{G\,DB}\\ \boldsymbol{h}_{B\,DB}\end{bmatrix},\boldsymbol{H}=\begin{pmatrix}\boldsymbol{H}_{R\,DB}\\ \boldsymbol{H}_{G\,DB}\\ \boldsymbol{H}_{B\,DB}\end{pmatrix},\boldsymbol{F}=\begin{pmatrix}\boldsymbol{F}_{DB} & \boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{F}_{DB} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{F}_{DB}\end{pmatrix} \tag{6.15}$$

The first term in (6.14) is the least-square data fidelity term. The second and third terms are the image-domain and blur-domain regularization terms, respectively. $\boldsymbol{g}_{DB}$ and $\boldsymbol{h}_{DB}$ are decomposed into the RGB channels. $\boldsymbol{H}_{i\,DB}$ and $\boldsymbol{F}_{DB}$ are the appropriate matrices constructed from $\boldsymbol{h}_{i\,DB}$ and $\boldsymbol{f}_{DB}$. In agreement with general blind MC image deconvolution, the blurs $\boldsymbol{h}_{R\,DB}$, $\boldsymbol{h}_{G\,DB}$ and $\boldsymbol{h}_{B\,DB}$ should be weakly coprime [122], [124], [161]. In most cases, this assumption can be satisfied. Although the optimization problem is formulated similar to (6.7), there are some differences in terms of the regularization schemes, in particular, with respect to the detailed image subbands and the multichannel PSFs, respectively. We will further discuss this issue in the next section. Once the detailed subbands of each channel have been estimated, we will combine them with the estimated LL subband, and synthesize them to produce the final restored color image.

6.5.2 Regularization Scheme

It is well known that MC blind image deconvolution is an ill-posed problem. TV technique ([136], [162] and [163]) discussed in Chapter 3 is adopted to suppress the noise while preserving the information of high-frequency subbands. The chosen regularization scheme is given by:

$$U_1(\boldsymbol{f}_{DB})=\frac{1}{2}\gamma\int_{\Omega}|\nabla\boldsymbol{f}_{DB}| \tag{6.16}$$

where γ is the regularization parameter, $\nabla\boldsymbol{f}_{DB}$ is the local gradient of $\boldsymbol{f}_{DB}$, and Ω is the image

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

support.

Next we will discuss the regularization scheme of the PSFs. To develop the regularization term for the PSFs, we observe that the following condition will be satisfied, in the absence of noise:

$$g_{i\,DB}(x,y)\otimes h_{j\,DB}(x,y)-g_{j\,DB}(x,y)\otimes h_{i\,DB}(x,y)=O \quad i,j=R,G,B \tag{6.17}$$

Therefore the condition can be transformed into minimization of a constraint term:

$$\text{Minimize: } \sum_{(x,y)\in\Omega}\sum_{\substack{i\neq j\\ i,j=R,G,B}}\left[g_{i\,DB}(x,y)\otimes h_{j\,DB}(x,y)-g_{j\,DB}(x,y)\otimes h_{i\,DB}(x,y)\right]^2 \tag{6.18}$$

The constraint term can then be expressed as a regularization term in the vector-matrix form, $U_2(\boldsymbol{h}_{DB})$ is given as:

$$U_2(\boldsymbol{h}_{DB})=\frac{1}{2}\zeta\left\|\boldsymbol{R}\boldsymbol{h}_{DB}\right\|_2^2 \tag{6.19}$$

where $\boldsymbol{R}$ is constructed as follows and $\boldsymbol{G}_{i\,DB}$ is the convolution matrix formed by the blurred image $g_{j\,DB}$.

$$\boldsymbol{R}=\begin{pmatrix}\boldsymbol{G}_{G\,DB} & -\boldsymbol{G}_{R\,DB} & \boldsymbol{O}\\ \boldsymbol{G}_{B\,DB} & \boldsymbol{O} & -\boldsymbol{G}_{R\,DB}\\ \boldsymbol{O} & \boldsymbol{G}_{B\,DB} & -\boldsymbol{G}_{G\,DB}\end{pmatrix} \tag{6.20}$$

6.6 Optimization Procedure

In this Section, we will describe the minimization of the cost functions corresponding to approximate subband deconvolution in (6.7), and detailed subband deconvolution in (6.14). The overall cost function that is used to estimate the original image and the blurring function is

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

non-convex. This is because the cost function comprises the image-domain and blur-domain terms. However, the image- and blur-domain cost functions are convex quadratic functions with positive semidefinite Hessian matrices. This implies that the minima of the projected cost function in each domain can be reached. Various studies such as [157], [164] and [165] have demonstrated the effectiveness of AM in addressing blind image restoration.

6.6.1 Alternating Minimization in LL Subband Deconvolution

The optimization procedure used to perform LL subband deconvolution in (6.7) can be summarized as follows:

(i). Initialization:

Initialize $f_i^0(x,y) = g_i(x,y)$ and $h_{ii}^0(x,y)$ to delta functions, where $i = R,G,B$.

(ii). At the m-th iteration:

a) Minimize the image-domain cost function to estimate the image:

$$\boldsymbol{f}^m = \arg\min_{\boldsymbol{f}} \left(\frac{1}{2} \left\| \boldsymbol{g} - \boldsymbol{H}^{m-1} \boldsymbol{f} \right\|_2^2 + \frac{1}{2} \left\| \sqrt{\Lambda} \boldsymbol{Q} \boldsymbol{f} \right\|_2^2 \right) \tag{6.21}$$

This is equivalent to solving for $\boldsymbol{f}^m$ using the following linear equation:

$$\left[(\boldsymbol{H}^{m-1})^T \boldsymbol{H}^{m-1} + (\sqrt{\Lambda}\boldsymbol{Q})^T (\sqrt{\Lambda}\boldsymbol{Q}) \right] \boldsymbol{f}^m = (\boldsymbol{H}^{m-1})^T \boldsymbol{g} \tag{6.22}$$

b) Minimize the blur-domain cost function to estimate the blur:

$$\boldsymbol{h}^m = \arg\min_{\boldsymbol{h}} \left(\frac{1}{2} \left\| \boldsymbol{g} - \boldsymbol{F}^m \boldsymbol{h} \right\|_2^2 + \frac{1}{2} \left\| \sqrt{\Psi} \boldsymbol{E} \boldsymbol{h} \right\|_2^2 \right) \tag{6.23}$$

This is equivalent to solving for $\boldsymbol{h}^m$ using the following linear equation:

$$\left[(\boldsymbol{F}^m)^T \boldsymbol{F}^m + (\sqrt{\Psi}\boldsymbol{E})^T (\sqrt{\Psi}\boldsymbol{E}) \right] \boldsymbol{h}^m = (\boldsymbol{F}^m)^T \boldsymbol{g} \tag{6.24}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(iii). Stop if convergence or a maximum number of iterations is reached; otherwise, go to (ii).

(iv). Obtain the deblurred LL subband by applying the analysis filter on the reconstructed channel.

We adopt the CG optimization to solve (6.22) and (6.24). The CG optimization has been discussed in detail in Chapter 3.

6.6.2 Alternating Minimization in Detailed Subband Deconvolution

The optimization procedure used to perform detailed subband deconvolution in (6.14) can be summarized as follows:

(i). Initialization:

Initialize $\boldsymbol{f}_{DB}^{0}(x,y) = \boldsymbol{g}_{i\,DB}(x,y)$ and $h_{i\,DB}^{0}(x,y)$ to delta functions, where $i = R,G,B$.

(ii). At the n-th iteration:

a) Minimize the image-domain cost function to estimate the image:

$$\boldsymbol{f}_{DB}^{n} = \arg\min_{\boldsymbol{f}_{DB}}\left(\frac{1}{2}\left\|\boldsymbol{g}_{DB} - \boldsymbol{H}^{n-1}\boldsymbol{f}_{DB}\right\|_2^2 + \frac{\gamma}{2}\int_{\Omega}|\nabla \boldsymbol{f}_{DB}|\right) \tag{6.25}$$

This is equivalent to solving for $\boldsymbol{f}_{DB}^{n}$ using the following equation:

$$\left((\boldsymbol{H}^{n-1})^{T}\boldsymbol{H}^{n-1} + \gamma \boldsymbol{T}\right)\boldsymbol{f}_{DB}^{n} = (\boldsymbol{H}^{n-1})^{T}\boldsymbol{g}_{DB} \tag{6.26}$$

where $\boldsymbol{T}$ can be constructed by using the method described in Chapter 3.

b) Minimize the blur-domain cost function to estimate the blur:

$$\boldsymbol{h}_{DB}^{n} = \arg\min_{\boldsymbol{h}_{DB}}\left(\frac{1}{2}\left\|\boldsymbol{g}_{DB} - \boldsymbol{F}^{n}\boldsymbol{h}_{DB}\right\|_2^2 + \frac{\zeta}{2}\left\|\boldsymbol{R}\boldsymbol{h}_{DB}\right\|_2^2\right) \tag{6.27}$$

This is equivalent to solving for $\boldsymbol{h}_{DB}^{n}$ using the following linear equation:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$\left[(\boldsymbol{F}^n)^T \boldsymbol{F}^n + \zeta \boldsymbol{R}^T \boldsymbol{R}\right] \boldsymbol{h}_{DB}^n = (\boldsymbol{F}^n)^T \boldsymbol{g}_{DB} \tag{6.28}$$

(iii). Stop if convergence or a maximum number of iterations is reached; otherwise, go to (ii).

The optimization procedure of the cost function in the SIMO model is similar to that of the SISO model. Finally, based on the framework discussed in Section 6.2, we could determine the restored color image by applying the synthesis filters on all the estimated subbands obtained from the SISO and SIMO deconvolution.

The main differences between the proposed method and the double-regularization SISO (DR-SISO) approach in [135] can be summarized as follows:

(i). The proposed method takes interchannel blurring into consideration, while the DR-SISO method ignores the interchannel blurring.

(ii). The proposed method utilizes the correlation information of high-frequency components between different color channels to formulate SIMO deconvolution. The DR-SISO method does not exploit this information.

We will use the DR-SISO method as the baseline for comparison of computational complexity. As the proposed restoration scheme needs to handle four subband deconvolution: low-frequency (LL) subband and high-frequency (LH, HL and HH) subbands, the computational complexity of our algorithm is nearly four times that of the DR-SISO method. Since the exact number of arithmetic operations is difficult to determine, we will use experimental simulation time to demonstrate the computational complexity. The simulation environments are as follows: Windows XP, MATLAB 7.1, CPU P4-3.4 GHz, and 1G RAM. The AM iteration was set to 10 for the experiment. The computational time for our method with an average test image of dimension 256×256 is 43s. The average computation time of the DR-SISO method is 11s. This agrees with our earlier analysis. Nevertheless, it should be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

pointed out that the proposed method offers superior reconstructed color image when compared with the DR-SISO method, as shown in the following experiments.

6.7 Experiments

In this Section, we demonstrate the effectiveness of the proposed algorithm by performing blind deconvolution on various color images under different environments. For approximate subband (LL) deconvolution, a SISO image deconvolution algorithm is used on blurred RGB channels. The deblurred LL subband of each channel can be obtained by performing wavelet decomposition on each deblurred color channel. For detailed subbands (HL, LH, HH) deconvolution, the blurred image was firstly decomposed into the respective RGB channels. Then, they underwent wavelet decomposition to produce the LH, HL, and HH wavelet subbbands for each channel. An undecimated wavelet transform with the low-pass filter $a_{lp}(.)=\frac{1}{4}[1\ \ 2\ \ 1]$ and high-pass filter $a_{hp}(.)=\frac{1}{4}[1\ \ -2\ \ 1]$ was chosen. We performed detailed subband deconvolution using SIMO modeling. After their respective deconvolution using the SISO and SIMO models, the LL, LH, HL, and HH subbands of each channel were then put together and synthesized using the synthesis filters $s_{lp}(.)=\frac{1}{8}[-1\ \ 2\ \ 6\ \ 2\ \ -1]$ and $s_{hp}(.)=\frac{1}{8}[1\ \ 2\ \ -6\ \ 2\ \ 1]$ to render the final restored image.

6.7.1 Blind Color Deconvolution on Synthesized Color Image

To demonstrate the effectiveness of the proposed method, we first use a synthesized color image as shown in Fig. 6.7 (a). The spectral correlation for subband coefficients in Table 6.2 is obtained using the method in Section 6.3.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table 6.2: Spectral correlation coefficients for the synthesized color image

Red (R) / Green (G)				Blur (B) / Green (G)			
LL	LH	HL	HH	LL	LH	HL	HH
0.7731	0.9791	0.9771	0.9861	0.7475	0.9565	0.9559	0.9671

In this experiment, the RGB channels of the color image were blurred by intrachannel and interchannel PSFs. The chosen intrachannel PSFs for RGB channels were Gaussian blurs with standard deviation of $\sigma = 2.0,\ 2.5,\ 3.0$ and with support $3\times3,\ 5\times5,\ 7\times7$ respectively. The interchannel PSFs consisted of 7×7 Gaussian blur with variance of 1.5. The ratio of magnitude for interchannel to intrachannel degradation was set to 5% in this experiment. The blurred color image is shown in Fig. 6.7(b). We performed blind color image deconvolution using the proposed algorithm, and compared the results with those obtained using the DR-SISO method in [135]. In this chapter, we selected the regularization parameters from the experience accumulated over experiments. First, we followed the algorithm [124] to estimate an order-of-magnitude of the regularization parameters. After fine-tuning within the estimated order of magnitude, the best regularization parameters were chosen when they produce visually appealing results. The simulation results show that the algorithm is robust towards different regularization parameters so long as they fall within a reasonable range. To ensure fairness, we also tried to estimate the regularization parameters by using our method to obtain the optimal result for the DR-SISO method. We adopted the following values for the regularization parameters in this experiment: $\lambda_i=10^{-4}$, $\beta_i = 10^7$ in the SISO deconvolution, $\gamma=10^{-6}$, $\zeta=10^5$ in the SIMO deconvolution for the HH subbands and $\gamma=10^{-5}$ for the LH and HL subbands.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

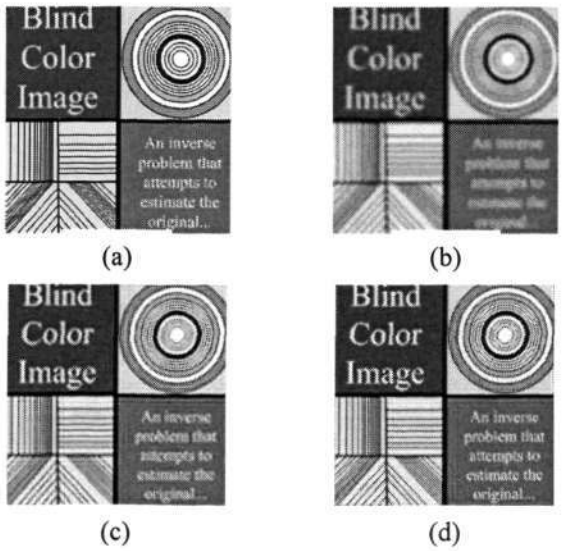


Fig. 6.7: Blind color image deconvolution on synthesis color image. (a) Original image, (b) Blurred image, (c) Restored image using the DR-SISO method [135], (d) Restored image using the proposed method.

The restored color images by the DR-SISO method and the proposed method are shown in Fig. 6.7(c) and Fig. 6.7(d), respectively. When compared with the DR-SISO method, it is observed that the proposed method offers clearer edges near the circles, lines and words. The subjective observation is confirmed by objective performance measures where the proposed method renders an PSNR=18.5 as compared to an PSNR=17.1 by the DR-SISO method. The proposed method is able to outperform the DR-SISO method as it utilizes the correlation information among different color channels.

6.7.2 Blind Color Deconvolution on Natural Images with Additive Noise

We illustrate the capability of the proposed method to handle noisy degraded color images in this section. The images used in this experiment are shown in Fig. 6.8. The blurs were used as same as the previous experiment. A 30dB SNR AWGN was added to each color image channel. In this experiment, we adopted the following regularization parameters in the image domain:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$\lambda_i = 0.005$ in the SISO deconvolution, and $\gamma=10^{-5}$ for the HH subband in the SIMO deconvolution and $\gamma=10^{-4}$ for the LH and HL subbands. The final results using the proposed and DR-SISO methods are tabulated in Table 6.3. From the results, it is observed that the $NMSE(\boldsymbol{h})$ obtained using both methods are similar because the intrachannel blurs are estimated using the SISO deconvolution in both cases. However, it should be noted that the proposed method takes the interchannel degradation into consideration in the detailed subband restoration. This is given by $NMSE(\boldsymbol{h}_{DB})$ in the table. The superiority of our method is further confirmed by the $PSNR(\boldsymbol{f})$ in the table. Further, the "Circuit board" image was used to illustrate the quality of the restored images in Fig. 6.9. It can be seen that the restored image using the proposed method has recovered the overall sharpness of the image. Comparison with the result obtained using the DR-SISO method indicates the superiority of the proposed method.

Fig. 6.8: Images used in the experiment.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(a) (b) (c)

Fig. 6.9: Blind color image deconvolution on Gaussian blurred color image with additive noise. (a) Noisy blurred image, (b) Restored image using the DR-SISO method [135], (c) Restored image using the proposed method.

Table 6.3: Performance of blind color image deconvolution on Gaussian blurred color images with additive noise.

Image	Different Gaussian PSF (30dB)				
	DR-SISO		Our proposed method		
	NMSE($\boldsymbol{h}$)	PSNR(f)	NMSE($\boldsymbol{h}$)	NMSE($\boldsymbol{h}_{DB}$)	PSNR(f)
Boat	0.0168	24.17	0.0142	0.0116	26.36
Fence	0.0174	26.89	0.0169	0.012	28.72
Woman	0.0251	25.61	0.0249	0.0203	26.69
Satellite	0.0213	20.28	0.0207	0.0125	22.45
Tilehouse	0.0221	21.36	0.0211	0.009	24.68
Circuit Board	0.0196	14.38	0.0173	0.0148	16.91

6.7.3 Blind Color Image Deconvolution on Real-life Image

A real-life "bookshelf" image is captured by a web camera, while in-focus is set to be the ground truth. It is shown in Fig. 6.10 (a). Next, we captured an out-of-focus image, and the blurred color image is given in Fig. 6.10(b). It is observed that the image suffers from blurring, especially near the text region. We run the proposed method on the blurred real-life image, and the restored color image is given in Fig. 6.10(d). From the figure, it can be seen that the considerable clarity of the image has been recovered. There is also no noticeable ringing and color artifact in the restored image. Further, the result obtained using the DR-SISO method is given in Fig. 6.10(c). Comparing the results obtained using both methods with the ground truth,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

it is shown that the proposed approach is superior in handling real-life blind color image deconvolution. It is able to recover more visual clarity from the blurred image, particularly near the text region.

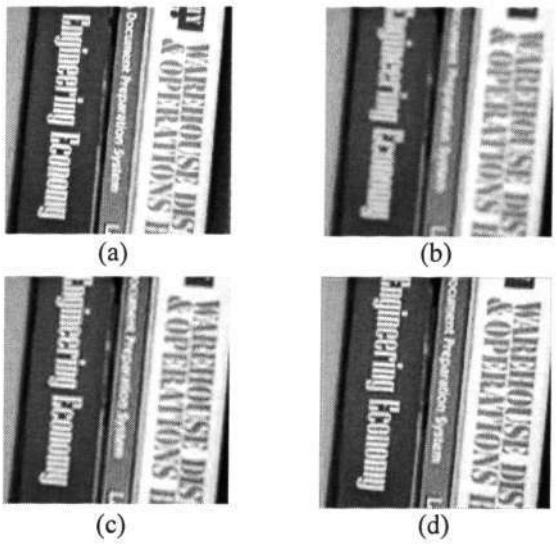


Fig. 6.10: Blind color image deconvolution on real-life image. (a) Ground truth, (b) Blurred image, (c) Restored image using the DR-SISO method [135], (d) Restored image using the proposed method.

6.8 Summary

In this chapter, we present a new algorithm to address blind color image deconvolution based on a hybrid framework of SISO and SIMO models. We emphasize the importance of exploiting both spatial and spectral correlations in color image deconvolution and use a combination of SISO and SIMO blind algorithms to solve the problem. In the SISO model, we develop a full-channel alternating minimization scheme to perform deconvolution. In the SIMO model, the dominant interchannel spectral correlation is exploited to perform deconvolution. Experimental results show that the proposed method is able to provide good restored color images where there is limited information about the blurring functions.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In many computer vision applications, the imaging system can see the world by various sensors, such as still or video cameras. The resolution of image is crucial for the performance of the system, such as license plate readers or surveillance systems. This thesis investigates the techniques for HR image reconstruction through image SR and deconvolution algorithms. In the following subsections, we summarize the work of this thesis in three parts: multi-frame SR, single-frame face SR and color image deconvolution.

7.1.1 Multi-frame Super-resolution

This thesis has addressed the multi-frame SR from two points of view: blur identification and motion estimation.

A main issue arising from the multi-frame SR problem is the availability of the knowledge on the blurring occurred during the image formation process. In many real-world applications, it is often difficult to know the blurring completely *a priori*. This motivates the study of blind image SR to incorporate blur identification into the SR algorithms. In this thesis, a new soft MAP estimation framework has been developed to address this problem. As opposed to most disjoint blind SR methods that performing SR reconstruction followed by blind image deconvolution on the blurred HR image, the new algorithm offers a joint process of blur

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

identification and HR image reconstruction. The MAP framework based on alternating minimization (AM) estimates the blur and the HR image iteratively. This provides a unified framework where a more reliable blur identification from the HR image can be achieved, which in turn renders a better HR image reconstruction. In addition, as opposed to the conventional notion of hard decision for the blurs, the proposed algorithm employs a soft blur learning scheme to integrate the parametric information of the blurring function into the algorithm. The best-fit parametric model of the blurring function can be estimated and incorporated into the blur prior to induce reinforcement learning towards it. The scheme can be tailored to include other blur types if some prior parametric knowledge of the blur is available. Experimental results show that the new method is effective in performing blind image SR where there is limited information about the blurring functions.

Another critical step in multi-frame SR is accurate registration, i.e. motion estimation, of the LR images. Unlike conventional SR works that focus on translational motion only, we adopt a more generic motion model that includes both translation as well as rotation in Chapter 4. The adopted motion model is more realistic. As opposed to the current two-stage SR methods that perform registration on the LR images the image registration in the proposed joint framework is performed iteratively using the progressively estimated HR image. This is promising as more accurate motion parameters can be determined, thereby enhancing the performance of the HR reconstruction. Further, the new method can overcome the shortcoming of the iterative AM framework, as the cost function is not projected onto the image and motion parameter domains one at a time. Instead, an iterative scheme based on a nonlinear least squares method is developed to estimate the motion parameters and the HR image simultaneously. Moreover, it is noted that the problem in this joint scheme is no longer linear with respect to some motion parameters. We solve this issue by deriving the Jacobian matrix, which enable the nonlinear least squares method to be used in the joint SR framework. Experimental results have shown

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

that the proposed method is effective in performing image registration and SR for simulated as well as real-life images.

7.1.2 Single-frame Face Super-resolution

Unlike multi-frame SR that performs HR image reconstruction by fusing multiple LR images, single-frame SR algorithm estimates missing HR details from a single observed LR image by learning from training examples. In this thesis, we focus our attention on face problem domain. Different from generic image SR, the common characteristics of a human face is essential for face image SR. Single-frame face SR problem is challenging, as there is less observed data available and we have to take into account the specific features of human faces.

In this thesis, we have developed a new single-frame face SR method by using the techniques: PCA and LLE. An improved PCA-based method is used to synthesis a global HR face, followed by a patch-based residue prediction step for compensating detailed facial features. Compared with the conventional PCA-based method, the new method can be applied to new data more effectively, and a better HR image that bears closer resemblance to the original face while preserving usual characteristics of a human face. Furthermore, the estimation of the residual patch does not depend on a single nearest neighbor in the training set. Instead, it depends simultaneously on multiple nearest neighbors in a way similar to LLE for manifold learning. Experimental results have shown the effectiveness of the proposed method.

7.1.3 Color Image Deconvolution

Image deconvolution aims to restore a blurred and noisy image, without increasing the pixel resolution. Blind color image deconvolution has been investigated in this thesis. It is considered as a MIMO deconvolution problem, where intrachannel and interchannel blurring are taken into consideration. We have presented a new algorithm to address blind color image deconvolution

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

based on a hybrid framework of SISO and SIMO models. We emphasize the importance of exploiting both spatial and spectral correlations in color image deconvolution and use a combination of SISO and SIMO blind algorithms to alleviate the intractable original MIMO problem. In the SISO model, we develop a full-channel alternating minimization scheme to perform deconvolution. In the SIMO model, the dominant interchannel spectral correlation is exploited to perform deconvolution. Compared with classical color image deconvolution methods, the new method incorporates the blur identification into image deconvolution. Experimental results have shown that the proposed method is able to provide good restored color images where there is limited information about the blurring functions.

7.2 Future Work

Based on the research in this thesis, we point out the potential directions for future research in this section. The future of HR reconstruction technology appears to be bright, although much work remains to be done to reach a mature technology.

7.2.1 Multi-frame Face Super-resolution

Studies in this thesis have shown that the proposed multi-frame generic image SR and learning-based single-frame face SR methods are able to provide satisfactory HR results under certain conditions. So an intuitive extension of the current work is a multi-frame face SR when multiple LR face observations are available. If we only apply the multi-frame generic image SR method on the LR faces, the reconstructed HR result may neglect the common characteristics of human faces, since the *prior* of the HR image is based on generic images. If we had more specific information about the target image, our estimate could be better. Further, it has been shown in [113] and [166] that fundamental limits for the reconstruction-based multi-frame SR

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

are bounded with respect to the magnification factors under different conditions. Beyond the limits, the reconstructed HR image may deviate significantly from the ground truth. On the other hand, if we only apply the single-frame face SR method on single-face frame, we may leave the available information of other frames unused. In view of this, it is advantageous to combine these two frameworks to yield a hybrid SR reconstruction technique, thereby gaining the benefits of both algorithms. In addition, one possible extension of face SR is the employment of some adaptive bilateral filtering works in the post-processing step. It is observed that the artifacts of the result by the proposed face SR method mainly locate around the face boundary. Thus, it is desirable to use the location-adaptive bilateral filtering algorithm to improve the final result.

7.2.2 Color Image Super-resolution

In this thesis, SR problem has been investigated mainly for grayscale images. One possible direction for future research is to investigate the SR problem for color images. As discussed in color image deconvolution, the color channels (RGB) are not independent but rather highly correlated. Therefore, how to fully exploit this property for color image SR is of interest. Besides that, it is noted that the color digital cameras also suffer from color mosaicking [159], [167]. It is known that most color imaging device use a single CCD equipped with a CFA (commonly using Bayer pattern) to acquire color images. At the location of each pixel only one color sample is taken. To interpolate the other colors from neighboring samples, it is often referred to as "*CFA demosaicking*". Merely applying a simple CFA demosaicking algorithm prior to SR may lead to suboptimal performance. Recently, there is a growing interest in a joint framework by merging SR and CFA demosaicking into one single process [168], [169]. However, to the best of our knowledge, there is still limited work addressing the problem of estimating motion between Bayer filtered images for color image SR, which remains to be fully

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

analyzed. Therefore, we believe that the future research in this area will be beneficial for color image SR.

7.2.3 Video Super-resolution

Image SR aims to reconstruct a single still HR image from a set of LR images or a single LR image, while video SR tries to reconstruct a HR video sequence from a LR video sequence. The SR techniques described in this thesis may be applied to video SR by using a shifting window of processed frames [7]. The "*sliding window*" determines the set of LR frames to be processed to produce the output HR frame at each time. Then the window is moved forward to produce successive HR frames in the output sequence. However, the disadvantage of this method is that the reconstruction of the HR frames in the sequence is independent. In other words, it does not make use of the previously estimated HR frame. In addition, the computational requirement for the sliding window-based method is high. Recently, a few papers have suggested different recursive estimation methods for video SR [170], [171], [172] by employing the temporal evolution of the process. Therefore, an interesting extension of the research in this thesis includes the incorporation of the recursive estimation algorithms in video SR.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Author's Publications

Journal Papers:

[1] K.-H. Yap, Yu He, Yushuang Tian, and L.-P. Chau, "A nonlinear L_1-norm approach for joint image registration and super-resolution" accepted by *IEEE Signal Processing Letters,* 2009.

[2] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "A soft MAP framework for blind super-resolution image reconstruction," *Image and Vision Computing*, vol. 27, pp. 364-373, Mar. 2009.

[3] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "A novel hybrid model framework to blind color image deconvolution," *IEEE Transaction on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 38, pp. 86-880, July 2008.

[4] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "A nonlinear least square technique for simultaneous image registration and super-resolution," *IEEE Transaction on Image Processing,* vol. 16, pp. 2830-2841, Nov. 2007.

[5] Li Chen, K.-H. Yap and Yu He, "Subband synthesis for color filter array demosaicking" *IEEE Transaction on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 38, pp. 485-492, Mar. 2008.

[6] Li Chen, K.-H. Yap, Yu He, "Efficient Recursive Multichannel Blind Image Restoration" *EURASIP Journal on Applied Signal Processing*, Article ID 19675, 2007.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Conference Papers:

[1] Yu He, K.-H. Yap, and L.-P. Chau, "A Learning Approach for Single-Frame Face Super-Resolution", in *proceeding of the IEEE International Symposium on Circuits and System,* , Taipei, Taiwan, May 2009, pp 770-773.

[2] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "A new color image regularization scheme for blind image deconvolution," in *Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, USA, Mar. 2008, pp. 1253-1256.

[3] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "Joint image registration and super-resolution using nonlinear least squares method," in *Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Honolulu, Hawaii, USA, April 2007, vol. 1, pp. 561-564.

[4] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "Blind super-resolution image reconstruction using a maximum a posteriori estimation," in *Proceeding of the International Conference on Image Processing*, Atlanta, USA, Oct. 2006, pp. 1729-1732.

[5] Yu He, K.-H. Yap, Li Chen, and L.-P. Chau, "Blind color image deconvolution based on wavelet decomposition," in *Proceeding of the International Conference on Image Processing*, Genova, Italy, Sep. 2005, pp. 766-769.

[6] Li Chen, K.-H. Yap, Yu He, "Color filter array demosaicking using wavelet-based subband synthesis," in *Proceeding of the International Conference on Image Processing*, Genova, Italy, Sep. 2005, pp. 1002-1005.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Bibliography

[1] K. Aizawa, T. Komatsu, and T. Saito, "A scheme for acquiring very high resolution images using multiple cameras," in *proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, Mar. 1992.

[2] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, pp. 56-65, April 2002.

[3] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Optical Engineering*, vol. 37, pp. 247-260, Jan. 1998.

[4] 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, pp. 1646-1658, Dec. 1997.

[5] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Magazine*, vol. 14, pp. 24 - 41, Mar. 1997.

[6] M. T. Heath, *Scientific Computing: An Introductory Survey*. New York: McGraw-Hill Science Engineering, July 2001.

[7] S. Borman and R. Stevenson, "Spatial resolution enhancement of low resolution image sequences: a comprehensive review with directions for future research," *Laboratory for Image and Signal Analysis, University of Notre Dame, Technical Report*, 1998.

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

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[9] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 47-57, Aug. 2004.

[10] S. Chaudhuri, *Super-resolution Imaging*: Springer, 2001.

[11] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Advances in Computer Vision and Image Processing*, vol. 1, pp. 317-319, 1984.

[12] 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, pp. 1013-1027, June 1990.

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

[14] S. P. Kim and W. Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," in *proceeding of the IEEE International Conference on Acoustics Speech and Signal Processing*, Toronto, Canada, May 1991.

[15] N. K. Bose, H. C. Kim, and H. M. Valenzuela, "Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframes," in *proceeding of the IEEE conference on Acoustics, Speech and Signal Processing*, Minneapolis, Apr. 1993.

[16] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore: Johns Hopkins University Press, 1989.

[17] C. E. Davila, "Efficient recursive total least squares algorithms for FIR adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 42, pp. 268-280, Feb. 1994.

[18] R. S. Prendergast and T. Q. Nguyen, "Improving frequency domain super-resolution via undersampling model," in *proceeding of the IEEE International Conference on Image Processing*, Genova, Italy, Sep. 2005.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[19] K. Aizawa, T. Komatsu, and T. Saito, "Acquisition of very high resolution images using stereo cameras," in *proceeding of the SPIE, Visual Communications and Image Processing*, Boston, MA, Nov. 1991.

[20] P. Vandewalle, S. Susstrunk, and A. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on Applied Signal Processing*, Article ID 71459, 2006.

[21] S. Lertrattanapanich and N. K. Bose, "High resolution image formation from low resolution frames using Delaunay triangulation," *IEEE Transaction on Image Processing*, vol. 11, pp. 1427-1441, Dec. 2002.

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

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

[24] N. R. Shah and A. Zakhor, "Resolution enhancement of color video sequences," *IEEE Transaction on Image Processing*, vol. 8, pp. 879-885, June 1999.

[25] M. Irani and S. Peleg, "Motion analysis for image enhancement resolution, occlusion, and transparency," *Journal of Visual Communication and Image Representation*, vol. 4, pp. 324-335, Dec. 1993.

[26] 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 *proceeding of the IEEE International Conference on Image Processing*, Austin, Texas, Sept. 1994.

[27] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high resolution image from multiple degraded mis-registered low resolution images," in *proceeding of the SPIE, Visual Communications and Image Processing*, Chicago, Sept. 1994.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[28] B. C. Tom, K. T. Lay, and A. K. Katsaggelos, "Multichannel image identification and restoration using the expectation-maximization algorithm," *Optical Engineering*, vol. 35, pp. 241-254, Jan. 1996.

[29] B. C. Tom and A. K. Katsaggelos, "Multi-channel image identification and restoration using the expectation-maximization algorithm," in *proceeding of the SPIE, Applications of Digital Image Processing XVII*, San Diego, July 1994.

[30] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, Nov. 1984.

[31] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, pp. 1621-1633, Dec. 1997.

[32] H. He and L. P. Kondi, "Resolution enhancement of video sequences with simultaneous estimation of the regularization parameter," *Journal of Electronic Imaging*, vol. 13, pp. 586-596, July 2004.

[33] J. Tian and K.-K. Ma, "A MCMC approach for Bayesian super-resolution image reconstruction," in *proceeding of the IEEE International Conference on Image Processing*, Genoa, Italy, Sep. 2005.

[34] J. Tian and K.-K. Ma, "Markov Chain Monte Carlo super-resolution image reconstruction with artifacts suppression," in *proceeding of the IEEE Asia Pacific Conference on Circuits and Systems*, Singapore, Dec. 2006.

[35] J. Tian and K.-K. Ma, "Markov Chain Monte Carlo super-resolution image reconstruction with simultaneous adaptation of the prior image model," in *proceeding of the Advances in Multimedia Information Processing - Pacific Rim Conference on Multimedia (PCM)*, Hangzhou, China, Nov. 2006.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[36] D. Gamerman, *Markov Chain Monte Carlo: stochastic simulation for Bayesian inference*. London: Chapman & Hall, 1997.

[37] J. Tian and K.-K. Ma, "Edge-adaptive super-resolution image reconstruction using a Markov Chain Monte Carlo approach," in *proceeding of the International Conference on Information, Communications and Signal Processing*, Singapore, Dec. 2007.

[38] R. Molina and B. D. Ripley, "Using spatial models as priors in astronomical images analysis," *Journal of Applied Statistics*, vol. 16, pp. 193-206, 1989.

[39] R. Molina, M. Vega, J. Abad, and A. K. Katsaggelos, "Parameter estimation in Bayesian high-resolution image reconstruction with multisensors," *IEEE Transaction on Image Processing*, vol. 12, pp. 1655-1667, Dec. 2003.

[40] M. V. Joshi, S. Chardhuri, and R. Panuganti, "A learning-based method for image super-resolution from zoomed observations," *IEEE Transaction on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 35, pp. 527-536, June 2005.

[41] N. A. Woods and N. P. Galatsanos, "Non-stationary approximate Bayesian super-resolution using a hierarchical prior model," in *proceeding of the IEEE International Conference on Image Processing*, Genova, Italy, Sep. 2005.

[42] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing*, vol. 5, pp. 996-1011, June 1996.

[43] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, pp. 233-242, May 1994.

[44] N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient superresolution image reconstruction algorithm," *IEEE Transactions on Image Processing*, vol. 10, pp. 573-583, April 2001.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[45] G. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, pp. 215-223, 1979.

[46] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Transactions on Image Processing*, vol. 10, pp. 1187-1193, Aug. 2001.

[47] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, pp. 1327-1344, Oct. 2004.

[48] 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, pp. 1715-1726, Nov. 1989.

[49] 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 *proceeding of the IEEE Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, Mar. 1992.

[50] A. J. Patti, M. I. Sezan, and A. M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transaction on Image Processing*, vol. 6, pp. 1064-1076, Aug. 1997.

[51] P. E. Eren, M. I. Sezan, and A. M. Tekalp, "Robust, object-based high-resolution image reconstruction from low-resolution video," *IEEE Transaction on Image Processing*, vol. 6, pp. 1446-1451, Oct. 1997.

[52] A. J. Patti and Y. Altunbasak, "Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants," *IEEE Transaction on Image Processing*, vol. 10, pp. 179-186, Jan. 2001.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[53] J. Yu, C. Xiao, and K. Su, "A method of Gibbs artifact reduction for POCS super-resolution image reconstruction," in *proceeding of the International Conference on Signal Processing*, Minneapolis, April 2007.

[54] I. Begin and F. R. Ferrie, "Blind super-resolution using a learning-based approach," in *proceeding of the International Conference on Pattern Recognition*, Cambridge, United Kingdom, Aug. 2004.

[55] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Transactions on Image Processing*, vol. 10, pp. 1299-1308, Sep. 2001.

[56] H. He and L. P. Kondi, "A regularization framework for joint blur estimation and super-resolution of video sequences," in *proceeding of the IEEE International Conference on Image Processing*, Genova, Italy, Sept. 2005.

[57] M. K. Ng, J. Koo, and N. K. Bose, "Constrained total least-squares computations for high-resolution image reconstruction with multisensors," *Journal of Imaging Science and Technology*, vol. 12, pp. 35-42, Jan. 2002.

[58] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Institute of Physics Inverse Problems*, vol. 19, pp. R1-R26, 2003.

[59] N. A. Woods, N. P. Galatsanos, and A. K. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Transactions on Image Processing*, vol. 15, pp. 201-213, Jan. 2006.

[60] J. Chung, E. Haber, and J. Nagy, "Numerical methods for coupled superresolution," *Inverse Problems*, vol. 22, pp. 1261-1272, 2006.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[61] D. Robinson, S. Farsiu, and P. Milanfar, "Optimal registration of aliased images using variable projection with applications to super-resolution," *The Computer Journal*, doi:10.1093/comjnl/bxm007, April 2007.

[62] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli, "Super-Resolution from unregistered and totally aliased signals using subspace methods," *IEEE Transactions on Signal Processing*, vol. 55, pp. 3687-3703, July 2006.

[63] P. Vandewalle, L. Sbaiz, M. Vetterli, and S. Susstrunk, "Super-Resolution from highly undersampled images," in *proceeding of the International Conference on Image Processing*, Genova, Italy, Sept. 2005.

[64] E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Transactions on Image Processing*, vol. 12, pp. 826-837, July 2003.

[65] H. He and L. P. Kondi, "An image super-resolution algorithm for different error levels per frame," *IEEE Transactions on Image Processing*, vol. 15, pp. 592-603, March 2006.

[66] N. Nguyen and P. Milanfar, "A wavelet-based interpolation-restoration method for superresolution," *Circuits, Systems, and Signal Processing, Special issue on advanced signal and image reconstruction*, vol. 19, pp. 321-338, Aug. 2000.

[67] N. K. Bose and K. J. Boo, "High-resolution image reconstruction with multisensors," *International Journal of Imaging Systems Technology*, vol. 9, pp. 294-304, Dec. 1998.

[68] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen, "Wavelet deblurring algorithms for spatially varying blur from high-resolution image reconstruction," *Linear algebra and its applications*, pp. 139-155, 2003.

[69] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen, "Wavelet algorithms for high-resolution image reconstruction," *SIAM Journal on Scientific Computing*, vol. 21, pp. 1408-1432, 2003.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[70] L. Shen and Q. Sun, "Biorthogonal wavelet system for high-resolution image reconstruction," *IEEE Transactions on Signal Processing*, vol. 52, pp. 1997-2011, July 2004.

[71] N. K. Bose, S. Lertrattanapanich, and M. B. Chappalli, "Superresolution with second generation wavelets," *Signal Processing: Image Communication*, vol. 19, pp. 387-391, Aug. 2004.

[72] N. K. Bose and M. B. Chappalli, "A second generation wavelet framework for superresolution with noise filtering," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 84-89, Aug. 2004.

[73] M. B. Chappalli and N. K. Bose, "Simultaneous noise filtering and superresolution with second-generation wavelets," *IEEE Signal Processing Letters*, vol. 12, pp. 772-775, Nov. 2005.

[74] J. W. Hwang and H. S. Lee, "Adaptive image interpolation based on local gradient features," *IEEE Signal Processing Letters*, vol. 11, pp. 359-362, Mar. 2004.

[75] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Transaction on Image Processing*, vol. 10, pp. 1521-1527, Oct. 2001.

[76] D. D. Muresan and T. W. Parks, "Adaptive quadratic image interpolation," *IEEE Transaction on Image Processing*, vol. 13, pp. 690-698, May 2004.

[77] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level version," *International Journal of Computer Version*, vol. 40, pp. 25-47, 2000.

[78] W. T. Freeman and E. C. Pasztor, "Markov networks for superresolution," in *proceeding of the Annual Conference of Information Sciences and System*, Mar. 2000.

[79] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image Analogies," in *proceeding of the SIGGRAPH, ACM Press*, Los Angeles, Aug. 2001.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[80] J. Sun, N. N. Zheng, H. Tao, and H. Y. Shum, "Image hallucination with primal sketch priors," in *proceeding of the IEEE International Conference of Computer Vision and Pattern Recognition*, Madison, Wisconsin, June 2003.

[81] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323-2326, Dec. 2000.

[82] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, June 2004.

[83] T. L. Chang, T. L. Liu, and J. H. Chuang, "Direct energy minimization for super-resolution on nonlinear manifolds," in *proceeding of the European Conference on Computer Vision*, Graz, Austria, July 2006.

[84] M. F. Tappen, B. C. Russell, and W. T. Freeman, "Exploiting the sparse derivative prior for super-resolution and image demosaicing," in *proceeding of the International Workshop on Statistical and Computational Theories of Vision*, Nice, France, Oct. 2003.

[85] S. Roth and M. J. Black, "Fields of experts: a framework for learning image priors," in *proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, CA, June 2005.

[86] G. E. Hinton, "Products of experts," in *proceeding of the International Conference on Artificial Neural Networks*, Edinburgh, UK, Sep. 1999.

[87] E. Haber and L. Tenorio, "Learning regularization functionals," *Inverse Problems*, vol. 19, pp. 611-626, 2003.

[88] K. I. Kim, M. Franz, and B. Scholkopf, "Kernel hebbian algorithm for single-frame super-resolution," in *proceeding of the European Conference on Computer Vision*, Prague, May 2004.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[89] C. V. Jiji, M. V. Joshi, and S. Chauduri, "Single-frame image super-resolution using learned wavelet coefficients," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 105-112, Sep. 2004.

[90] C. V. Jiji, S. Chuaudhuri, and P. Chatterjee, "Single frame image super-resolution: should we process locally or globally," *Multidimensional Systems and Signal Processing*, vol. 18, pp. 123-152, Sep. 2007.

[91] H. Y. Kim, "Binary operator design by k-nearest neighbor learning with application to image resolution increasing," *International Journal of Imaging Systems and Technology*, vol. 11, pp. 331-339, May 2000.

[92] G. Dalley, B. Freeman, and J. Marks, "Single-frame text super-resolution: a Bayesian approach," in *proceeding of the International Conference on Image Processing*, Singapore, Oct. 2004.

[93] A. Corduneanu and C. J. Platt, "Learning spatially-variable filters for super-resolution of text," in *proceeding of the IEEE International Conference on Image Processing*, Genova, Italy, Sept. 2005.

[94] L. C. Pickup, S. J. Roberts, and A. Zisserman, "A sampled texture prior for image super-resolution," in *proceeding of the Advances in Neural Information Processing Systems*, USA, Dec. 2003.

[95] D. Datsenko and M. Elad, "Example-based single document image super-resolution: a global MAP approach with outlier rejection," *Multidimensional Systems and Signal Processing*, vol. 18, pp. 103-121, Sep. 2007.

[96] S. Baker and T. Kanade, "Hallucinating faces," in *proceeding of the International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, Mar. 2000.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[97] C. Liu, H. Y. Shum, and C. S. Zhang, "A two-step approach to hallucinating faces: global parametric model and local nonparametric model," in *proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Hawaii, Dec. 2001.

[98] X. G. Wang and X. O. Tang, "Hallucinating face by eigentransformation," *IEEE Transactions on System, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 35, pp. 425-434, Aug. 2005.

[99] C. V. Jiji and S. Chaudhuri, "PCA-based generalized interpolation for image super-resolution," in *proceeding of the Indian Conference on Vision, Graphics and Image Processing*, Indian, Dec. 2004.

[100] J. S. Park and S. W. Lee, "Enhancing low-resolution facial images using error back-projection for human identification at a distance," in *proceeding of the International Conference on Pattern Recognition*, Cambridge, UK, Aug. 2004.

[101] A. Chakrabati, A. N. Rajagopalan, and R. Chellappa, "Super-resolution of face images using kernel PCA-based prior," *IEEE Transaction on Multimedia*, vol. 9, pp. 888-892, June 2007.

[102] T. A. Stephenson and T. Chen, "Adaptive Markov random fields for example-based super-resolution of faces," *EURASIP Journal on Applied Signal Processing*, 2006.

[103] W. Liu, D. Lin, and X. Tang, "Face hallucination through dual associative learning," in *proceeding of the IEEE International Conference on Image Processing*, Genova, Italy, Sep. 2005.

[104] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine*, vol. 13, pp. 43-64, May 1996.

[105] D. Kundur and D. Hatzinakos, "Blind image deconvolution revisited," *IEEE Signal Processing Magazine*, vol. 13, pp. 61-63, Nov. 1996.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[106] A. M. Tekalp, H. Kaufman, and J. W. Woods, "Identification of image and blur parameters for the restoration of noncausal blurs," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, pp. 963-972, Aug. 1986.

[107] R. L. Lagendijk, A. M. Tekalp, and J. Biemond, "Maximum likelihood image and blur identification: a unifying approach," *Optical Engineering*, vol. 29, pp. 422-435, May 1990.

[108] R. L. Lagendijk, J. Biemond, and D. E. Boekee, "Identification and restoration of noisy blurred images using the expectation-maximization algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1180-1191, July 1990.

[109] S. J. Reeves and R. M. Merserear, "Blur identification by the method of generalized cross-validation," *IEEE Transaction on Image Processing*, vol. 1, pp. 301-311, July 1992.

[110] G. R. Ayers and J. C. Dainty, "Iterative blind image deconvolution method and its applications," *Optics Letters*, vol. 13, pp. 547-549, July 1988.

[111] D. Kundur and D. Hatzinakos, "A novel blind deconvolution scheme for image restoration using recursive filtering," *IEEE Transactions on Signal Processing*, vol. 46, pp. 375-390, Feb. 1998.

[112] Q.-C. Tao, D.-S. Luo, and X.-H. He, "An improved NAS-RIF algorithm based on the wavelet denoising and image segmentation," in *proceeding of the International Conference on Machine Learning and Cybernetics*, Shanghai, China, Nov. 2003.

[113] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1167-1183, Sep. 2002.

[114] R. Nakagaki and A. K. Katsaggelos, "A VQ-based blind image restoration algorithm," *IEEE Transaction on Image Processing*, vol. 12, pp. 1044-1053, Sep. 2003.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[115] N. P. Galatsanos, V. N. Mesarovic, R. M. Molina, and A. K. Katsaggelos, "Hierarchical Bayesian image restoration from partially-known blurs," *IEEE Transaction on Image Processing*, vol. 9, pp. 1784-1797, Oct. 2000.

[116] V. N. Mesarovic, N. P. Galatsanos, and M. N. Wernick, "Iterative LMMSE restoration of partially-known blurs," *Journal of the Optical Society of America A*, vol. 17, pp. 711-723, Apr. 2000.

[117] A. C. Likas and N. P. Galatsanos, "A variational approach for Bayesian blind image deconvolution," *IEEE Transactions on Signal Processing*, vol. 52, pp. 2222-2233, Aug. 2004.

[118] T. Schulz, "Multiframe blind deconvolution of astronomical images," *Journal of the Optical Society of America A*, vol. 10, pp. 1064-1073, May 1993.

[119] H. Liu and G. Xu, "Closed form blind symbol estimation in digital communications," *IEEE Transaction on Image Processing*, vol. 43, pp. 2714-2723, Jan. 1995.

[120] H. Pai and A. C. Bovik, "Exact multichannel blind image restoration," *IEEE Signal Processing Letters*, vol. 4, pp. 217-220, Aug. 1997.

[121] S. Pillar and B. Liang, "Blind image deconvolution using a robust GCD approach," *IEEE Transaction on Image Processing*, vol. 8, pp. 295-301, Feb. 1999.

[122] G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms," *IEEE Transaction on Image Processing*, vol. 8, pp. 202-219, Feb. 1999.

[123] G. Giannakis and R. Heath, "Blind identification of multichannel FIR blurs and perfect image restoration," *IEEE Transaction on Image Processing*, vol. 9, pp. 1877-1896, Nov. 2000.

[124] F. Sroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Transactions on Image Processing*, vol. 12, pp. 1094-1106, Sept. 2003.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[125] G. Panci, P. Carnpisi, S. Colonnese, and G. Scarano, "Multichannel blind image deconvolution using the Bussgang algorithm: spatial and multiresolution approaches," *IEEE Transaction on Image Processing*, vol. 12, pp. 1324-1337, Nov. 2003.

[126] B. R. Hunt and O. Kubler, "Karhunen-Loeve multispectral image restoration, part I: Theory," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, pp. 592-600, Jun. 1984.

[127] N. P. Galatsanos and R. T. Chin, "Digital restoration of multichannel images," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 415-421, Mar. 1989.

[128] N. P. Galatsanos, A. K. Katsaggelos, and R. T. Chin, "Least squares restoration of multichannel images," *IEEE Transactions on Signal Processing*, vol. 39, pp. 2222-2236, Oct. 1991.

[129] N. P. Galatsanos and R. T. Chin, "Restoration of color images by multichannel Kalman filtering," *IEEE Transactions on Signal Processing*, vol. 39, pp. 2237-2252, Oct. 1991.

[130] W. Zhu, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized multichannel restoration using cross-validation," *Graphical Models and Image Processing*, vol. 57, pp. 38-54, Jan. 1995.

[131] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Transaction on Image Processing*, vol. 1, pp. 2237-2252, Jul. 1992.

[132] A. K. Katsaggelos, K. T. Lay, and N. P. Galatsanos, "A general framework for frequency domain multichannel signal processing," *IEEE Transaction on Image Processing*, vol. 2, pp. 417-420, July 1993.

[133] R. R. Schultz and R. L. Stevenson, "Stochastic modeling and estimation of multispectral image data," *IEEE Transaction on Image Processing*, vol. 4, pp. 1109-1119, Aug. 1995.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[134] G. Angelopoulos and I. Pitas, "Multichannel Wiener filters in color image restoration," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 4, pp. 83-87, Feb. 1994.

[135] T. W. S. Chow, X.-D. Li, and K. T. Ng, "Double regularization approach for blind restoration of multichannel imagery," *IEEE Transaction on Circuit and Systems, Part I: Fundamental Theory and Applications*, vol. 48, pp. 1075-1085, Sep. 2001.

[136] T. F. Chan and C. K. Wong, "Total variation blind deconvolution," *IEEE Transactions on Image Processing*, vol. 7, pp. 370-375, Mar. 1998.

[137] H. C. Andrews and B. R. Hunt, *Digital image restoration*. New Jersey: Prentice-Hall, 1977.

[138] D. Rajan and S. Chaudhuri, "An MRF-based approach to generation of super-resolution images from blurred observations," *Journal of Mathematical Imaging and Vision*, vol. 16, pp. 5-15, Jan. 2002.

[139] C. Vogel and M. Oman, "Iterative methods for total variation denoising," *SIAM Journal on Scientific Computing*, vol. 17, pp. 227-238, Jan. 1996.

[140] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. New Jersey: Prentice-Hall, 1976.

[141] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, pp. 188-200, Mar. 2002.

[142] Y. Wang, J. Ostermann, and Y.-Q. Zhang, *Video processing and communications*. New Jersey: Prentice Hall, 2002.

[143] F. Sroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Transaction on Image Processing*, vol. 14, No. 7, pp. 874-883, July 2005.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[144] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouky, B. M. Salam, and F. E. A. El-Samle, "Regularized super-resolution reconstruction of images using wavelet fusion," *Optical Engineering*, vol. 44, pp. 1-10, Sep. 2005.

[145] G. Jacquemod, C. Odet, and R. Goutte, "Image resolution enhancement using subpixel camera displacement," *Signal process*, vol. 24, pp. 139-146, Jan. 1992.

[146] Y. He, K.-H. Yap, L. Chen, and L.-P. Chau, "Blind super-resolution image reconstruction using a maximum a posteriori estimation," in *proceeding of the IEEE International Conference on Image Processing*, Atlanta, GA, Oct. 2006.

[147] H. Fu and J. Barlow, "A regularized structured total least squares algorithm for high-resolution image reconstruction," *Linear algebra and its applications*, vol. 391, pp. 75-98, Nov. 2004.

[148] J. B. Rosen, H. Park, and J. Glick, "Structured total least norm for nonlinear problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, pp. 14-30, Jan. 1999.

[149] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Transaction on Image Processing*, vol. 5, pp. 1266-1271, Aug. 1996.

[150] Y. Keller, A. Averbuch, and M. Israeli, "Pseudopolar-based estimation of large translations, rotations, and scalings in images," *IEEE Transaction on Image Processing*, vol. 14, pp. 12-22, Jan. 2005.

[151] M. Elad and D. Datsenko, "Example-based regularization deployed to super-resolution reconstruction of a single image," *The Computer Journal*, vol. 50, pp. 1-16, April 2007.

[152] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, pp. 71-86, 1991.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[153] L. K. Saul and S. T. Roweis, "Think globally, fit locally: unsupervised learning of low dimensional manifolds," *Journal of Machine Learning Research*, vol. 4, pp. 119-155, June 2003.

[154] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *proceeding of the IEEE International Conference on Computer Vision*, Bombay, India, Jan. 1998.

[155] P. Philips, H. Moon, P. Rauss, and S. Rizvi, "The FERET evaluation methodology for face-recognition algorithms," in *proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, June 1997.

[156] L. K. Saul and S. T. Roweis, "An introduction to locally linear embedding," *website: "http://www.cs.toronto.edu/~roweis/lle"*, Jan. 2001.

[157] Y.-L. You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," *IEEE Transaction on Image Processing*, vol. 5, pp. 416-428, Mar. 1996.

[158] P. Blomgren and T. F. Chan, "Color TV: total variation methods for restoration of vector valued images," *IEEE Transaction on Image Processing*, vol. 7, pp. 304-309, Mar. 1998.

[159] B. Gunturk, Y. Altunbasak, and R. M. Merserea, "Color plane interpolation using alternating projections," *IEEE Transaction on Image Processing*, vol. 11, pp. 997-1013, Sep. 2002.

[160] N. X. Lian, V. Zagorodnov, and Y.-P. Tan, "Edge-preserving image denoising via optimal color space projection," *IEEE Transaction on Image Processing*, vol. 15, pp. 2575-2587, Sept. 2006.

[161] J. Bescos and J. H. Altamirano, "Digital restoration models for color imaging," *Applied Optics*, vol. 27, pp. 419-425, Jan. 1988.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[162] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Transaction on Image Processing*, vol. 7, pp. 813 - 824, Jun. 1998.

[163] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Journal of Physics D: Applied Physics*, vol. 10, pp. 259-268, Nov. 1992.

[164] L. Chen and K.-H. Yap, "A soft double regularization approach to parametric blind image deconvolution," *IEEE Transaction on Image Processing*, vol. 14, pp. 624-633, May 2005.

[165] T. F. Chan and C.-K. Wong, "Convergence of the alternating minimization algorithm for blind deconvolution," *Linear Algebra Application*, vol. 316, pp. 259-285, Sept. 2000.

[166] Z. Lin and Y.-Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 83-98, Jan. 2004.

[167] W. Lu and Y. P. Tan, "Color filter array demosaicking: new method and performance measures," *IEEE Transaction on Image Processing*, vol. 12, pp. 1194-1210, Oct. 2003.

[168] T. Gotoh and M. Okutomi, "Direct super-resolution and registration using raw CFA images," in *proceeding of the International Conference on Computer Vision and Pattern Recognition*, Washington, DC, Jul. 2004.

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

[170] M. Elad and A. Feuer, "Super-resolution reconstruction of image sequences," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 817-834, Sep. 1999.

[171] M. Elad and A. Feuer, "Superresolution restoration of an image sequence: adaptive filtering," *IEEE Transaction on Image Processing*, vol. 8, pp. 387-395, Mar. 1999.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[172] J. Tian and K.-K. Ma, "A new state-space approach for super-resolution image sequence reconstruction," in *proceeding of the IEEE International Conference on Image Processing*, Genoa, Italy, Sep. 2005.