Super Resolution Image Reconstruction based on Efficient Priors

Pulak Purkait

Electronics and Communication Sciences Unit

Indian Statistical Institute

A thesis submitted to Indian Statistical Institute for partial fulfillment of the requirements for the degree of

Doctor of Philosophy

2013

I would like to dedicate this thesis to my parents who have given lots of love and support to the continuance of my entire education.

Acknowledgements

First and foremost I offer my sincerest gratitude to my supervisor, Prof. Bhabatosh Chanda, who has supported me throughout my thesis with his patience and knowledge whilst allowing me the room to work in my own way. I attribute this degree to his encouragement and effort and without him this thesis would not have been completed or written. One simply could not wish for a better or friendlier supervisor.

I thank all Professors of ECSU for their kind cooperation. I wish to express my gratitude to Prof. Nikhil Ranjan Pal and Prof. Dipti Prasad Mukherjee for their invaluable advises and continuous encouragement. I am grateful to Dr. Pinakpani Pal for being so friendly, cheerful and courageous. I would like to thank Prof. Aditya bagchi and Prof. Srimanta Pal for their valuable suggestions and Dr. Swagatam Das for continuous encouragement to be an achiever. I appreciate our lab assistant Dilip Kumar Gayen for helping on any software and hardware related issues.

In my daily work I have been blessed with a cheerful group of fellow students and researchers of Electronics and Communication Sciences Unit. Soummitra Samanta were like our team manager. He took all the responsibilities on every lab event and during all of our tours. Mrinmoy Ghorai, Rajesh Kumar and Partha Pratim mohanta was like my elder brothers, supported me on every step during my work. I would like to thank Sujoy Biswas, Dr. Sitansu Kumar Das, Dr. Snehasish Mukherjee, Dr. Bibhas Chandra Dhara and Dr. Sanjay Kumar Saha for their continuous support and valuable discussion.

It would be incomplete without thanking my close friends during my stay in MTech Hostel (Swarup, Sanjay, Kali, Butu, Aritra, Sandeep and rest of all) and in RS Hostel (Suvendu, kaushik, Tapas, Avijit, Sudip, Kaushik, Sourav, Tridip, Kushal, Sandeep and so many).

The Indian Statistical Institute, Kolkata has funded my studies and the Electronics, and Communication Sciences Unit has provided the infrastructural support needed to produce and complete my thesis.

Finally, I thank my parents for their continuous support throughout my studies and for placing me where I am now.

.....
(Pulak Purkait)

Abstract

Observed images of a scene are usually degraded by blurring due to atmospheric turbulence and inappropriate camera settings. The images are further degraded by the various noises present in the environment and the system. Moreover, sometimes we end up with low-resolution image(s) of the scene due to hardware limitations (e.g., sensor size) of a digital camera. Therefore, it is essential to get a sharp clean high-resolution (HR) image from multiple low-resolution (LR) image frames or sometimes from even a single LR image frame. These is popularly known as super-resolution (SR) and one of the classical ill-posed inverse problem in image processing.

Regularization is an well-known method to convert an ill-posed problem to a well-posed one and obtain a stable solution as discussed in Chapter 1. Researcher has been studying these methods for general inverse problem in last few decades. Total Variation or Bilateral Total variation based regularization methods are already proposed and studied for ill-posed Super Resolution (SR) problems. However, the SR image reconstructed by these methods produce ringing artifacts near strong edges. Second, the extension of SR Imaging to SR video always desire faster SR reconstruction process.

In Chapter 2, we first develop a gain-control based locally adaptive regularization method for multi-frame SR reconstruction for relatively more detail reconstruction while suppressing the ringing artifacts. We present an iterative process for the model and perform a series of numerical experiments to show the evidence of good performance.

Multiscale morphological operators are studied extensively in the literature for image processing and feature extraction purpose. In Chapter

3, based on multiscale morphology, we model a non-linear regularization method for edge-preserving SR image reconstruction. We formulate multi-frame SR image reconstruction problem as a de-blurring problem and then solve the inverse problem using Bregman iterations. The proposed algorithm can efficiently suppress inherent noise of LR image as well as noise generated during SR image estimation.

In Chapter 4, we also establish a key relationship between the regularization methods with a general edge-preserving noise filtering method which leads to an efficient adaptive regularization based on the kernel of geodesic distance. Here we extend the morphologic regularization proposed earlier to an adaptive one. We propose (i) a new robust adaptive geodesic regularization and (ii) adaptive morphologic regularization methods based on geodesic distance, that can suppress the noise more efficiently while preserving the edges. We have experimentally established the efficiency and superiority of the proposed regularization methods for SR image reconstruction.

Because of the problem of having sufficient number of LR frames of the same scene a model-based high quality single-frame SR is attracting more attentions these days. The output image quality due to recent methods are not up-to that mark, and so still these are not commercialized though there is a huge commercial demand. Instead of classical interpolation based image zooming, where resolution is increased by inserting pixels, in Chapter 5, we have proposed a patch based image zooming technique, where each patch is replaced by estimated larger patch with sharper details. Our approach is based on the fundamental idea that a LR image patch could be generated from any of the many possible HR image patches. Therefore it would be more natural to use multiple predictors to generate a suitable HR patch from a given LR patch instead of single one. In this thesis, we build a generic framework to estimate HR image from a LR one using an adaptive prior (select the predictor locally) based on the local statistics of LR patches. We use natural image patch prior as the

HR patch statistics. We partition the natural images into documents and group them to discover the inherent topics using probabilistic Latent Semantic Analysis (pLSA) and also learn the dual dictionaries of HR and LR image patch pairs for each of these topics using sparse dictionary learning technique. Then for a test image we infer locally which topic each LR patch corresponds to and then we use the corresponding dual dictionary to generate HR patch. At the end we aggregate all those patches to get predicted HR image. Though produces very good results, this method is computationally expensive, due to determination of topic and sparse representation.

To develop a faster image zooming algorithm, in Chapter 6, we propose a novel fuzzy rule based prediction framework. Here also we assume that a LR image patch may be generated from any of the many possible HR image patches. Therefore it would be natural to develop different rules to each of the possibilities of HR patches from a given LR patch and then combine them according to rule strength to get the desired HR patch. The fuzzy rules and the rule parameters are learned from the database of LR-HR patch pairs. This results in an efficient mapping from LR patch space to HR patch space.

The performance of the proposed single frame SR methods are tested qualitatively on different test images, and also compared with representative conventional as well as the relevant state-of-the-art image zooming techniques. Experimental results show that the proposed methods yield superior performance over the competing state-of-the-art methods and are capable of efficiently reconstruction of thin lines, edges, fine details and texture in the image.

This is an attempt to explore the powerful mathematical morphology tool in the field of ill-posed inverse reconstruction problems. The morphologic regularization method proposed here are tested only on multi-frame SR reconstruction problem, but one may extend this regularization to other ill-posed inverse problems of image processing as well (e.g., deblurring). The proposed single-frame SR methods can

also be straight away extended to Video SR technique, where either we can apply SR technique on individual frame to get SR video, or we can divide the video into overlapping or non-overlapping 3D blocks (instead of 2D patches for a single-frame) and use a similar algorithm to increase the spatial as well as temporal resolution. One may also be interested in applying the proposed methods to the other image reconstruction applications, viz. image mosaicing. Moreover, as far as our knowledge goes, the attempt of generation of HR images using fuzzy rules is first application of fuzzy rules in high dimension data. Certainly, there is a huge space to explore more of image processing applications using the proposed fuzzy rule-based approach.

Contents

List of Figures

List of Tables

Nomenclature

Roman Symbols

K	Number of observed low-resolution images
l_x	Sampling period of the LR image along vertical direction
l_y	Sampling period of the LR image along horizontal direction
L_x	Resolution factor along vertical direction
L_y	Resolution factor along horizontal direction
R	Index matrix corresponding to sub-pixel shifts

Greek Symbols

$\oplus$	Morphological dilation
$\ominus$	Morphological erosion
$\bullet$	Morphological closing
$\circ$	Morphological opening
Γ	High-pass operator
α	Parameter of the smoothing kernel
Υ	Regularization operator
λ	Regularizing parameter

α_c Fuzzy-rule strength for c^{th} cluster

μ_c Normalized fuzzy-rule strength for c^{th} cluster

τ_c The mixture weight for c^{th} cluster

Superscripts

n Estimated after n^{th} iteration

T Matrix transpose

Subscripts

k Low-resolution image index

s Index representing size of disk

Other Symbols

A_K Incorporates effect of down-sampling, blurring and warping matrix

B^p_Υ Bregman distance corresponding to convex functional $\Upsilon(.)$

D Down-sampling matrix

$D_{ij,\,lm}$ Distance of the pixel $(i+l, j+m)$ from the center pixel (i, j)

D^h_k Dictionary portion corresponding to HR component

D^l_k Dictionary portion corresponding to LR component

$\downarrow_s$ Down-sampling with resolution factor s

$\uparrow_s$ Up-sampling with resolution factor s

e_k Additive noise in the k^{th} observed image

F_k Geometric warp matrix

Geo_{lm} Geodesic distance from current $(i, j)^{th}$ pixel to $(i+l, j+m)^{th}$ pixels

A Membership matrix of fuzzy-rule base linear model

C	Number of Cluster and hence fuzzy-rules
D	Consequent parameters of fuzzy-rule base system
I_g	Gain mapping
H	Blurring matrix
C_a	Adaptive morphological closing
C_s	Morphological closing with disk size structuring element of radius s
D_s	Morphological dilation with disk size structuring element of radius s
E_s	Morphological erosion with disk size structuring element of radius s
O_a	Adaptive morphological opening
O_s	Morphological opening with disk size structuring element of radius s
$P(z_k)$	The topic prior probability
R_c^{TS}	The c^{th} rule of the fuzzy-rule base system
Σ_c	The variance covariance matrix of c^{th} cluster of fuzzy-rule base
S_x^l	Shift operator matrix to present l pixel shift in horizontal direction
S_y^m	Shift operator matrix to present m pixel shift in vertical direction
d_j	j^{th} document in the corpus
$n(w_i, d_j)$	Stores the frequency of occurrences of the word w_i in the document d_j
w_i	i^{th} word of the Dictionary
z_k	Hidden variable represent k^{th} topic in the corpus
$\mathbf{v}_c$	The c^{th} cluster center of fuzzy-rule base
W	Denotes an index set incorporates an window around a pixel
X	Estimated high-resolution image

$\mathbf{X}^{(0)}$ Initial Approximation of HR image

Y_k k^{th} low resolution image

$\underline{\mathbf{Y}}$ Upsampled and Ensembled all LR images

Acronyms

ART Algebraic Reconstruction Technique

Breg Optimized with Bregman Iterations

BTV Bilateral total variation

BV Bounded variation

CCD Charge coupled device

CFT Continuous Fourier transform

CMOS Complementary metal-oxide-semiconductor

CS Compressed Sensing

CWM Center-Weighted Median Filters

DCT Discrete Cosine transform

DFT Discrete Fourier transform

EM Expectation Maximization

FT Fourier transform

FMM Fast Marching Method

FPC Fixed Point Continuation

GCBTV Gain controlled bilateral total variation

GMM Gaussian Mixture Models

Grd Optimized with gradient descent

HDTV High Defination Television

HR High Resolution

KL Kullback-Leibler divergence

LABTV Locally adaptive bilateral total variation

LDA Latent Dirichlet Allocation

LR Low Resolution

LR Low Resolution

LSE Least square estimation

MAP Maximum a Posteriori

SE Structuring Element

MATLAB Matrix Laboratory Software

Morph Morphologic Regularization

NFS Neuro-fuzzy Systems

NLM Non-local Means

OS Operating System

PDSV Phase-driven spatially variant regularization

pLSA Probabilistic Latent Semantic Analysis

POCS Projections on a convex set

PSNR Peak Signal-to-Noise Ratio

RMSE Root Mean Square Error

SKR Steering Kernel Regularization

SPAMS SPArse Modeling Software

SPM Spatial Pyramid Matching

SR Super Resolution

SSIM The Structural SIMilarity

ST Structure Tensor

SVD Singular Value Decomposition

TSM Takagi-Sugeno model

TV Total variation

Chapter 1

Introduction

1.1 Why Super Resolution?

In almost every application of digital image processing, it is desirable to generate an image that has a very high-resolution. This includes remote sensing, medical imaging, petroleum exploration, military information gathering and high definition television (HDTV). A high resolution (HR) image can contribute to a better appearance or classification of regions. The resolution of an image solely depends on the resolution of the image acquisition device, i.e., camera given its location of camera in the scene. The term 'resolution' refers to area of the scene represented by a single pixel and the higher resolution refers to less area resulting in more spatial detail in the image. So it is always desirable to have HR image for most of the computer vision applications and also for better viewing. A modern image sensor is typically a charge-coupled device (CCD) or a complementary metal-oxide-semiconductor (CMOS) active-pixel sensor. These sensors are typically arranged in a two-dimensional array to capture two-dimensional image signals. The sensor size or equivalently the number of sensor elements per unit area in the first place determines the spatial resolution of the image. A direct method for increasing the pixel resolution may be by sensor manufacturing technique. However, as the resolution of the sensor increases, so does the cost of the sensor. Second, increasing the chip size or packing more sensors in a single chip leads to increase in capacitance, which makes it difficult to speed

up the charge transfer. Moreover, the image details (high-frequency bands) are also limited by the optics, due to lens blurs (associated with the sensor point spread function (PSF)), lens aberration effects and aperture diffractions. Thus constructing imaging chips and optical components to capture very HR images is prohibitively expensive and is not practical in most real applications. So using algorithmic image reconstruction technique to obtain a HR image from observed low resolution (LR) image(s) has become a promising solution to the problem. The technique is popularly known as 'Super Resolution' (SR) technique. The topic remains an area of active research for the last two decades because of its many interesting applications including synthetic zooming of region of interest in forensic, surveillance, remote sensing, and also for conversion from NTSC/PAL video signal to HDTV signal.

SR are the techniques that construct HR image from observed LR image(s), thereby increasing the high-frequency components and suppressing the degradations caused by the imaging process of the LR camera. The basic idea behind SR is to combine the non-redundant information contained in the LR frame(s) with domain specific knowledge to generate a HR image. A closely related technique with SR is the single-image interpolation approach, which can be also used to increase the image size. However, since there is no additional information provided, the quality of the single-image interpolation is very much limited due to the ill-posed nature of the problem.

1.2 Literature survey

There are several popular approaches to SR image reconstruction [69, 15, 12]. A number of fundamental assumptions are made about image formation and quality, which in turn lead to different SR image reconstruction algorithms. These assumptions include the type of geometric transformation relating the LR input images, and the type of noise. There is also choice to be made whether the goal is to produce the very best SR image possible, or an acceptable SR image as quickly as possible. Moreover SR algorithms may vary depending on whether multiple LR images are available or only single LR image is available. Accordingly, SR image reconstruction methods may be broadly categorized into two classes: (i) Multi-

frame SR methods, and (ii) Single-frame SR methods. In the classical multi-frame SR [29, 31, 46], multiple LR images of the same scene with different sub-pixel shifts are taken as input. Since infinitely many HR images are possible that may results in the same set of LR images when down-sampled and degraded, SR image reconstruction is primarily known as an ill-posed problem. A set of constraints is imposed on the unknown HR image along with a regularization/prior term to convert the ill-posed SR problem to a well-posed one.

1.2.1 Multi-frame SR algorithms

Popular Multi-frame SR reconstruction algorithms can be further divided into two categories: Frequency domain algorithms and Spatial domain algorithms.

1.2.1.1 Frequency domain algorithms

Among frequency domain algorithms, different frequency transformations are employed. Tsai and Huang [113] proposed an algorithm for the SR reconstruction by estimating the relative shifts between observations. They related the continuous Fourier transform of the original scene to the discrete Fourier transforms of the observed LR images. Based on this algorithm, a series of improved SR reconstruction algorithms had been proposed [108, 89, 81]. Those methods are based on three fundamental principles: i) the shifting property of the Fourier transform (FT), ii) the aliasing relationship between the continuous Fourier transform (CFT) and the discrete Fourier transform (DFT), iii) the original scene is band-limited. These properties allow the formulation of a system of equations relating the aliased DFT coefficients of the observed images to samples of the CFT of the unknown scene. These equations are solved yielding the frequency domain coefficients of the original scene, which may then be recovered by inverse DFT. Formulation of the system of equations requires knowledge of the translational motion between frames to sub-pixel accuracy. Each observation image must contribute independent equations, which places restrictions on the inter-frame motion that contributes useful data.

The ability to cope with noise in the input images was added in the SR algorithm by Kim *et al.* [51]. A discrete cosine transform (DCT)-based method

was proposed by Rhee and Kang [89]. They reduce memory requirements and computational costs by using DCT instead of DFT. They also apply multichannel adaptive regularization parameters to overcome ill-posedness such as underdetermined cases or insufficient motion information cases. Tekalp *et al.* [108] generalized the technique to cope with both noise and blur in the inputs due to the imaging process. Tom and Katsaggelos [111, 110] took a two-phase SR approach, where the first step was to register, de-blur and de-noise the LR images, and the second step was to interpolate them together onto a HR image grid.

Wavelet models are also applied to the problem, taking a similar overall approach as Fourier-domain methods. Nguyen and Milanfar [76] proposed an efficient algorithm based on representing the LR images using wavelet coefficients and related these coefficients to those of the desired HR image. Bose *et al.* [3] proposed a method based on second generation wavelets, leading to a fast algorithm. They proposed the recursive total least squares method for SR reconstruction to reduce effects of registration errors. However, while this shows good results even in the presence of high level of input noise, the outputs still display wavelet like high frequency artifacts.

Theoretical simplicity is a major advantage of the frequency domain approach. That is, the relationship between LR images and the HR image is clearly demonstrated in the frequency domain. The frequency method is also convenient for parallel implementation capable of reducing hardware complexity. However, the observation model is restricted to only global translational motion and LSI blur. Due to the lack of data correlation in the frequency domain, it is also difficult to apply the spatial domain *apriori* knowledge for regularization.

1.2.1.2 Spatial domain algorithms

Among spatial domain algorithms, representative works include non-uniform interpolation based algorithms [58, 75]. An advantage of these methods is their low computational cost, which makes them suitable for real-time applications. However, degradation models are not applicable in these methods if the blur and the noise characteristics are different for different LR images. The methods based on projections on a convex set (POCS) [103, 83] utilize the spatial domain obser-

vation model and include *apriori* information. Though the methods are simple, their disadvantages are non-uniqueness of solutions, slow convergence rate and high computational cost.

Iterative back projection based methods [46] performed SR reconstruction in a straightforward way. However, these methods do not yield unique solution due to the ill-posed nature of the inverse problem, and some parameters are difficult to choose. On the other hand, Bayesian maximum a posteriori (MAP) estimation based methods [29, 118, 99] explicitly use the *apriori* information in the form of a prior probability density on an HR image and provide a unique solution using a rigorous theoretical framework. Regularization based methods [11, 95, 49], learning based SR methods [37, 98, 10] and space-time SR method [98] usually are better in adaptability and lead to better SR reconstruction than frequency domain approaches. So they have become popular in recent years. In [99], a MAP based joint formulation was proposed that judiciously combined motion estimation, segmentation and SR reconstruction together. This formulation was used to solve a complex SR problem in which the scenes contained multiple independently moving objects. However, this method fails to produce good SR images in real scenario. Super-resolution reconstruction method for color images are also proposed in [33, 91].

1.2.2 Single-frame SR algorithms

SR reconstruction of another class of methods generate HR image from a single LR image or frame. These algorithms again can be roughly categorized into two major approaches: (i) Interpolation-based methods and (ii) Patch-based methods.

(i) Interpolation-based methods: While simple interpolation methods such as Bilinear or Bicubic interpolation tend to generate overly smooth images with ringing and jagged artifacts, the interpolation by exploiting the natural image priors generally produces much better results. Dai *et al.* [20] represented the local image patches using the background/foreground descriptors and reconstructed the sharp discontinuity between the two. Sun *et al.* [105] explored the gradient profile prior for local image structures and applied it to super-resolution. Such approaches are effective in preserving the edges in zoomed image. However,

they are limited in modeling the visual complexity of the real images.

(ii) Patch-based methods: Another category of single-frame SR methods are based on machine learning techniques, which attempt to capture the co-occurrence prior between LR and HR image patches. These methods typically require a large database of HR and LR patch pairs and a learning mechanism to learn the correspondence map. These methods are called "Example-Based SR" [37, 30, 38, 121] or "image hallucination" [50]. Here, correspondence between LR and HR image patches is learned from a collection of LR and HR image pairs and then applied to a new LR image to reconstruct its most likely HR version. These algorithms are, in general, based on image edge prior [30, 20] or gradient profile prior [105]. However, unlike classical SR, the high resolution details reconstructed ("hallucinated") by example-based SR are not guaranteed to provide the true (unknown) high resolution details. Methods based on learning edge models are also proposed [20, 34, 43]. The goal of these methods is to magnify (zoom) an image while maintaining the edge sharpness and the texture details in the image.

Glasner *et al.* [38] merged the concepts of both single-frame SR and multi-frame SR for HR image reconstruction from a single-frame. They down-scaled the input LR image to further coarser scale and then learned the correspondence between the LR and the HR patch pairs. In the recent years SR algorithms with sparse image prior [50, 24] have been receiving more attention due to advancement of Sparse Coding techniques [62, 63, 64]. It has been applied to many other related inverse problems in image processing, such as compression [84], de-noising [28] and restoration [61, 64]. Rigorous study on image statistics suggests that an image patch can be well-represented as a sparse linear combination of elements from an appropriately chosen over-complete dictionary [84, 61]. The effectiveness of such sparsity prior is demonstrated by Yang *et al.* [121] for both general image SR and the special case of face hallucination. The extended version [120] of the work, uses a cheap bicubic interpolation for the smooth patches and sparse coding for the texture, edge and corner patches. The method use neural network based training algorithm for fast prediction of saprse code. The effectiveness of such sparsity prior is demonstrated by Yang *et al.* [121].

1.3 Multi-frame SR problem formulation

Observed images of a scene are usually degraded by unknown blurs due to atmospheric turbulence and inappropriate camera settings. The images are then down-sampled by a factor determined by the intrinsic camera parameters. Those are further degraded due to noise present in the environment. Figure 1.1 shows a typical observation model relating the HR image with LR video frames, as introduced in the literature [69, 33, 12]. The input of the imaging system is continuous natural scenes that may be well approximated as band-limited signals. These signals may be contaminated by atmospheric turbulence before reaching the imaging system. Sampling the continues signal beyond the Nyquist rate generates the high resolution digital image (a) we desire. In case of multi-frame SR setting, usually there exists some kind of motion between the camera and the scene to capture. The inputs to the camera are multiple instances of the scene, connected by possibly local or global shifts, leading to image (b). Going through the camera, these motion related high-resolution frames are subjected to different kinds of blurring effects, such as optical blur and motion blur. These blurred images (c) are then downsampled at the image sensors (e.g., CCD detectors) into pixels, by an integral of the image area falling onto each sensor area. These down-sampled images are further affected by the sensor noise and color filtering noise. Thus, the frames captured by the LR imaging system are blurred, decimated, and noisy versions of the underlying true scene.

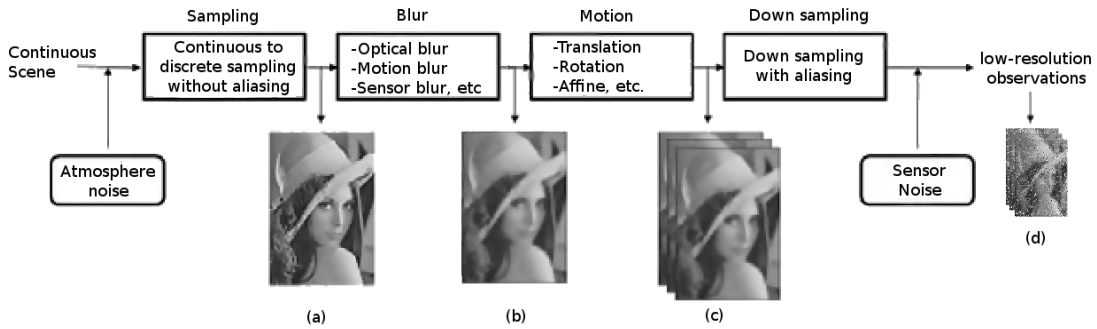


Figure 1.1: The observation model of a real imaging system relating a high resolution image to the LR observation frames with motion between the scene and the camera.

Each low-resolution frame is a decimated, aliased observation of the true scene as a member of the sequence of images taken over time. SR reconstruction is possible only if there exists sub-pixel motions between these LR frames such that each LR frame contains supplementary information about the underlined HR image, and thus the ill-posed upsampling problem can be better conditioned. A pictorial example of the scheme with three sub-pixel shifted images of a scene is shown in Fig. 1.2, where the sub-pixel shifts of each frame $[\delta_x, \delta_y]$ along x-(vertical) and y-(horizontal) direction respectively. In the imaging process, the camera captures several LR frames, which are downsampled from the HR scene with sub-pixel shifts relative to each other. SR construction reverses this process by aligning the LR observations to sub-pixel accuracy and combining them into a HR image grid (interpolation), thereby overcoming the imaging limitation of the camera.

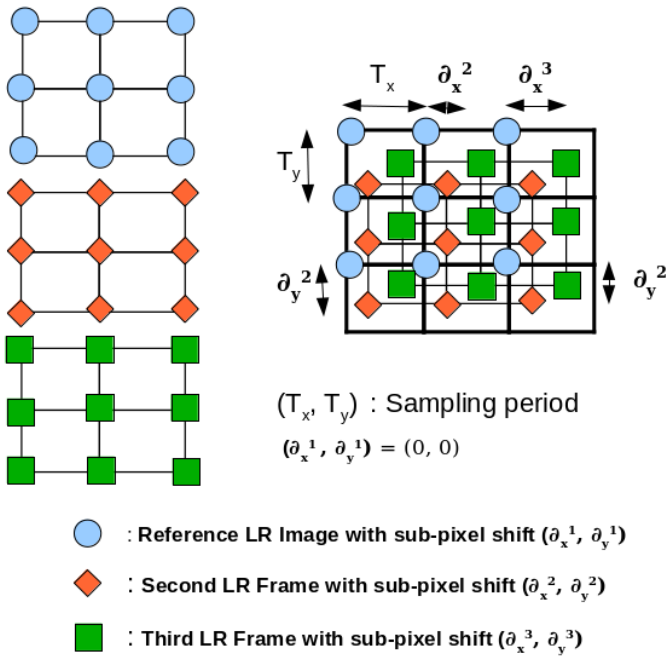


Figure 1.2: Illustrates the registration of three sub-pixel shifted images. The image with pixel co-ordinates marked by '•' is assumed to be reference image.

Let us denote the continuous two-dimensional (2D) image by $f(x, y)$ and the

discrete HR image by f_{HR} of size $l_xL_x \times l_yL_y$. We also denote the k^{th} LR discrete image of size $L_x \times L_y$ by f_k^{LR} having resolution factor l_x and l_y along x and y direction respectively. Suppose f_{HR} is obtained by sampling $f(x, y)$, then they are related as

$$f_{HR}(r, c) = f(rs_x, cs_y) \tag{1.1}$$

where s_x and s_y are sampling periods of HR images in $x-$ and $y-$ directions respectively. Similarly, kth LR image is obtained from $f(x, y)$ as

$$f_k^{LR}(r, c) = f(rl_x^{'} + \delta_x^k, cl_y^{'} + \delta_y^k), \qquad \forall k = 1, 2, ..., K \tag{1.2}$$

where $l_x^{'}$ and $l_y^{'}$, the sampling periods of LR images, are given by $l_x^{'} = s_xl_x$ and $l_y^{'} = s_yl_y$. Here K is the total number of LR images, δ_x^k and δ_y^k represent the shifts in the $x-$ and $y-$ directions of the k-th LR image with respect to the reference HR image satisfying $0 \leq \delta_x^k < l_x^{'}$ and $0 \leq \delta_y^k < l_y^{'}$. Hence instead of HR image, if we consider one of the LR images as reference image, then corresponding shifts satisfying $0 \leq \delta_x^k < l_x^{'}$ and $0 \leq \delta_y^k < l_y^{'}$ are called sub-pixel shifts. The relationship between the LR images and the HR image can be formulated as [33, 51, 103, 110, 111]

$$\mathbf{Y}_k = DH_k^{cam}F_kH_k^{atm}\mathbf{X} + \mathbf{e}_k, \qquad \forall k = 1, 2, ..., K \tag{1.3}$$

where $\mathbf{Y}_k$, $\mathbf{X}$ and $\mathbf{e}_k$ represent lexicographically ordered column vectors of kth LR image, HR image and additive noise respectively. Thus, dimension of $\mathbf{Y}_k$ and $\mathbf{e}_k$ is $M = L_xL_y$ and that of $\mathbf{X}$ is $N = l_xl_yL_xL_y$. H_k^{atm} is atmospheric blurring matrix and H_k^{cam} is camera lens/CCD blurring matrix each of size $N \times N$. F_k is a geometric warp matrix representing relative motion of LR images given by a linear shift. D is the down-sampling matrix of size $M \times N$ and k is the index of the LR images. $\mathbf{e_k}$ is generally assumed to be additive noise. Under some assumptions blur and geometric warp matrices commute [33], then Eq. (1.3) can be rewritten as

$$\mathbf{Y}_k = DF_kH_k^{cam}H_k^{atm}\mathbf{X} + \mathbf{e}_k, \qquad \forall k = 1, 2, ..., K \tag{1.4}$$

In conventional imaging systems (such as video cameras), camera lens/CCD

blur has more significant effect than the atmospheric blur, which is very dominating in case of for astronomical or satellite images. In former cases, H_k^{cam} is the dominant blurring and we can ignore H_k^{atm} for SR modeling. Moreover, assuming that the LR images are taken under same environmental condition and using same sensor, $H_k^{cam}H_k^{atm}$ becomes same for all k and may be denoted simply by H. Then Eq. (1.4) reduces to

$$\mathbf{Y}_k = DF_kH\mathbf{X} + \mathbf{e}_k, \qquad \forall k = 1, 2, ..., K \tag{1.5}$$

Finally, statistical nature of $\mathbf{e_k}$ are identical for all LR images and is usually taken to be normally distributed. We assume that the parameters of the distribution are known. As a result, in discrete domain the LR images are related to the HR image as

$$\mathbf{Y}_k = A_k\mathbf{X} + \mathbf{e}, \qquad \text{for} \quad k = 0, 1, 2, ..., K-1 \tag{1.6}$$

where $A_k = DF_kH$ of size $M \times N$ is the complete transformation matrix combining different kinds of blur, motion and down-sampling. Now A_k varies image to image because of F_k. Our goal is to estimate $\mathbf{X}$ given K LR images $\mathbf{Y}_k$, model of A_k (or, in other words, models of D, H and F_k) by solving the system of linear equations described in Eq. (1.6).

1.3.1 Regularization for iterative algorithm

As described in the previous section, Eq. (1.5), which leads to Eq. (1.6), represents an ill-posed under-determined system of linear equations. Such a system results in infinite number of solutions for $\mathbf{X}$. The intuition behind the regularization is to select a suitable solution that satisfies some image quality criterion. Regularization has already been used in conjunction with iterative techniques for restoration of noisy and degraded images. It basically introduces an additional information in order to solve an ill-posed problem or to prevent an over-fitting. Moreover, regularization helps removing artifacts from the estimated image and also improves the rate of convergence. Different types of regularizing functions can be found in literature. Some limit the total energy, while some other forces spatial smoothness by limiting high frequency energy [33]. The latter one is more

common. Now a stable solution to image reconstruction is expected to maintain spatial homogeneity of gray-level, i.e., a pixel value should be consistent with that of its neighborhood. To obtain such a stable solution, a specific regularization $\Upsilon(X)$ needs to be imposed on the observation model. The regularization $\Upsilon(X)$ can incorporate prior knowledge of the desirable HR solution, e.g., degree of smoothness. This additional criterion that favors well-behaved solutions can be converted to a generalized minimization cost function [29], i.e.,

$$\hat{\mathbf{X}} = Arg \min_{X} \left[\sum_{k=1}^{K} \rho(\mathbf{Y}_k, A_k\mathbf{X}) + \lambda\Upsilon(\mathbf{X}) \right] \tag{1.7}$$

where λ is the Lagrangian constant, called the regularization parameter, that controls the emphasis between two terms: the data error term (the first one) and the regularization term (the second one). The regularization process is emphasized (large value of λ) when only a small number of input LR images are available or the fidelity of the observed data is low due to registration error and noise. On the other hand, if large number of LR images are available and the amount of noise is small then a small value of the regularization parameter λ leads to a good solution.

One of the commonly used regularizing function is the Tikhonov cost function [11, 59, 76, 77, 127, 29] for obtaining stable estimate of de-noised image by minimizing

$$\Upsilon(\mathbf{X}) = ||\Gamma\mathbf{X}||_2 \tag{1.8}$$

where Γ is a high-pass operator such as derivative or Laplacian. The motivation behind this regularization method is to limit high-frequency energy or forcing smoothness. As both noisy and edge contain high-frequency energy, so the resultant de-noised image does not contain sharp edges. One of the first so called edge preserving regularization methods for de-noising and de-blurring is the total variation (TV) method [91]. The TV criterion penalizes the L_1 norm of the gradient defined as

$$\Upsilon(\mathbf{X}) = ||\Gamma\mathbf{X}||_1 \tag{1.9}$$

Farsiu *et al.* [33] introduced the bilateral total variation (BTV) model for regu-

larization by combining the total variation and the bilateral filter which is computationally cheap and preserves edges to some extent. The regularizing function for BTV is given as

$$\Upsilon(\mathbf{X}) = \sum_{l=-w}^{w} \sum_{m=-w}^{w} \alpha^{|m|+|l|} |\mathbf{X} - S_x^l S_y^m \mathbf{X}|_1 \tag{1.10}$$

where $l+m \geq 0$, S_x^l, and S_y^m are shift operator matrices to present l and m pixel shifts in horizontal and vertical directions respectively. The term w represents the window size and α $(0 < \alpha < 1)$ is the weighting coefficient. Equation (1.10) basically represents the weighted sum of first order gradients at different scales in both the directions and the contribution of gradient magnitude decreases as the scale of the gradient operator increases. Thus the degree of noise suppression as well as the degree of edge preservation depends on the value of α. Large value of α helps to reduce noises but also blurs edges in estimated image, while the small value preserves the sharpness of edge of the estimated image but leaves behind some noise.

1.4 Single frame priors

Single frame SR image reconstruction is useful when only one LR image is available (i.e. $K = 1$), therefore problem (1.6) can be written as:

$$\mathbf{Y} = DH\mathbf{X} + \mathbf{e} \tag{1.11}$$

where D, H and $\mathbf{e}$ represents down-sampling matrix, blurring operator and noise respectively. In this case no registration algorithm is required, since there is no other frame than the given frame. Hence the problem becomes extremely ill-posed. Instead of developing more complex regularization for iterative algorithms, use some edge preserving interpolation based technique or learning the image prior from the natural image database efficiently produces pleasant output. Bicubic interpolation is frequently used to increase the number of pixels in an image. However, it (or even sinc interpolation) cannot recover original high frequency details of the scene if the scene is not sampled at a rate higher than the Nyquist

frequency. It is a common experience that images re-sized by interpolation appear smooth and have artifacts near the sharp edges. Hence, an approach that recovers the "true" HR image details from its LR version(s) is essential. The most of the recent techniques described in Section 1.2.2, are based on learning the inverse mechanism of LR input image and estimated HR image from an external image database of LR and HR image pairs. The number of learning parameters can be significantly reduced by learning small patch-pairs instead of whole image. The LR image is first divided into small overlapping patches and then predict HR patch corresponding to each LR patch and then combine the output of overlapping HR patches to get the actual HR image.

1.5 Motivation and contributions

The main focus of this thesis is learning priors for SR methods in spatial domain, which are very effective to solve ill-posed multi-frame and single-frame SR reconstruction problem using a regularized iterative algorithm and as well as using external image database respectively.

Since multi-frame SR reconstruction is an ill-posed problem, our aim is to estimate HR image guided by a *priori* knowledge (e.g., degree of smoothness) represented by a regularization term into the process of SR reconstruction to obtain a stable solution. Tikhonov regularization (1.8) based on L_2 norm is one of the most common regularization method for SR reconstruction. It imposes smoothness in the reconstructed image, but at the same time loses some details (e.g., sharp edges) present in the LR images. Another interesting algorithm proposed by Farsiu *et al.* [33] employs bilateral total variation (BTV) (1.10) and L_1 norm both for regularization and data fusion. Their method is more robust and can preserve details (e.g., edges) better than Tikhonov regularization method. However, the approach fails considering the partial smoothness in an image, i.e., the method is not locally adaptive, and blurs the edges to some extent in the process of SR reconstruction. So it cannot balance the suppression of noise against the preservation of image details. In a recent work Li *et al.* [59] have reduced the aforementioned shortcomings to some extent. They have used a locally adaptive bilateral total variation (LABTV) operator for the regularization.

They have also added a new term in regularization for gradient consistency to achieve further improvement. They measured the regularization term with L_p norm ($1 < p < 2$), while data error term and gradient error term are measured with L_1 norm. However, even though all these regularization systems lead to a stable solution, a periodic noise with ringing artifacts are still present in the estimated SR image whenever we emphasize preserving details of the HR image.

Recently, two other regularization terms are proposed for SR image reconstruction, viz. "non-local means regularization" (NLM) [87] and "steering kernel regression" (SKR) [107]. NLM is based on the idea that if a patch occurs inside an image it is more probable that the patch would occur at other locations within the same image. So it smooths the patch according to photometric similarity with the neighboring patches, whereas SKR is based on the structure tensor and smooths the image along the edges during the SR image reconstruction.

However, even though all these regularizations lead to a stable solution, it is always interesting to explore different kind of regularization methods for betterment. So, developing an edge-preserving regularization method that can suppress noise in degraded images and also the ringing artifacts evolved during SR reconstruction of image without sacrificing sharp edge is yet to achieve. In this thesis, we work toward this goal by using different kind of regularization techniques apart from conventional techniques. We call it as morphologic regularization as they are defined based on mathematical morphologic operators. We also study the relationship between different regularization methods and non-linear filters which leads to more efficient morphologic regularization techniques. In this thesis, we address various computational and optimization issues that are associated with the solution of the system of equations for SR reconstruction using non-linear and adaptive morphologic regularizations. We explore Bregman iterations to address this optimization problem efficiently. It is experimentally shown that the proposed method can achieve better result than existing ones in terms of qualitatively and qualitatively. Moreover, it is shown experimentally and numerically that proposed techniques are faster than or comparable to the existing state-of-the-art methods.

In the context of single-frame SR, which discards most of the limitations of multi-frame SR (e.g. availability of multiple frames, global transformation of all

the LR frames corresponding to a fixed reference frame), still is not commercialized. So developing more fast and efficient algorithm has a huge commercial demand and is always attractive within image processing community. In this thesis, we propose a novel single-frame SR technique using probabilistic latent semantic analysis (pLSA). A natural image usually consists of smooth regions, textures, and sharp boundaries. We consider them as a different 'topics'. They act differently for a small region while generating corresponding LR version. Therefore, it would be more natural if we model each of the topics separately for prediction mechanism from LR to HR. From a collection of natural images we learn those topics in an unsupervised way and develop separate predictors for each of the topics. In this thesis, we also develop a fuzzy rule based system for SR. Here underline assumption is that a small LR patch may be generated from one of many possible HR patches. We use distinct fuzzy rules which generate various possible patches and then combine them using a 'T-norm' to get the estimated HR patch corresponding to a LR one. We also show that the proposed method is less computationally expensive than the existing ones.

1.6 Organization of this thesis

This thesis contributes on developing and learning priors in multi-frame and single-frame SR. It organized as follows:

In Chapter 2, we develop locally adaptive regularization with morphologic gain-Map for multi-frame SR. As discussed in Section 1.3.1, most of the existing regularization/smoothness prior for multi-frame image SR minimize first or second order norm of image gradient with fixed regularization parameter λ for all the pixels (1.7). Since the random noise and the edges may both contribute same amount high frequency components. Therefore, sharp edges are blurred during the smoothness process. In this chapter, we develop a locally adaptive regularization by introducing a gain value at each pixel which controls the smoothness process depending upon the intensity profile over the pixels neighborhood. This novel gain-map is computed based on multi-scale morphologic operators, namely dilation and erosion.

In Chapter 3, we develop an efficient smoothness prior based on multiscale

Morphological filters and solve the non-linearity of the regularization function using efficient sub-gradient method. There has been a lot of work devoted for different kinds of regularization techniques as described in Section 1.3.1. However, almost all of them is based on the first or second order image derivative operators which are linear operators. Therefore exploring different kinds of, especially non-linear, smoothness prior that can preserve edge more reliably is always interesting. The previous chapter motivates us to develop an edge preserving regularization technique based on the multi-scale morphologic operators. However, the non-linearity of the problem makes it hard to solve and moreover it is time consuming. Here, we develop a regularization technique based on morphological operators and come up with an efficient subgradient method to solve this difficulty. Bregman iterations are used for fast convergence of the corresponding energy functional.

In Chapter 4, a key relationship between non-linear low pass-filter and the existing regularization techniques is developed. That leads to formulate a novel and efficient regularization technique based on the kernel of geodesic distance. Here we also propose an adaptive morphological regularization, where the structuring element of the morphological operators varies from pixel to pixel, based on geodesic kernel which performs much better than the non-adaptive regularizations.

Motivated by inherent problem of multi-frame SR, in Chapter 5, we build Single-frame SR with multiple Dictionary learning. In the context of single frame SR, patch-based dual dictionary technique [121] and local edge statistics are explored successfully [38]. However, in most of the applications authors develop an one-to-one relationship between LR and HR patches. Whereas, we observe that a LR patch can be generated from many possible HR patches. Hence, to solve the SR problem efficiently a one(LR)-to-many(HR) map is required. In this chapter, we address this problem efficiently using probabilistic latent semantic analysis (pLSA) [44]. In this work multiple sparse dual dictionary are learned from a set of natural LR and HR image pairs based on the multiple possibilities of HR patch corresponding to a single LR patch, are used for patch prediction, and we make pLSA to select one of the dictionary to generate a HR patch corresponding to a LR patch.

In Chapter 6, we develop another high-quality single frame SR technique

using fuzzy rule based prediction framework. Proposed approach is based on the fundamental idea that a LR image patch could be generated from any of the many possible HR image patches as state before. Therefore it would be natural to assign certain fuzziness to each of the possible HR patches. We develop a prediction system that learns different fuzzy rules for various LR-HR patch correspondence and also learns the natural image patch prior from an external database. We do so by collecting a large amount of the LR-HR natural image patch pairs from an existing database, then generate fuzzy rules to learn the antecedent and consequent parameters to get an efficient mapping from LR patch space to HR patch space.

In Chapter 7, we discuss about the key issues addressed in the previous chapters and also suggest some possible future directions of research in this field. We also discuss about the impact of this thesis in the domain SR and image restoration.

Chapter 2

Morphologic Gain Controlled Regularization

In most of the existing regularization or smoothness prior for multi-frame image SR, we minimize first or second order norm of image gradient with a fixed regularization parameter λ as mentioned in Section (1.3.1). Since both the random noise and the edges contribute to high frequency components, sharp edges would be removed while trying to remove random noise during the smoothness process. In this chapter we find out an way to improve the quality of the reconstructed SR image by suppressing the ringing effect. So that algorithm produces results that can suppress aforementioned periodic noise evolved during iterative process without effecting the edges.

We recall the relation between LR images and the corresponding HR image as described in Eq. (1.6)

$$\mathbf{Y}_k = A_k\mathbf{X} + \mathbf{e}, \qquad \text{for} \quad k = 0, 1, 2, ..., K-1 \tag{2.1}$$

where $A_k = DHF_k$ of size $M \times N$ is the complete transformation matrix combining different kinds of blur, motion and down-sampling. Now A_k varies from LR image to LR image because of F_k. We assume that geometric warp matrix F_k represents only translational motion causing sub-pixel shifts. So to compute F_k we determine the sub-pixel shifts δ_x^k and δ_y^k by taking one of the LR images as reference image. For the purpose of image registration we can use one of the

various sophisticated algorithms available in literature [7, 128]. The reconstruction of HR image, given a set of LR images, is a reverse process. Our goal is to estimate $\mathbf{X}$, given K LR images $\mathbf{Y}_k$, and the model of A_k by solving the system of linear equations described in Eq. (2.1).

2.1 L1 error based estimation of SR image

Since the observed LR images are noise-corrupt and more importantly as the number of LR image $K < N/M = L_x L_y$, we estimate a SR image $\hat{\mathbf{X}}$, which when degraded and down-sampled satisfies Eq. (2.1) in, say, least-square sense, i.e.,

$$\sum_{k=1}^{K} \left(||\mathbf{Y}_k - A_k \hat{\mathbf{X}}||^2 - ||\mathbf{e}||^2 \right) = 0 \tag{2.2}$$

Since $||\mathbf{e}||^2 = N(\mu^2 + \sigma^2)$ is a positive constant for sample mean μ and variance σ^2, Eq. (2.2) is equivalent to

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left[\sum_{k=0}^{K-1} \left(||\mathbf{Y}_k - A_k \mathbf{X}||^2 \right) \right] \tag{2.3}$$

or, in terms of its first derivative, we find an estimate $\hat{\mathbf{X}}$ to satisfy

$$\frac{\partial}{\partial \hat{X}} \sum_{k=0}^{K-1} \left(||\mathbf{Y}_k - A_k \hat{\mathbf{X}}||^2 \right) = \mathbf{0} \tag{2.4}$$

where $\mathbf{0}$ represents a null vector. Simplifying Eq. (2.4) we get

$$\sum_{k=0}^{K-1} (A_k)^T \left(A_k \hat{\mathbf{X}} - \mathbf{Y}_k \right) = \mathbf{0} \tag{2.5}$$

Thus we end up with solving a set of KM linear equations, where M is the size of the LR image. The (.) part of Eq. (2.3) measures the dissimilarity between observed LR image and degraded-and-downsampled version of estimated HR image as L_2 norm. Considering any dissimilarity measure, Eq. (2.3) may be re-written

as, in general,

$$\hat{\mathbf{X}} = Arg\min_{\mathbf{X}} \left[\sum_{k=0}^{K-1} \rho\left(\mathbf{Y}_k, A_k\mathbf{X}\right) \right] \tag{2.6}$$

where $\rho(.)$ is a distance function. Usually, ρ is defined by the L_p norm $(1 \leq p \leq 2)$ of the residual, i.e., $\rho(\mathbf{Y}_k, A_k\mathbf{X}) = ||A_k\mathbf{X} - \mathbf{Y}_k||_p$. According to [29], in case of ill-posed problems L_1 norm leads to a more robust result in error estimation than L_2 norm. We adopt this norm in the present work. In that case, Eq. (2.6) is written as

$$\hat{\mathbf{X}} = Arg\min_{\mathbf{X}} \left[\sum_{k=0}^{K-1} |A_k\mathbf{X} - \mathbf{Y}_k| \right] \tag{2.7}$$

and the estimate $\hat{\mathbf{X}}$ may be obtained by equating the derivative of $\sum_{k=1}^{K} |A_k\mathbf{X} - \mathbf{Y}_k|$ to $\mathbf{0}$, which yields (similar to Eq. (2.5))

$$\sum_{k=0}^{K-1} (A_k)^T sign(A_k\mathbf{X} - \mathbf{Y}_k) = \mathbf{0} \tag{2.8}$$

Because of ill-condition nature of the problem, the set of linear equations presented in Eq. (2.8) are unstable. That means a small perturbation in $\mathbf{Y}_k$ due to noise results in a non-ignorable deviation from the solution $\mathbf{X}$. So solving $\mathbf{X}$ directly based on a few observed LR images $\mathbf{Y}_k$ and the degradation matrices A_k is not a pragmatic approach. We try to solve these equations using a standard iterative back-projection algorithm.

2.1.1 Iterative algorithm

As suggested by Eq. (2.8), reconstruction of HR image $\mathbf{X}$ can be achieved by solving K sets of linear equations, where K is the number of LR images. The iterative approach is found useful because of their superior performance in the above context where $\mathbf{X}$ has large dimension. In that direction the algebraic reconstruction technique (ART) [40] and expectation maximization algorithm (EM) [117, 82] are the algorithms widely used in the community due to their simplicity, efficiency and performance. The algorithm we adopt here for the task is based on the method proposed by Kaczmarz [48], which was rediscovered

and successfully used by Gordon *et al.* [40] in computer tomography. ART formulates dissimilarity measure using L_2 norm (similar to Eq. (2.3)) and updates each element of $\mathbf{X}$ iteratively using gradient descend technique. We modify that algorithm, considering L_1 norm as dissimilarity measure (see Eq. (2.7)), where j-th element of $\hat{\mathbf{X}}$ at $(n+1)$-th iteration is defined as

$$X_j^{(n+1)} = X_j^{(n)} - \beta^{(n)} \frac{(A_{k,j})^T}{||A_{k,j}||^2} sign(A_k \mathbf{X}^{(n)} - \mathbf{Y}_k), \quad \text{for} \;\; j = 0, 1, 2, \ldots, N-1 \;\; \text{and} \;\; k = 0, 1, 2, \ldots, K-1 \tag{2.9}$$

where $A_{k,j}$ is the j-th column of A_k. The iterative parameter $\beta^{(n)}$ represents the step size in the n-th iteration and it reduces as iteration proceeds, i.e., $\beta^{(n)}$ becomes smaller as $\hat{\mathbf{X}}$ approaches the solution. Iteration number n is incremented as all LR images are considered once and is related to LR image index k as : $n = \lceil k/K \rceil$. Initial estimate for the HR image, i.e., $\mathbf{X}^{(0)}$ is usually taken as the average of upsampled version of all available LR images. This modified ART is sufficiently fast compared to traditional gradient descent technique as in each iteration the estimated image is modified based on the current LR image only. So this method is more effective when number of available LR images is low.

The modified ART like algorithm (2.9) using above mentioned regularization, can be represented as

$$X_j^{(n+1)} = X_j^{(n)} + \beta^{(n)} \frac{(A_{k,j})^T}{||A_{k,j}||^2} sign(\mathbf{Y}_k - A_k \mathbf{X}^{(n)}) + \lambda \underbrace{\sum_{l=-w}^{w} \sum_{m=-w}^{w}}_{l+m \geq 0} \alpha^{|m|+|l|} [I - S_y^{-m} S_x^{-l}] sign(X^n - S_x^l S_y^m X^n)\} \tag{2.10}$$
$$\text{for} \;\; j = 0, 1, 2, \ldots, N-1 \;\; \text{and} \;\; k = 0, 1, 2, \ldots, K-1.$$

where $A_{k,j}$, k, β, etc. are the same as defined for Eq. (2.9). S_x^{-l} and S_y^{-m} define the transpose of matrices S_x^l and S_y^m respectively and have a shifting effects in the directions opposite to that of S_x^l and S_y^m.

For better control of α, Li *et al.* [59] defined a new regularization term, namely

locally adaptive bilateral total variance (LABTV) regularization, as

$$\Upsilon_{LABTV}(X) = \sum_{l=-w}^{w} \sum_{m=-w}^{w} \frac{1}{p_{\bar{x}}(m,l)} \phi^{|m|+|l|} ||X - S_x^l S_y^m X||_{p_{\bar{x}}(m,l)} \tag{2.11}$$

where the range of the k^{th} entry of $p_{\bar{x}}(m,l)$ is $1 \leq p_{x_k}(m,l) \leq 2$ and its value varies adaptively with pixels. The weighting matrix $\phi(m,l)$ contains adaptive weighting coefficients computed using the fuzzy entropy based on neighborhood homogeneity.

2.2 Proposed gain controlled adaptive edge preserving regularization

In this section we intend to build an edge-preserving technique for SR image reconstruction. Both the noise and edge contribute to high frequency components. Now if both these components are treated equally, as is done in estimating HR image by means of Eq. (2.10), they are usually removed due to regularization method ensuring smoothness. Here our intention is to develop a regularization method which results in HR images with sharp edges and textures while suppressing the noise. Second problem is that even if we use a sophisticated regularization technique (e.g., Eq. (2.10)), a periodic ringing noise appears around the edges in estimated image. We observe that this ringing effect is most distracting in smooth regions because human perception can tolerate at most small scale ringing in highly textured regions. The problem may be surmounted by employing some adaptive technique where relative effect of regularization term and the data error term at each pixel of estimated HR image is controlled depending on its neighborhood information. For example, if the neighborhood of a pixel is smooth then the regularization term (designed from smoothness criterion) is emphasized to force the said pixel to maintain the smoothness in that image region; on the other hand, if the neighborhood contains an edge the data error term is emphasized to preserve signal characteristics. There are a few methods for such adaptive gain control. A representative example from this class of regularization techniques is due to Li *et al.* [59] who have suggested a locally adaptive bilateral

total variation (LABTV) regularization to control the α term of the Eq. (1.10). In the proposed work, assigning weightage between the data error term and the regularization term is obtained using multi-scale morphology and is described in detail in the next section.

Note that the strong ringing is mainly caused by high contrast edges and magnitude of ringing is proportional to the magnitude of image gradient. Based on these observations, we propose an edge-preserving gain-controlled regularization technique for iterative HR reconstruction. We allow more emphasis on improvement in high frequency regions than smoother regions, so that we get rid of periodic ringing artifacts in smooth region. Moreover, in successive iteration the pixel value in high frequency region changes significantly and thus emphasis on those regions results in speed up of the process during iteration.

For modification of each pixel we assign a gain value to each pixel depending on (inversely proportional to) the smoothness of that pixel's neighborhood. It is a multiplier to suppress the propagation of the ringing effect during each iteration without blurring the edges. So we modify the above iterative regularization algorithm by introducing a gain map I_g^n corresponding to each pixel of n^{th} estimated HR image. Incorporating this modification on (2.10), the reconstructed SR image in $(n+1)$-th iteration can be written as:

$$X_j^{(n+1)} = X_j^{(n)} + \beta^{(n)}\{[I(j,j) + I_g^{(n)}(j,j)]\frac{(A_{k,j})^T}{||A_{k,j}||^2} sign(\mathbf{Y}_k - A_k\mathbf{X}^{(n)}) + \lambda[I(j,j) - I_g^{(n)}(j,j)] \underbrace{\sum_{l=-w}^{w}\sum_{m=-w}^{w}}_{l+m\geq 0} \alpha^{|m|+|l|}[I - S_y^{-m}S_x^{-l}]sign(\mathbf{X}^{(n)} - S_x^l S_y^m \mathbf{X}^{(n)})\},$$
$$\text{for} \quad j = 0, 1, 2, \ldots, N-1 \quad \text{and} \quad k = 0, 1, 2, \ldots, K-1. \tag{2.12}$$

where $n = \lceil k/K \rceil$ is the number of periodic iteration considering all the LR images once, I is the identity matrix of size $N \times N$ where N is the dimension of $\mathbf{X}^{(n)}$. Gain map $I_g^{(n)}$ is also a diagonal matrix of same size as that of I and contains smoothness (or in reverse sense edge-ness), information for each pixels of $\mathbf{X}^{(n)}$. It's diagonal term is zero if corresponding pixel of estimated image

$\mathbf{X}^{(n)}$ lies in a smooth region, is one if it lies in high frequency region (like, edge) and is in-between zero and one depending on the nature of neighborhood of the candidate pixel. Thus, in case of pixels in smooth region, both the coefficients of the data error term $(I + I_g^{(n)})$ and the regularization are 1. On the other hand, the former becomes 2 and the latter becomes 0 for strong edge pixel. In essence, in smooth regions both image data and smoothness gets equal emphasis, while in high frequency regions the image data is much more emphasized than maintaining smoothness as a natural choice. Sometimes, the value of $(I + I_g^{(n)})$ may be normalized by introducing a weighting coefficient ω in the gain control coefficient. The Eq. (2.12) then may be written as

$$X_j^{(n+1)} = X_j^{(n)} + \beta^{(n)}\{[\omega I(j,j) + (1-\omega)I_g^{(p)}(j,j)]\frac{(A_{k,j})^T}{||A_{k,j}||^2}sign(\mathbf{Y}_k - A_k\mathbf{X}) + \lambda[I(j,j) - I_g^{(p)}]\underbrace{\sum_{l=-w}^{w}\sum_{m=-w}^{w}}_{l+m\geq 0}\alpha^{|m|+|l|}[I - S_y^{-m}S_x^{-l}]sign(\mathbf{X}^{(n)} - S_x^l S_y^m \mathbf{X}^{(n)})\},$$
$$\text{for} \quad j = 0,1,2,\ldots,N-1 \quad \text{and} \quad k = 0,1,2,\ldots,K-1 \tag{2.13}$$

where ω is a parameter that controls the influence of the gain map, and selection of its value is not very critical and is mostly camouflaged by λ and β. Since gain map $I_g^{(n)}$ is generated from $\mathbf{X}^{(n)}$, the estimated image after n-th iteration, incorporating it in the last two terms of Eqs. (2.12) and (2.13) makes the iterative reconstruction process locally adaptive. The details of generating gain map I_g^n using multi-scale morphology follows.

2.2.1 Construction of gain map

Since we want to suppress the contrast of ringing in smooth regions while avoiding suppression of sharp edges, the gain map should be small (ideally zero) in smooth regions and large (ideally 1, in the [0,1] range) in high frequency regions. It is known that dilated minus eroded image [102] gives the edge magnitude of the original image. It is also known that dilation inflates bright noise, while erosion inflates the dark ones. However, morphological opening and closing removes

bright and dark noise, respectively, without affecting the edge sharpness. Now an image, in general, may contain both detail and noise at different scales. So we adopt multi-scale morphology based approach to construct the gain map I_g. Suppose B represents a disc structuring element of unit size. Then sB defined as

$$sB = \underbrace{((B \oplus B) \oplus B) \ldots) \oplus}_{(s-1) \text{ times}} B$$

represents a disk structuring element of size s and morphological operation with sB incorporates multi-scale character by varying the value of s. For example, opening with sB removes all bright (positive) noise of size less than s, erosion by sB shrinks the object by s pixels, etc. The algorithm for generating gain map I_g using sB for $s = 1, 2, \ldots, S$ may be summarized as follows :

Step 1: Initialize image $F = \mathbf{0}$.

Step 2: Iterate the following steps for $s = 1$ to S

- Obtain morphological opening $O_s(X^{(n)})$ and closing $C_s(X^{(n)})$ from $X^{(n)}$, estimated image after n-th iteration, i.e,

$$\begin{aligned} O_s(X^{(n)}) &= X^{(n)} \circ sB \\ C_s(X^{(n)}) &= X^{(n)} \bullet sB \end{aligned}$$

  where B is a disk structuring element of unit size and s is scale parameter.

- Apply gray-scale morphological erosion on opening image $O_s(X^{(n)})$ with the structuring element S, i.e.,

$$F_1 = O_s(X^{(n)}) \ominus sB$$

- Apply gray-scale morphological dilation on the closing image $C_s(X^{(n)})$ with disc structuring element S, i.e.,

$$F_2 = C_s(X^{(n)}) \oplus sB$$

- Calculate the intermediate gradient image

$$F = F + (F_2 - F_1)$$

Step 3: Normalize the elements of F to [0,1] as

$$F(i,j) = \text{Sigmoid}\left(F(i,j) - \frac{F_{max} + F_{min}}{2}\right)$$

where F_{max} and F_{min} are maximum and minimum values in F and Sigmoid is a logistic function defined as

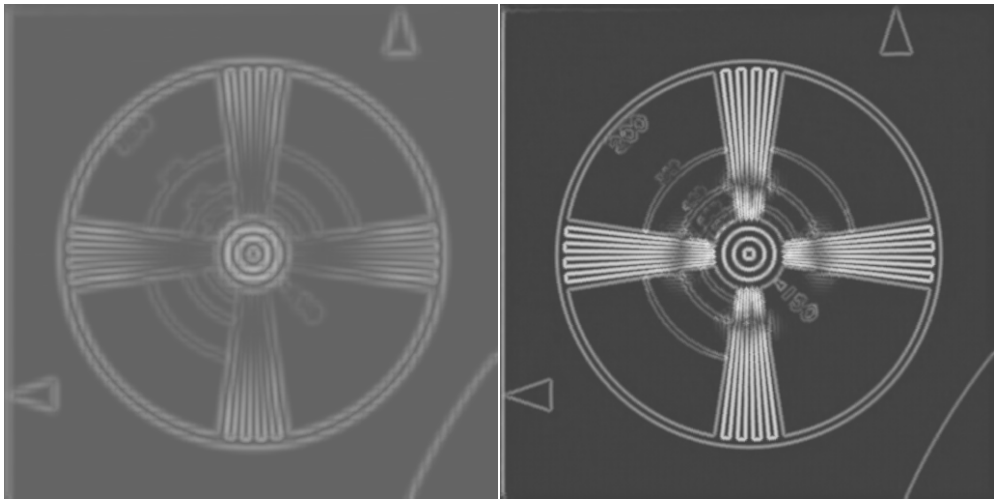
$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-\tau x}}$$

Step 4: Then form a diagonal matrix I_g whose diagonal elements are obtained from F by ordering its pixels lexicographically.

The maximum size S of structuring element is chosen depending upon the contrast of the estimated image and it could be $2 - 3$. If the image is of high contrast S is chosen to be small. Here we choose the value of S adaptively. As Initial estimated image is blurred, a high value of S is chosen, and as the iteration proceeds the value of S decreases.

Since $(F_2 - F_1)$ contains significant value in the high-frequency region, and small value in relatively smooth region. So resultant F is the matrix having high value in the middle of a thick edge and then gradually decreases on both sides. So I_g is constructed in such a way that strong edge or high frequency regions are emphasized.

Figure 2.1 shows an example of gain control image during estimation of HR chart image of size 512×512 from 8 LR chart images of same size 102×102 with different sub-pixel shifts. Figure 2.1(a) is the gain image computed from initial estimate of HR image and Fig. 2.1(b) shows the gain image after 10^{th} iteration, here value of S is taken as 3.



(a) (b)

Figure 2.1: An example of gain map I_g during estimation of HR chart image. (a) the gain map at the first iteration, and (b) the gain map at the 10^{th} iteration.

2.3 Experimental results and discussion

In this section we evaluate and analyze the performance of the proposed SR method and compare the same with that of the other relevant state-of-the-art methods. In this discussion the proposed algorithm (2.13) is referred to as 'GCBTV', while the closest competitors, known as the robust SR method [33] and a locally adaptive technique [59], are referred to as 'BTV'(bilateral total variation) and 'LABTV'(locally adaptive bilateral total variation) respectively. 'Bicubic' and 'TV' represent bicubic interpolation and total variation regularization methods. Another recently proposed regularization techniques Phase-driven spatially variant regularization ('PDSV') [109]. Other methods are also mentioned in a similar fashion.

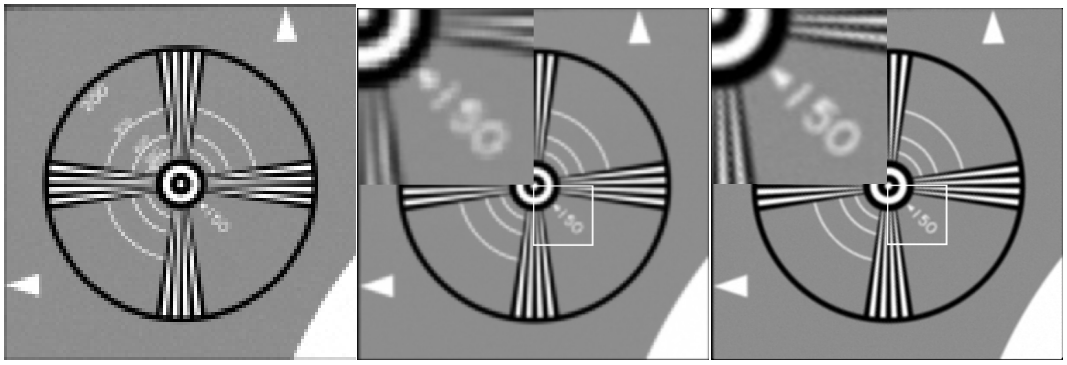
For performance evaluation of the SR image reconstruction methods a typical 512×512 graylavel HR image (called 'chart' Image) is chosen [see Fig. 2.2(a)]. We synthetically generate some LR images from this HR Image and later reconstruct the HR image only from those generated LR images. Then we compute PSNR with respect to the original HR image as a quantitative measure of quality of the reconstructed image. To generate the LR images we have considered only

translational motion of camera resulting sub-pixel shifts. An example of such LR images is shown in Fig. 2.2(b) where top-left quadrant displays the zoomed version (for better manual evaluation) of the portion of the bottom-right quadrant marked by a white box. Other images also have similar kind of display. The blurring is chosen as 5×5 weighted averaging where weights are taken from Gaussian function with standard deviation $\sigma = 2.5$ and the blurring matrix H is formed accordingly. The downsampling factor is chosen to be 5 and the matrix D is constructed to achieve this. Theoretically we need 25 LR images for exact reconstruct of SR image under ideal and noise free environment. However, in our experiment we observe that 8 LR images are sufficient to reconstruct SR image of acceptable quality (say, PSNR $>$ 30 dB) with resolution factor 5. The detail experiment with different number of LR images is shown in Fig. 2.3. Hence, in all our experiment we have taken 8 randomly chosen LR images and reconstruct SR image with resolution factor 5, and the same setup is followed throughout this experimental section unless stated otherwise.

The parameters for each algorithm are chosen (in limited number of trials) to produce visually most appealing results. For fair evaluation purpose, we have applied each algorithm for SR reconstruction on different sets of 8 LR images with different parameters and the best result of each algorithm is chosen as the output of the algorithm for this experiment. In our reconstruction algorithm (2.13), we have chosen model parameters as $S = 2$, $\alpha = 0.5$, $\beta = 0.5$ and $\omega = 0.7$, respectively, and got superior result in most of the experiments. In the iterative algorithm (2.13), initial estimate $\mathbf{X}^{(0)}$ is taken as just average of all the up-sampled LR images.

In the first experiment, LR images contain small amount of Gaussian noise with zero mean and standard deviation $\sigma = 2$. To start with, 8 LR images are upsmapled and averaged to obtain initial estimation as shown in Fig. 2.2(b). The results of different approaches are shown in Fig. 2.2(c-i). A white bounding box is marked on each image and is displayed on the top-right corner after zooming for careful study of the quality.

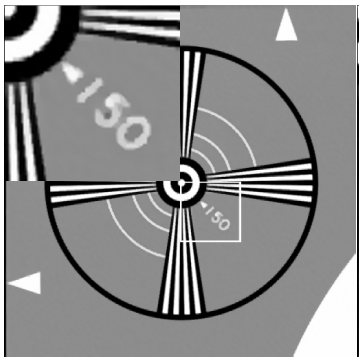
We have tested the SR image reconstruction algorithm on some other images and the results for boat, man and lenna images are shown in Figs. 2.3-2.5 respectively. In each of the images the proposed method is able to reconstruct much



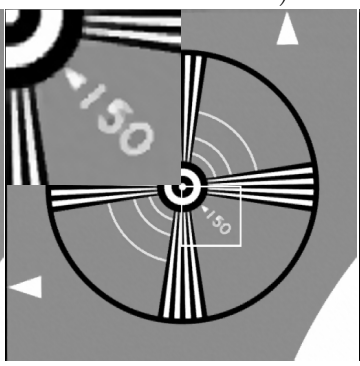
(a) *LR Image*

(b) *Initial HR Image* (PSNR = 22.43, SSIM = 0.9365)

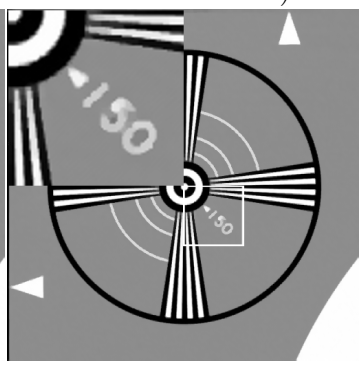
(c) *Bicubic* (PSNR = 21.70, SSIM = 0.9470)



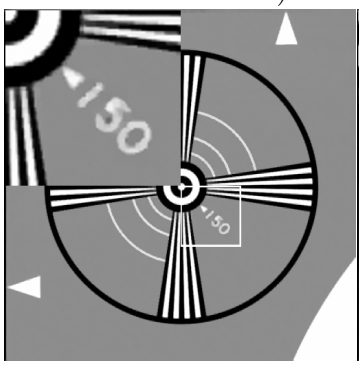
(d) *TV* (PSNR = 29.81, SSIM = 0.9893)



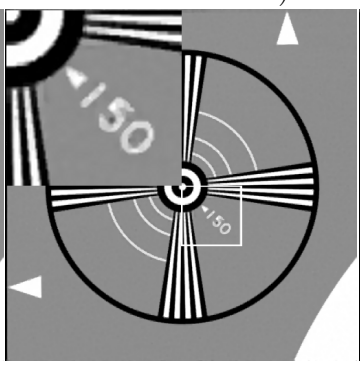
(e) *GCTV* (PSNR = 30.62, SSIM = 0.9904)



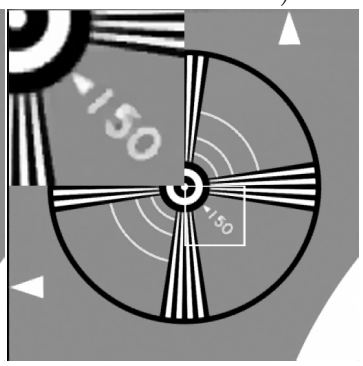
(f) *BTV* (PSNR = 29.91, SSIM = 0.9899)



(g) *PDSV* (PSNR = 30.81, SSIM = 0.9928)



(h) *LABTV* (PSNR = 30.70, SSIM = 0.9925)



(i) *GCBTV* (PSNR = 30.88, SSIM = 0.9931)

Figure 2.2: The results of various SR image reconstruction techniques: (a) One of the LR chart images with Gaussian noise (mean $\mu = 0$ and standard deviation $\sigma = 2$), (b) upsampled and merged 8 LR images, (c) HR image reconstruction using bicubic interpolation considering those 8 LR image (d) and (f) SR reconstructed image with TV and BTV Morphological regularization, respectively, and (e) and (i) are the corresponding improvement with gain controlled technique. Note that (f)-(i) show the result of Farsiu *et al.* [33], Tian *et al.* [109], Li *et al.* [59], and that of proposed method, respectively.

more detail than existing state-of-art reconstruction techniques. We achieve significant improvement in PSNR for each of the images. We also observe that in addition of quality of the image, the speed of convergence of the iterative process is boost up. Moreover it takes only $50-75$ iteration to achieve the results shown in Figs. 2.2-2.5. Most noticeable observation is that the proposed gain control method improves over both the results due to TV and BTV significantly. Proposed method reconstructs SR images with higher PSNR than that of the recent techniques due to Li *et al.* [59](LABTV) and Tian *et al.* [109](PDSV). Hence, it may be inferred that the proposed gain control method is helpful to obtain SR images of better quality and also to remove ringing effect while preserving edge.

We plot a bar diagram of PSNR values over different number of LR images of chart image of Fig. 2.2 in Fig. 2.3. Since typical values for the PSNR in lossy image and video compression are taken to be grater than 30db, we consider this limit as satisfying HR image quality for SR image reconstruction. This experiment suggests that 8 LR images may be enough to get SR image with resolution factor is 5. This observation is useful because large number of LR images (i.e. 25 LR images for resolution factor 5) may not be available always.

In the next experiment we evaluate the performance of our reconstruction model under mis-prediction of motion parameter. We choose two LR images among eight LR images and change their sub-pixel shifts deliberately by adding small random number lying in the range [-0.5, 0.5]. Sub-pixel shifts of the other LR images are preserved. In Fig. 2.7 we show the results of this experiment. Figures 2.7(a) and 2.7(b) are the reconstructed images using PDSV and LABTV regularization respectively. Figure 2.7 (c) shows the reconstructed image using our method. This shows that even if we commit some error in motion estimation, PSNR of our GCBTV regularization technique is higher than state-of-the-art regularization techniques. Thus experiment illustrates that our method is more robust under mispredicted motion model.

In the final experiment, we deal with a compressed sequence of 9 real images of size (96×128) (containing translational motion approximately) from a commercial video camera (3COM, Model no. 3718); (courtesy of Sina Farsiu and Dirk Robinson [68]). The (unknown) camera PSF is assumed to be a 3×3 Gaussian kernel with standard deviation (σ) equal to 1.5. We have used the method

(*a*) *LR Image*

(*b*) *Initial HR Image* (PSNR = 21.79, SSIM = 0.9012)

(*c*) *Bicubic* (PSNR = 26.18, SSIM = 0.8913)

(*d*) *TV* (PSNR = 30.09, SSIM = 0.9519)

(*e*) *GCTV* (PSNR = 30.13, SSIM = 0.9525)

(*f*) *BTV* (PSNR = 30.17, SSIM = 0.9521)

(*g*) *PDSV* (PSNR = 30.14, SSIM = 0.9533)

(*h*) *LABTV* (PSNR = 30.02, SSIM = 0.9507)

(*i*) *GCBTV* (PSNR = 30.19, SSIM = 0.9528)

Figure 2.3: Comparison of reconstruction results of boat image of our technique with the existing regularization techniques : (i) a LR image, (ii) HR image reconstructed using bicubic interpolation, (iii) SR image reconstructed using TV regularization, (iv) gain controlled TV regularization, (v) BTV regularization [33], (vi) PDSV regularization [109], (vii) LABTV regularization [59] and (viii) gain controlled BTV regularization[our method] respectively. A small cropped portion is cropped and zoomed before displayed in the top-left corner of each images for better visual clarity.

(a) *LR Image*

(b) *Initial HR Image* (PSNR = 20.37, SSIM = 0.8952)

(c) *Bicubic* (PSNR = 27.28, SSIM = 0.8948)

(d) *TV* (PSNR = 30.11, SSIM = 0.9475)

(e) *GCTV* (PSNR = 30.19, SSIM = 0.9481)

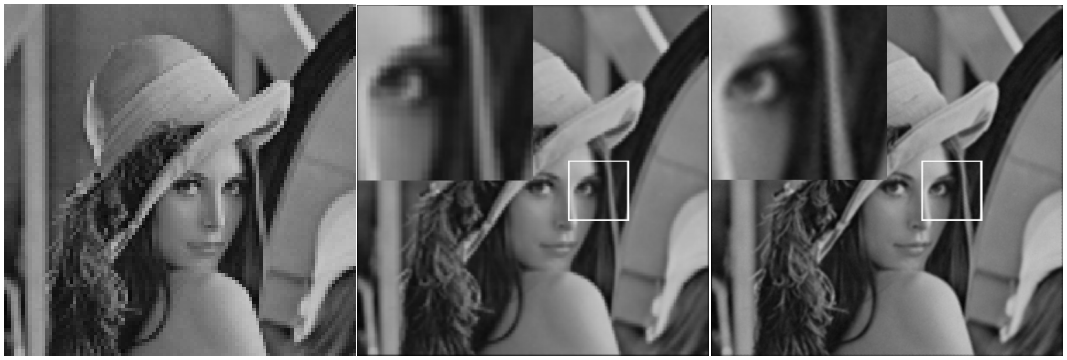
(f) *BTV* (PSNR = 30.17, SSIM = 0.9481)

(g) *PDSV* (PSNR = 30.17, SSIM = 0.9458)

(h) *LABTV* (PSNR = 30.07, SSIM = 0.9486)

(i) *GCBTV* (PSNR = 30.25, SSIM = 0.9488)

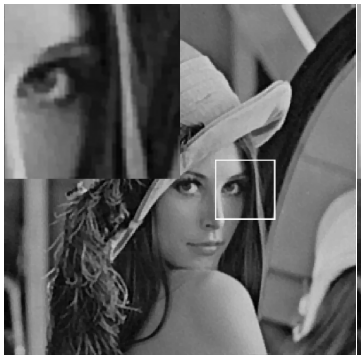
Figure 2.4: Comparison of reconstruction results of man image of our technique with the existing regularization techniques : (i) a LR image, (ii) HR image reconstructed using bicubic interpolation, (iii) SR image reconstructed using TV regularization, (iv) gain controlled TV regularization, (v) BTV regularization [33], (vi) PDSV regularization [109], (vii) LABTV regularization [59] and (viii) gain controlled BTV regularization[our method] respectively. A small cropped portion is cropped and zoomed before displayed in the top-left corner of each images for better visual clarity.



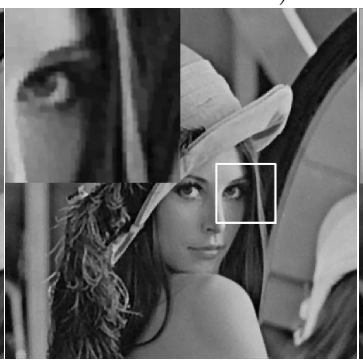
(a) *LR Images*

(b) *Initial HR Image* (PSNR = 23.61, SSIM = 0.9038)

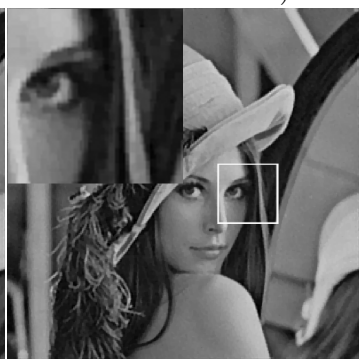
(c) *Bicubic* (PSNR = 28.24, SSIM = 0.9294)



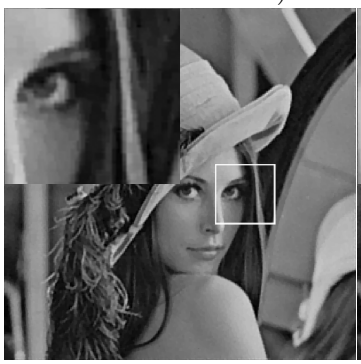
(d) *TV* (PSNR = 31.77, SSIM = 0.9591)



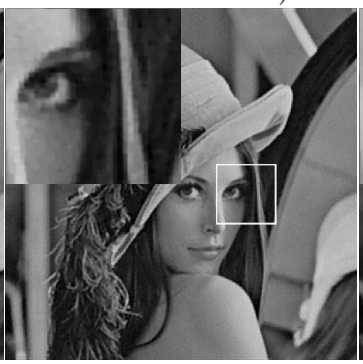
(e) *GCTV* (PSNR = 31.81, SSIM = 0.9593)



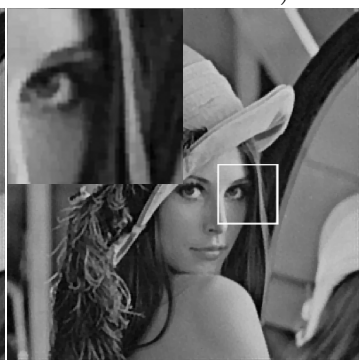
(f) *BTV* (PSNR = 31.80, SSIM = 0.9591)



(g) *PDSV* (PSNR = 31.75, SSIM = 0.9598)



(h) *LABTV* (PSNR = 31.54, SSIM = 0.9531)



(i) *GCBTV* (PSNR = 31.83, SSIM = 0.9592)

Figure 2.5: Comparison of reconstruction results of Lena image of our technique with the existing regularization techniques : (i) a LR image, (ii) HR image reconstructed using bicubic interpolation, (iii) SR image reconstructed using TV regularization, (iv) gain controlled TV regularization, (v) BTV regularization [33], (vi) PDSV regularization [109], (vii) LABTV regularization [59] and (viii) gain controlled BTV regularization[our method] respectively. A small cropped portion is cropped and zoomed before displayed in the top-left corner of each images for better visual clarity.

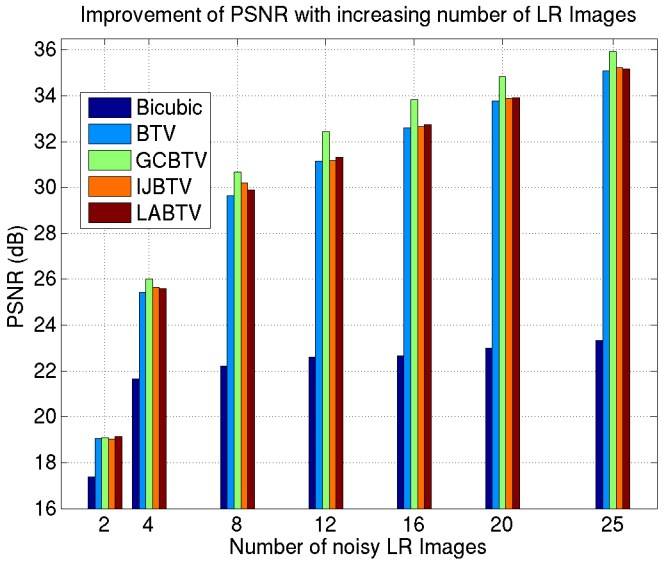
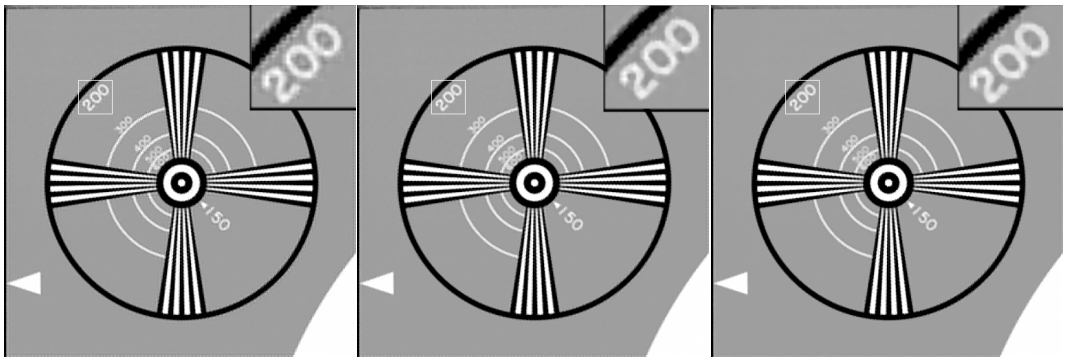


Figure 2.6: Comparison of reconstruction result with different number of noisy LR images while resolution factor is five



(a) *PDSV* (PSNR = 27.88, SSIM = 0.9206)

(b) *LABTV* (PSNR = 28.15, SSIM = 0.9459)

(c) *GCBTV* (PSNR = 28.39, SSIM = 0.9475)

Figure 2.7: Comparison of reconstruction result of chart image for misinterpreted motion model (subpixel-shifts) (a) SR image reconstructed using PDSV regularization [109] (b) using LABTV regularization [59] and (c) using our GCBTV regularization method for HR image reconstruction. A small cropped portion is cropped and zoomed before displayed in the top-right corner of each images.

(a) $LRframe$ (a) $PDSV$

(c) $LABTV$ (d) $GCBTV$

Figure 2.8: Comparison of reconstruction result of real Emily image sequence (a) a sample LR image frame (b) SR image reconstructed using PDSV regularization [109] (b) using LABTV regularization [59] and (c) using our GCBTV regularization method for HR image reconstruction. A small cropped portion is cropped and zoomed before displayed in the top-left corner of each images.

described by Vanderwalle *et al.* [114] to compute the motion vectors assuming translational motion. Figure 2.8(a) is one of these LR images and Fig. 2.8(b)-(d) are the result of the PDSV [109], LABTV [59] and our GCBTV method respectively. We compare the result in Fig. 2.8 visually and it is clear that artifacts occur in Fig. 2.8(b)-(c) due to error in blur parameter and approximation in motion vector are much less in Fig. 2.8(d).

2.4 Summary

In this chapter, we have presented a gain controlled regularized SR image reconstruction algorithm that can enhance the quality of HR image produced from a set of noisy blurred LR images. The iterative algorithm based on algebraic reconstruction technique (ART) and expectation maximization is used for reconstruction in computer tomography using L_2 norm. One of our contributions is to apply this method for two-dimensional SR image reconstruction and also to modify the algorithm for L_1 norm as dissimilarity measure. This makes the convergence of the algorithm much faster than the conventional iterative methods employed in SR image reconstruction. Our second contribution is to propose a mathematical morphology based edge-preserving adaptive regularization method for super-resolution imaging through a robust gain control. We have shown that the proposed gain control method preserves edges and suppresses the propagation of ringing artifacts occurred during iterative process. An experimental comparative study reveals that the proposed method performs better in terms of quality than that of some recent works.

Chapter 3

Morphologic Regularization through Bregman Iteration

The conventional regularization terms for SR image reconstruction lead to a stable solution as discussed in Section 1.3.1 and also in the last Chapter. However, a periodic noise with ringing artifacts are still present in the generated HR image whenever we try to capture details of the HR image. The problem is lessened to some extend as discussed in Chapter 2. Note that, there we employed a gain-controlled function over the well-known BTV regularization. Moreover, the existing iterative SR methods are painfully slow. Hence it is desirable to have faster and more efficient reconstruction. We do this by employing better optimization technique as well as appropriate regularization term. For example, with gradient descent optimization technique the LABTV [59] regularization outperforms BTV [33] which gives better result than TV [91] regularization; on the other hand, based on TV regularization Marquina and Osher [66] obtain superior result employing Bregman iteration than the usual gradient descent techniques. So we envisage that even better result would be obtained by combining Bregman iteration and a more sophisticated regularization method that can suppress noise in LR images and ringing artifacts occurred during the reconstruction of HR image.

In this chapter, we propose a new regularization technique based on non-linear multi-scale morphologic operator and also formulate SR reconstruction problem

as a deblurring problem. Morphological operator is a well-known tool that extracts structure from an image [102]. It is also used in image denoising [71, 73], image segmentation [116, 72] and image fusion [86, 22] successfully. In Chapter 2, we have developed an adaptive regularization technique by assigning a gain value at every pixel based on morphological operators. Here we intend to build an edge-preserving morphologic regularization technique for SR image reconstruction. The conventional regularization methods try to limit high-frequency energy and thus ensure smoothness. However, as both the noisy pixels and edge pixels contain high-frequency energy, the resultant de-noised image does not contain sharp edges. It is well known that morphological opening and closing removes bright and dark noise, respectively, without affecting the edge sharpness [71, 102]. Now an image, in general, may contain both detail and noise at different scale. So we propose a multiscale morphology based regularization to remove noise and artifacts, while preserving the edge sharpness.

Second, usually iterative SR image reconstruction techniques use gradient descent approach to obtain optimal solution. In Chapter 2, we have used modified ART like algorithm for this task. However, this is an extension of gradient descent technique. Since gradient descent approach is very slow, our objective here is to find out a way to make faster convergence of iterative process for real time applications. Marquina and Osher [66] are first to use Bregman iteration for fast SR image reconstruction with TV regularization. Here we develop a stack of Bregman iterations for multiscale morphological regularization using efficient subgradients. Thus, our proposed algorithm produces results that are free from aforementioned periodic noise and also enjoys speedy convergence.

3.1 Revisiting problem formulation for SR reconstruction

The reconstruction of HR image from given a set of LR images is a reverse process as discussed in Section 1.3. Given K LR images $\mathbf{Y}_k$ and the model parameters of D, H and F_k, our goal is to estimate $\mathbf{X}$ by solving the system of linear equations

described in Eq.(1.5).

$$\mathbf{Y}_k = DF_kH\mathbf{X} + \mathbf{e}_k, \qquad \forall k = 1, 2, ..., K \tag{3.1.1}$$

In Chapter 2, we solve the above system of equations directly using modified gradient descent technique. However, in most of the studies [33, 110, 111] the SR reconstruction is addressed in two steps:

- Finding a blurred HR image $\hat{\mathbf{Z}} \approx H\mathbf{X}$ from the LR measurements $\mathbf{Y}_k$.
- Estimating the de-blurred image $\hat{\mathbf{X}}$ from $\hat{\mathbf{Z}}$.

Instead of addressing the problem in two steps, i.e., first estimating blur image and then estimating actual HR image from blur image, we reformulate SR reconstruction as a de-blurring problem and minimize the error in a single step.

In Eq. (3.1.1) D and H are same for all LR images, since we have assumed that all LR images have same resolution factor and same blurring operator. We avoid down-sampling and then up-sampling at each iteration of iterative reconstruction algorithm [see Eq. (2.13)] by up-sampling each LR images $\mathbf{Y}_k$ and fill the pixels with zero where values are unknown ensuring that these pixel values will not participate in estimating the HR image.

Suppose $\underline{\mathbf{Y}}_k$ denotes the up-sampled k-th LR image obtained through reverse effect of DF_K of Eq. (3.1.1), then

$$\underline{\mathbf{Y}}_k = R_k(H\mathbf{X} + \underline{\mathbf{e}}_k), \qquad \forall k = 1, 2, ..., K \tag{3.1.2}$$

where R_k is an index matrix, in general, a banded matrix where all elements $R(i, j)$ is zero if $|i - j|$ is greater than the bandwidth, and all other elements in a row is $\frac{1}{|i-j|+1}$ if corresponding element is present in $\underline{\mathbf{Y}}_k$ and zero otherwise. For brevity, let us take R_k to be a diagonal matrix (i.e., bandwidth is equal to zero), whose diagonal elements are one if corresponding elements of $\underline{\mathbf{Y}}_k$ have come from $\mathbf{Y}_k$ and zero otherwise. In other words, if an element of $\underline{\mathbf{Y}}_k$ is not known from $\mathbf{Y}_k$, that element is assumed to be zero, and all elements of the corresponding row of R_k are also zero. Thus assumed elements of $\underline{\mathbf{Y}}_k$ cannot participate in estimating $\mathbf{X}$ based on Eq. (3.1.2). Suppose by ensembling all of the available $\underline{\mathbf{Y}}_k$s we get

$\underline{\mathbf{Y}}$ [as shown in Fig. 3.1(f)], then

$$\underline{\mathbf{Y}} = R(H\mathbf{X} + \underline{\mathbf{e}}) \tag{3.1.3}$$

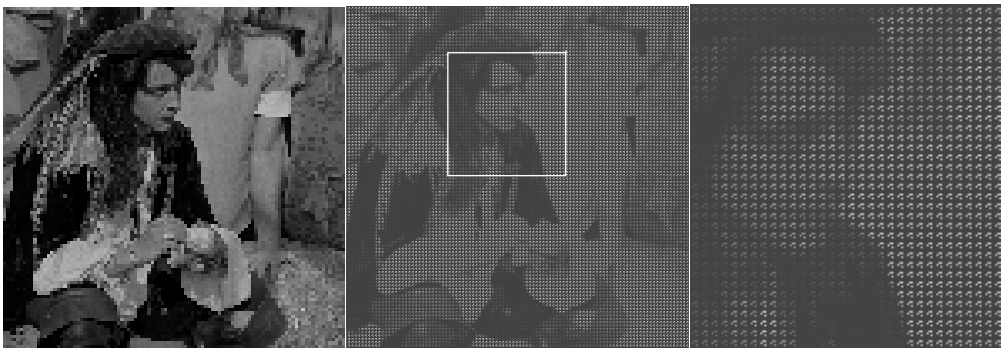
Now $\underline{\mathbf{Y}}$ contains maximum information about HR image $\mathbf{X}$ available from LR images $\mathbf{Y}_k$ and most of diagonal elements of R are one. In fact, if $K = N = l_x l_y$, R becomes an identity matrix and all the elements of $\underline{\mathbf{Y}}$ are known. If there is only sub-pixel shifts for LR images, after up-sampling there would be only one nonzero value for a pixel for all distinct up-sampled LR images. As a result we get $\underline{\mathbf{Y}}$ as a blurred version of actual HR image $\mathbf{X}$ except some pixel values would be missing, since we take only a few LR images and index matrix R keeps track of those missing values.

Thus if there is $l_x l_y$ distinct LR images, and warping matrix F_k consists of only translational pixel shifts, then R may be dropped from Eq. (3.1.3) and SR reconstruction of $\mathbf{X}$ becomes similar to de-blurring method of a blurred image $\underline{\mathbf{Y}}$. In Fig. 3.1(a) shows one of four LR images $\mathbf{Y}_k$ (i.e., $k = 1, 2, 3, 4$) with resolution factor five with different pixel shifts, and Figs. 3.1(b)-(c) show corresponding Y and cropped portion respectively. So by up-sampling followed by reversing pixel shifts and finally merging $\mathbf{y}^k$s, we get blurred image $\underline{\mathbf{Y}}$ where the unknown pixels are filled with zero.

Since Eq. (3.1.3) consists of less number of non-identical equations than the dimension of the solution space and coefficient matrix RH is usually a singular matrix during SR reconstruction some unwanted noise would be introduced in the solution space. As a result estimated image $\hat{\mathbf{X}}$ will have some ringing noise in high frequency region. So we need to regularize the solution space to get noise-free image. The intuition behind this regularization method is to limit the high-frequency energy forcing the smoothness.

3.1.1 L2 error based estimation of SR image

Since observed LR images are noise-corrupt and more importantly $K < N/M = L_x L_y$ or, in other words, all diagonal elements of R are not 1, SR image reconstruction becomes an ill-posed problem. Second, number of unknowns, i.e., number of Pixels in $\mathbf{X}$ is usually very large. So obtaining direct solution of Eq. (3.1.3) may



(a) (b) (c)

Figure 3.1: Illustrates up-sampling from LR images, (a) one of 4 LR Images with down-sampling factor 5, (b)-(c) combined all 4 up-sampled LR images with corresponding pixel shifts, and (c) cropped portion of the up-sampled image (b)

not be feasible using direct inverse methods. Instead, we estimate an HR image $\hat{\mathbf{X}}$, which when degraded minimizes $\rho(RHX, \underline{Y})$.i.e.

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left[\rho(RH\mathbf{X}, \underline{\mathbf{Y}})\right] \tag{3.1.4}$$

The operator ρ is a distance measure between the observation and an estimate and is defined as $\rho(U, V) = \frac{1}{p}\|U - V\|_p^p$, where $1 \leq p \leq 2$. Thus the minimization of data error term $\rho(RHX, \underline{Y})$ tries to preserves edge information and texture details in the estimated HR image. If we choose $p = 2$, then the estimated HR image satisfies Eq. (3.1.3) in, say, least-square sense, i.e.,

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left[\|(RH\mathbf{X} - \underline{\mathbf{Y}})\|_2^2\right] \tag{3.1.5}$$

or, in terms of its first derivative, we find an estimate X to satisfy

$$H^T R^T (RH\mathbf{X} - \underline{\mathbf{Y}}) = 0 \tag{3.1.6}$$

Thus we end up with solving KM number of linear equations as given in Eq. (3.1.6), where K and M are the number and size of LR images respectively. Please note that the noise term $\mathbf{e}$ is not considered in arriving the solution from these equations. Also note that the (.) part of Eq. (3.1.5) measures the dissimilarity as

L_2 norm between up-sampled observed LR image and the marked pixels of re-degraded estimated HR image. Because of ill-condition nature of the problem, the solution set of linear equations presented in Eq. (3.1.6) is unstable. That means a small perturbation in $\underline{\mathbf{Y}}$ due to noise results in a non-ignorable error in the solution $\mathbf{X}$. So to obtain stable solution of $\mathbf{X}$ based on a few observed LR images $\mathbf{Y}^k$ and the degradation matrix H, it becomes necessary to impose regularization on the solution.

3.1.2 Regularization for SR reconstruction algorithm

The intuition behind the regularization is to ensure that the reconstructed image satisfies some quality criterion as discussed in Section 1.3.1. From Bayesian point of view, regularization technique corresponds to imposing certain prior distributions on model parameters. For example, regularization may help to ensure smoothness and removes artifacts from the final answer, and also to improve rate of convergence. Thus our intention is to develop a regularization method which results in HR images with sharp edges and fine textures while suppressing the noise with low computational cost. Regularization has already been used in conjunction with iterative techniques for restoration of noisy and degraded images [11, 59, 49, 76, 77, 127]. It involves introducing additional information in order to solve an ill-posed problem and prevents over-fitting.

State-of-the-art regularization

To obtain a stable solution, suppose a specific regularization operator $\Upsilon(\mathbf{X})$ is imposed on the estimated HR image $\mathbf{X}$. The regularization operator $\Upsilon(\mathbf{X})$ can incorporate prior knowledge of the desirable HR solution, e.g., degree of smoothness. Therefore SR image reconstruction (3.1.4) can simply be reformulated as:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \Upsilon(\mathbf{X}) \text{ such that } \rho(RH\mathbf{X}, \underline{\mathbf{Y}}) \leq \sigma \tag{3.1.7}$$

where σ is the noise variance in LR images usually assumed to be small. In the case of SR reconstruction from noise free LR images above formulation reduces

to

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \Upsilon(\mathbf{X}) \text{ such that } RH\mathbf{X} = \underline{\mathbf{Y}} \tag{3.1.8}$$

However, solving Eq. (3.1.8) directly based on a few observed LR images where the number of unknowns is large is not a pragmatic approach. So we try to solve these equations using a suitable iterative method. In Section 3.3 we discuss about this issue in detail and show how we can incorporate the effect of noise using the concept of adding "noise back" in Bregman iteration. These additional constraints in Eqs. (3.1.7) and (3.1.8) that favor well-behaved solution can be converted to a generalized minimization cost function [49, 127], i.e.,

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left[\rho(RH\mathbf{X}, \underline{\mathbf{Y}}) + \mu\Upsilon(\mathbf{X})\right] \tag{3.1.9}$$

where μ is the Lagrangian constant called the regularization parameter that controls the emphasis between the data error term (first term) and the regularization term (second term). In terms of L_2 norm, the Eq. (3.1.9) becomes:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left[\frac{1}{2}\|RH\mathbf{X} - \underline{\mathbf{Y}}\|_2^2 + \mu\Upsilon(\mathbf{X})\right] \tag{3.1.10}$$

Choosing regularization parameter μ for optimum solution of (3.1.10) is a non-trivial task. For example, a large value of μ preserve sharp edges as data error term is less emphasized and on the other hand, small value of μ may amplify unwanted ringing artifacts as smoothness criteria gets less importance. In general above minimization problem can be solved by usual iterative gradient descent [11, 29, 27, 33, 99] method. During iterative reconstruction, since both the noise and edges contribute to high frequency component, they both are usually removed simultaneously during the regularization process enforcing smoothness. The problem may be surmounted partially by choosing a decreasing sequence of regularizing parameter to enforce smoothness at the early stages but finally getting HR image with with sharp edges. In that case, of course, another parameter determining the 'rate of decrease' of regularization parameter has to be introduced. Clearly, if rate of decrease is high then there may not be any effect due to regularization and, on the other hand, if it is low then the solution may not satisfy the equality criterion. In Section 3.3, employing an iterative algorithm

that minimizes the regularization error, the equality constraint (3.1.8) is satisfied for a wide range of value of regularization parameter.

One of the commonly used regularization method is the Tikhonov cost function [11, 59, 76, 77, 127, 29], for making stable estimated de-noised image (1.8). In Section 1.3.1, we also discussed different types of state-of-the-art regularization methods existing in literature. Most of those regularization terms $\Upsilon(X)$ are function of $\Gamma\mathbf{X}$ where Γ is an high-frequency operator, or function of $(\mathbf{X} - S_x^l S_y^m \mathbf{X})$ which is first order image derivatives in different resolution along x-axis and y-axis. Thus conventional regularization methods choose $\Upsilon(X)$ as high-frequency energy and minimize its p^{th} norm to ensure smoothness. In this chapter we define $\Upsilon(X)$ based on morphologic filters that preserve structures and suppress noise. An image, in general, may contain both detail and noise at different scales. So we propose a multiscale morphology based regularization to remove noise and artifacts, while preserving the edge sharpness in the next section.

Moreover, some of those regularization terms are differentiable. However, most of the edge preserving regularization (1.9)-(1.10) terms are non-differentiable and are harder to optimize. Usually, iterative SR image reconstruction methods based on differentiable regularization terms with L_p norm $(1 < p \leq 2)$ [29, 127, 59] use gradient descent approach to obtain optimal solution. On the other hand, quite a few techniques [26, 42, 90] are available in the literature to efficiently handle non-differentiable regularization term with L_1 norm (e.g., TV regularization). A group of solvers, evolved from Bregman iteration [78], is one of recently developed methods for such non-differentiable constraint optimization problems. Since gradient descent approach is painfully slow, our objective here is to find out a way to make faster convergence of iterative process for real time applications. Marquina and Osher [66] are first to use Bregman iteration for fast SR image reconstruction with TV regularization.

We introduce proposed multi-scale morphologic regularization method in Section 3.2. In Section 3.3, we review Bregman iteration and develop an algorithm for SR image reconstruction with proposed morphologic regularization. Section 3.4 presents the experimental results and compares our results with that of some existing regularization methods for SR image reconstruction. Finally, some concluding remarks are made in Section 3.5.

3.2 Morphologic regularization

Let B be a disk of unit size with origin at its center and sB be a disk structuring element (SE) of size s. Then the morphological dilation $D_s(\mathbf{X})$ of an image $\mathbf{X}$ of size $m \times n$ at scale s is defined as:

$$D_s(\mathbf{X}) = \begin{pmatrix} \max_{r\in(sB)_{(1)}}\{x_r\} \\ \max_{r\in(sB)_{(2)}}\{x_r\} \\ \ddots \\ \max_{r\in(sB)_{(mn)}}\{x_r\} \end{pmatrix} \tag{3.2.1}$$

where $(sB)_{(i)}$ is a set of pixels covered under SE sB translated to the i-th pixel x_i. Similarly morphological erosion $E_s(\mathbf{X})$ at scale s is defined as:

$$E_s(\mathbf{X}) = \begin{pmatrix} \min_{r\in(sB)_{(1)}}\{x_r\} \\ \min_{r\in(sB)_{(2)}}\{x_r\} \\ \ddots \\ \min_{r\in(sB)_{(mn)}}\{x_r\} \end{pmatrix} \tag{3.2.2}$$

Morphological opening $O_s(\mathbf{X})$ and closing $C_s(\mathbf{X})$ by SE sB are defined as:

$$O_s(\mathbf{X}) = D_s(E_s(\mathbf{X}))$$

$$C_s(\mathbf{X}) = E_s(D_s(\mathbf{X}))$$

In multiscale morphological image analysis [71, 72, 6, 125], we have seen that the difference between s^{th} scale closing and opening extracts noise particles and image artifacts in scale s and may be used for de-noising purpose. So, in this chapter, we propose the regularization function based on multi-scale morphology as:

$$\Upsilon(\mathbf{X}) = \sum_{s=1}^{S} \alpha^s \mathbf{1}^t [C_s(\mathbf{X}) - O_s(\mathbf{X})] \tag{3.2.3}$$

where $\mathbf{1}$ is a column vector consisting of all 1's and α is the weighting coefficient. To give more emphasis on the small scale for noise removal, the value of α is chosen from the interval $0 < \alpha < 1$. Therefore, with the proposed regularization

term the SR reconstruction problem (3.1.8) is reduced to :

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}}\{\sum_{s=1}^{S} \alpha^s \mathbf{1}^t[C_s(\mathbf{X}) - O_s(\mathbf{X})] \ : \ \|RH\mathbf{X} - \underline{\mathbf{Y}}\|_2 < \eta\} \tag{3.2.4}$$

where the regularization function is $\Upsilon(\mathbf{X}) = \sum_{s=1}^{S} \alpha^s \mathbf{1}^t(C_s(\mathbf{X}) - O_s(\mathbf{X}))$. In general, above constraint minimizing problem is solved by converting it into unconstrained optimization problem (3.1.8) as discussed in previous section and then by using gradient descent [11, 29, 33, 99] method to optimize it. Since the proposed regularization term is not differentiable everywhere, we use subgradient technique to solve the above reconstruction problem as discussed in the next section.

3.3 Subgradient methods and Bregman iteration

The first discover in the field of the iterative regularization for constrained optimization problems (3.1.8) with TV norm was proposed with great success is Bregman iteration. Bregman iteration is successfully used by Osher *et al.* [78] in the field of computer vision for finding the optimal value of energy functions in the form of a constrained convex functional. After that a class of efficient solvers has been proposed for constrained problems (3.1.8) and also unconstrained problems (3.1.10). Among them 'fixed point continuation' (FPC) [42] method is proposed to solve the unconstrained problem by performing gradient descent steps iteratively. Linearized Bregman algorithm [123] is derived by combining FPC and Bregman iteration to solve constrained problem in a more efficient way. Those methods are successfully used in sparse reconstruction problem viz. compressed sensing (CS) [26, 42, 123, 8, 79] and sparse coding [59, 9, 21, 14] due to their simplicity, efficiency and stability. However, they were not used to solve problems involving regularization term with high pass filter Γ. In fact, it was difficult to apply the FPC and linearized Bregman methods to image processing applications. Later Goldstein *et al.* [39] have developed 'split Bregman method' for more structured regularization in variational problems of image processing. It combines the

concept of Bregman iteration [78] and forward-backward operator splitting [18] to optimize constrained functional. Based on a constrained variational model that uses the total variation of the signal as a regularizing functional, Marquina and Osher [66] have formulated a model for SR . In this section, based on Bregman iteration and proposed morphologic regularization we develop an algorithm for SR image reconstruction problem.

3.3.1 Bregman iteration

Consider the following minimization problem:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \{\Upsilon(\mathbf{X}) \ : \ T(\mathbf{X}) = 0\} \tag{3.3.1}$$

where Υ and T are both convex functionals defined over $\mathbb{R}^n \rightarrow \mathbb{R}^+$. Now it is well known that finding optimal of the above minimization problem is difficult when $\Upsilon(\mathbf{X})$ is non-differentiable. Bregman iteration is a technique that does this job efficiently. Bregman iteration scheme is based on Bregman distance [5] and subgradient at a point. Bregman distance corresponding to convex functional $\Upsilon(.)$ from the point $\mathbf{X}$ to the point $\mathbf{V}$ is defined as

$$B^p_\Upsilon(\mathbf{X}, \mathbf{V}) = \Upsilon(\mathbf{X}) - \Upsilon(\mathbf{V}) - \langle p, \mathbf{X} - \mathbf{V} \rangle \tag{3.3.2}$$

where $p \in \partial\Upsilon$ is a subgradient of Υ at the point $\mathbf{V}$ and $\langle ., . \rangle$ represent the inner product of two vectors. A subgradient p of a convex functional Υ at a point $\mathbf{V}$ satisfies the following condition:

$$\Upsilon(\mathbf{X}) \geq \Upsilon(\mathbf{V}) + \langle p, \mathbf{X} - \mathbf{V} \rangle \tag{3.3.3}$$

for all points $\mathbf{X}$ on Υ neighboring $\mathbf{V}$. Thus if $\partial\Upsilon$ is the set of all possible subgradients at the point $\mathbf{V}$ and it would be a singleton set if Υ is differentiable at the point $\mathbf{V}$. In that case, it can be shown easily from the first order Taylor's expansion that the subgradient p is equivalent to $\nabla\Upsilon$. In Fig. 3.2, we have shown the Bregman distance $B^p_\Upsilon(\mathbf{X}, \mathbf{V})$ for a continuously differentiable function Υ.

Clearly Bregman distance is non-negative for a convex functional, but this

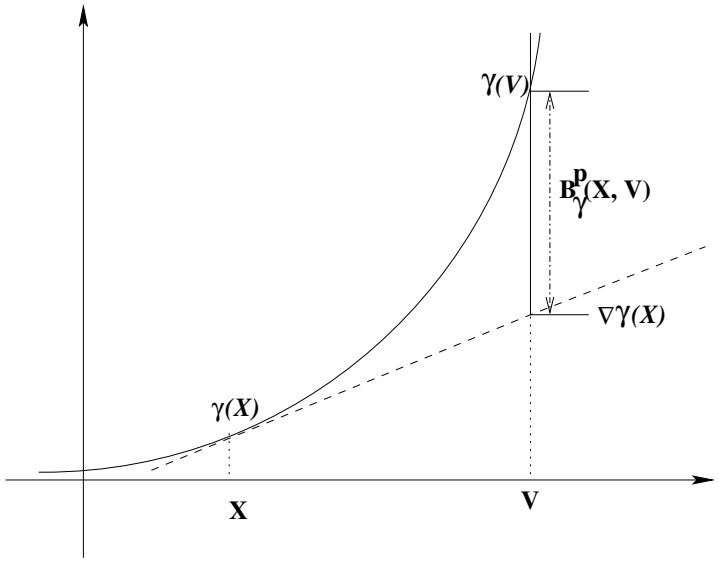


Figure 3.2: Illustrates the Bregman distance $B^p_\Upsilon(\mathbf{X}, \mathbf{V})$ corresponding to convex functional $\Upsilon(.)$ from the point $\mathbf{X}$ to the point $\mathbf{V}$.

is not a metric in usual sense as because it is not symmetric in general, i.e., $B^p_\Upsilon(\mathbf{X}, \mathbf{V}) \neq B^q_\Upsilon(\mathbf{V}, \mathbf{X})$. However, it does measure of closeness in the sense that $B^p_\Upsilon(\mathbf{X}, \mathbf{V}) \geq 0$ and $B^p_\Upsilon(\mathbf{X}, \mathbf{V}) \geq B^p_\Upsilon(\mathbf{W}, \mathbf{V})$ for $\mathbf{W}$ on the line segment connecting $\mathbf{X}$ and $\mathbf{V}$. Now the Bregman iterations [78, 66, 39] that solve the above constrained minimization problem (3.3.1) are as follows:

$$\begin{aligned} &\text{Initialize } \mathbf{X}^0 = p^0 = \mathbf{0} \\ &\begin{cases} \mathbf{X}^{(n+1)} = \arg\min_{\mathbf{X}} \{\mu B^{p^{(n)}}_\Upsilon(\mathbf{X}, \mathbf{X}^{(n)}) + T(\mathbf{X})\} \\ p^{(n+1)} = p^{(n)} - \nabla T(\mathbf{X}^{(n+1)}) \end{cases} \end{aligned} \tag{3.3.4}$$

where $B^{p^{(n)}}_\Upsilon$ is the Bregman distance corresponding to convex functional $\Upsilon(.)$ and is defined from point $\mathbf{X}$ to point $\mathbf{X}^{(n)}$. This scheme is called Bregman iterative regularization in the case if Υ is smoothness constraint. Osher *et al.* [78, 39] have discussed the convergence and stability of this scheme in detail.

Now, in the case of SR image reconstruction problem (3.1.8) under noise-free environment, we want to solve the constraint minimization problem of the form:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \Upsilon(\mathbf{X}), \text{ such that } RH\mathbf{X} = \underline{\mathbf{Y}} \tag{3.3.5}$$

where $H \in R^{N\times N}$ is a linear operator, $\underline{\mathbf{Y}} \in R^N$ is the observed sparse signal, $\mathbf{X} \in R^N$ is the signal we want to estimate, and Υ is a smoothness constraints (continuous convex function). Now in our case $T(\mathbf{X}) = \frac{1}{2}\|RH\mathbf{X} - \underline{\mathbf{Y}}\|_2^2$ for l_2 norm minimization, the corresponding Bregman equations are

$$\begin{cases} \mathbf{X}^{(n+1)} = \arg\min_{\mathbf{X}} \left[\mu(\Upsilon(\mathbf{X}) - \langle p^{(n)}, \mathbf{X}\rangle) + \frac{1}{2}\|RH\mathbf{X} - \underline{\mathbf{Y}}\|_2^2\right] \\ p^{(n+1)} = p^{(n)} - H^T R^T (RH\mathbf{X}^{(n+1)} - \underline{\mathbf{Y}}) \end{cases} \tag{3.3.6}$$

For any sequence $\{\mathbf{X}^{(n)}\}$, $\{p^{(n)}\}$ satisfying (3.3.6) for Υ continuous and convex, we have, for any $\bar{\mathbf{X}}$

$$\begin{aligned} B_\Upsilon^{p^{(n)}}(\bar{\mathbf{X}}, \mathbf{X}^{(n)}) - B_\Upsilon^{p^{(n-1)}}(\bar{\mathbf{X}}, \mathbf{X}^{(n-1)}) \leq \langle RH\bar{\mathbf{X}} - \underline{\mathbf{Y}}, RH\mathbf{X}^{(n-1)} - \underline{\mathbf{Y}}\rangle \\ - \|RH\mathbf{X}^{(n-1)} - \underline{\mathbf{Y}}\|_2^2 \end{aligned} \tag{3.3.7}$$

Equation (3.3.7) implies that the Bregman distance between $\mathbf{X}^{(n)}$ and any element $\bar{\mathbf{X}}$ satisfying $RH\bar{\mathbf{X}} = \underline{\mathbf{Y}}$ is monotonically decreasing. Second, (3.3.7) also implies that if $\bar{\mathbf{X}}$ is the "noise-free" approximation of the solution of (3.1.8), the Bregman distance between $\mathbf{X}^{(n)}$ and $\bar{\mathbf{X}}$ diminishes as long as

$$\|RH\mathbf{X}^{(n-1)} - \underline{\mathbf{Y}}\|_2^2 > \|RH\bar{\mathbf{X}} - \underline{\mathbf{Y}}\|_2^2 = \sigma \tag{3.3.8}$$

i.e. until we get too close to the noisy signal, where σ is some measure of noise and this gives the stopping criterion for SR reconstruction from noisy LR images.

Yin *et al.* [123] have shown that for the case of linear constraints $RH\mathbf{X} - \underline{\mathbf{Y}} = 0$ in (3.2.4), Bregman iterations (3.3.4) can be reduced to a more simplified form [39] with l_2 norm

$$\begin{cases} \mathbf{X}^{(n+1)} = \arg\min_{\mathbf{X}}\{\mu\Upsilon(\mathbf{X}) + \frac{1}{2}\|RH\mathbf{X} - \mathbf{Y}^{(n)}\|_2^2\} \\ \mathbf{Y}^{(n+1)} = \mathbf{Y}^{(n)} + (\underline{\mathbf{Y}} - RH\mathbf{X}^{(n+1)}) \end{cases} \tag{3.3.9}$$

In other words the error $(\underline{\mathbf{Y}} - RH\mathbf{X}^{(n)})$ in n-th estimation is added back to $\mathbf{Y}^{(n)}$ such that finally $RH\mathbf{X}^{(n)}$ satisfies converges to $\underline{\mathbf{Y}}$ and the constraint $RH\mathbf{X} - \underline{\mathbf{Y}} = 0$. Note that the first equation solve the unconstrained minimization problem

(3.1.10). This is analogue of 'adding back noise' in the ROF model for TV de-noising [78]. Therefore the scheme presented in (3.3.6) can be reduce into an equivalent simplified form of Bregman iteration as As, in general, there is no explicit expression for $\mathbf{X}^{(n+1)}$ to solve the unconstrained optimization subproblem (first equation) (3.3.9), we go further to solve it explicitly.

3.3.2 Proximal map

Consider the following unconstrained minimization problem:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}}(\mu\Upsilon(\mathbf{X}) + T(\mathbf{X})) \tag{3.3.10}$$

where $\mu > 0$. Combettes *et al.* [18] describe a forward-backward technique to minimize the sum of two convex functionals based on the proximal operator introduced by Moreau [70]. By classical arguments of convex analysis, the solution of (3.3.10) satisfies the condition:

$$\mu\partial\Upsilon(\mathbf{X}) + \partial T(\mathbf{X}) = 0$$

For any positive number γ, we have

$$(\mathbf{X} + \gamma\mu\partial\Upsilon(\mathbf{X})) - (\mathbf{X} - \gamma\partial T(\mathbf{X})) = 0$$

This leads to a forward and backward splitting algorithm:

$$\mathbf{X}^{(k+1)} = Prox_{\Upsilon}(\mathbf{X}^{(k)} - \gamma\partial T(\mathbf{X}^{(k)})), \tag{3.3.11}$$

where the proximal operator $Prox_{\Upsilon}(\mathbf{V})$ is defined as:

$$Prox_{\Upsilon}(\mathbf{V}) = \arg\min_{\mathbf{X}}\{\mu\Upsilon(\mathbf{X}) + \frac{1}{2\gamma}\|\mathbf{X} - \mathbf{V}\|_2^2\} \tag{3.3.12}$$

In our case $T(\mathbf{X}) = \frac{1}{2}\|RH\mathbf{X} - \underline{\mathbf{Y}}\|_2^2$. we can see that $\partial T(\mathbf{X}) = H^T R^T(RH\mathbf{X} - \underline{\mathbf{Y}})$ is β-Lipchitz, where $\beta = \lambda_{max}[(RH)^T RH]$ is the maximum eigenvalue of the square symmetric matrix $[(RH)^T RH]$. Therefore the solution of the minimization

problem (3.3.10) can be computed by the following two-step algorithm:

$$\begin{cases} \mathbf{U}^{(k+1)} = \mathbf{X}^{(k)} - \gamma H^T R^T (RH\mathbf{X}^{(k)} - \underline{\mathbf{Y}}) \\ \mathbf{X}^{(k+1)} = \arg\min_{\mathbf{X}} \{\mu \Upsilon(\mathbf{X}) + \frac{1}{2\gamma} \|\mathbf{X} - \mathbf{U}^{(k+1)}\|_2^2\} \end{cases} \tag{3.3.13}$$

This algorithm (3.3.13) can be used to solve the unconstrained minimization problem presented by first equation of (3.3.9) as described next.

3.3.3 Bregmanized operator splitting

In this section we develop an algorithm to solve the equality constrained minimization problem (3.3.1) with $T(\mathbf{X}) := RH\mathbf{X} - \underline{\mathbf{Y}}$ using Bregman iterations (3.3.9) and operator splitting (3.3.13), briefly introduced in the last two subsections. First of all Equality constraint in (3.1.8) is enforced with Bregman iteration as in Eq. (3.3.9) process as follows:

$$\begin{cases} \mathbf{X}^{(n+1)} = \arg\min_{\mathbf{X}} \left(\mu \Upsilon(\mathbf{X}) + \frac{1}{2} \|RH\mathbf{X} - \mathbf{Y}^{(n)}\|_2^2\right) \\ \mathbf{Y}^{(n+1)} = \mathbf{Y}^{(n)} + (\underline{\mathbf{Y}} - RH\mathbf{X}^{(n+1)}) \end{cases} \tag{3.3.14}$$

Now, the operator splitting technique is used to solve the unconstrained subproblem, i.e. first equation of (3.3.9) as follows:

$$\textbf{for } k \geq 0, \ \mathbf{X}^{(n+1,\ 0)} = \mathbf{X}^{(n)}$$
$$\begin{cases} U^{(n+1,\ k+1)} = \mathbf{X}^{(n,\ k)} - \gamma H^T R^T (RH\mathbf{X}^{(n,\ k)} - \mathbf{Y}^{(n,\ k)}) \\ \mathbf{X}^{(n+1,\ k+1)} = \arg\min_{\mathbf{X}} \{\mu \Upsilon(\mathbf{X}) + \frac{1}{2\gamma} \|\mathbf{X} - U^{(n+1,\ k+1)}\|_2^2\} \end{cases} \tag{3.3.15}$$

Ideally we need to iterate (3.3.15) infinite number of times [126] to obtain a convergent solution $\mathbf{X}^{(n+1)}$ for the said subproblem. Instead we propose to iterate them once for small amount of noise, which leads to the following algorithm:

$$\begin{cases} \mathbf{U}^{(n+1)} = \mathbf{X}^{(n)} - \gamma H^T R^T (RH\mathbf{X}^{(n)} - \mathbf{Y}^{n}) \\ \mathbf{X}^{(n+1)} = \arg\min_{\mathbf{X}} \{\mu \Upsilon(\mathbf{X}) + \frac{1}{2\gamma} \|\mathbf{X} - \mathbf{U}^{(n+1)}\|_2^2\} \\ \mathbf{Y}^{(n+1)} = \mathbf{Y}^{(n)} + (\underline{\mathbf{Y}} - RH\mathbf{X}^{(n+1)}) \end{cases} \tag{3.3.16}$$

For optimal solution of the second equation of (3.3.16), we have (replacing gradients by its corresponding subgradients)

$$\mu\frac{\delta\Upsilon(\mathbf{X})}{\delta\mathbf{X}}+\frac{1}{\gamma}(\mathbf{X}-\mathbf{U}^{(n+1)})=0$$

So the corresponding iteration can be written as

$$\mathbf{X}^{(n+1)}=\mathbf{U}^{(n+1)}-\mu^{'}\mid\frac{\delta\Upsilon(\mathbf{X})}{\delta\mathbf{X}}\mid_{\mathbf{X}^{(n)}}$$

where $\mu^{'}=\mu\gamma$ and $\mid\frac{\delta\Upsilon(\mathbf{X})}{\delta\mathbf{X}}\mid_{\mathbf{X}^{(n)}}$ is a subgradient of $\Upsilon(\mathbf{X})$ at $\mathbf{X}^{(n)}$.

Goldstein and Osher [39] have shown that the above algorithm (3.3.16) can also be used to approximate the minimizer for linear inequality constraint of the form $\|RH\mathbf{X}-\underline{\mathbf{Y}}\|_2^2<\eta$ for small value of η by imposing a stopping criterion based on inequality constraint that leads to an algorithm for solving inequality constrained minimization (3.1.8). Therefore we can present the proposed SR image reconstruction algorithm for small amount of noise using Bregman iteration and operator splitting as follows:

Proposed Iterative Algorithm for SR Image Reconstruction:

Initialize $\mathbf{Y}^{(0)}=n=0,\underline{\mathbf{Y}}$, $\mathbf{X}^{(0)}=\text{FillUnknown}(\underline{\mathbf{Y}})$;

$$\begin{aligned}&\textbf{While}(\|RH\mathbf{X}^{(n)}-\underline{\mathbf{Y}}\|_2^2>\eta)\\&\quad\begin{cases}\mathbf{U}^{(n+1)}=\mathbf{X}^{(n)}-\gamma H^TR^T(RH\mathbf{X}^{(n)}-\mathbf{Y}^n)\\\mathbf{X}^{(n+1)}=\mathbf{U}^{(n+1)}-\mu^{'}\mid\frac{\delta\Upsilon(\mathbf{X})}{\delta\mathbf{X}}\mid_{\mathbf{X}^{(n)}}\\\mathbf{Y}^{(n+1)}=\mathbf{Y}^{(n)}+(\underline{\mathbf{Y}}-RH\mathbf{X}^{(n+1)})\end{cases}\\&\quad n=n+1\\&\textbf{end}\end{aligned}\tag{3.3.17}$$

The module FillUnknown($\underline{\mathbf{Y}}$) in the initialization step of the above algorithm fills the unknown pixels in $\underline{\mathbf{Y}}$ by the corresponding known neighboring pixels. The parameter η is a predefined threshold chosen depending on variance of noise in LR images and significantly small for noise-free LR images. So we iterate until constraint in (3.1.8) is satisfied.

Note that, for LR images with high noise to satisfy inequality constraint in

(3.1.8), a fraction β (where $\beta < 1$) of the error $(\underline{\mathbf{Y}} - RH\mathbf{X}^{(n)})$ in n-th estimation is added back to $\mathbf{Y}^{(n)}$, i.e. last equation of (3.3.17) is replaced by $\mathbf{Y}^{(n+1)} = \mathbf{Y}^{(n)} + \beta(\underline{\mathbf{Y}} - RH\mathbf{X}^{(n+1)})$ and is iterated once after every l iterations of first two equation.

Since the proposed regularization function $\Upsilon(\mathbf{X})$ as defined in Eq. (3.2.3) consists of non-differentiable max and min operators $(D_s(\mathbf{X}), E_s(\mathbf{X}))$, we go for computing subgradients $|\ \frac{\delta\Upsilon(\mathbf{X})}{\delta\mathbf{X}}\ |_{\mathbf{X}^{(n)}}$ of the morphological regularization $\Upsilon(\mathbf{X})$ in the following subsection.

3.3.4 Subgradient methods

Now we try to develop the gradient of the dilated (3.2.1) and eroded (3.2.2) image with respect to its pixel values. As for example, let us consider two dimensional function $f(x_1, x_2) = \max\{x_1, x_2\}$, for this function partial derivatives exists at every points except the points along the line of intersection of the planes $g_1(x_1, x_2) = x_1$ and $g_2(x_1, x_2) = x_2$ as shown in the Fig. 3.3. i.e. $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$ exist at the points where one variable is strictly greater than the other one and it's values are as follows:

$$\frac{\partial f}{\partial x_1} = \begin{cases} 1 & \text{if } x_1 > x_2 \\ 0 & \text{if } x_1 < x_2 \\ \text{undefined}, & \text{elsewhere} \end{cases} \tag{3.3.18}$$

$$\frac{\partial f}{\partial x_2} = \begin{cases} 0 & \text{if } x_1 > x_2 \\ 1 & \text{if } x_1 < x_2 \\ \text{undefined}, & \text{elsewhere} \end{cases} \tag{3.3.19}$$

For this type of non-differentiable error function we use sub-gradient method [18, 19, 88, 97] for iterative convergence of the target vector. Now, $p \in \delta\Upsilon$ is said to be subgradient of a functional Υ at the point V if it satisfies the following

$$\Upsilon(X) - \Upsilon(V) - \langle X - V, p\rangle \geq 0$$

as defined in (3.3.3). Now we need to find out possible sub-gradients of the functional $f(x_1, x_2) = \max\{x_1, x_2\}$ at the points where it is not differentiable,

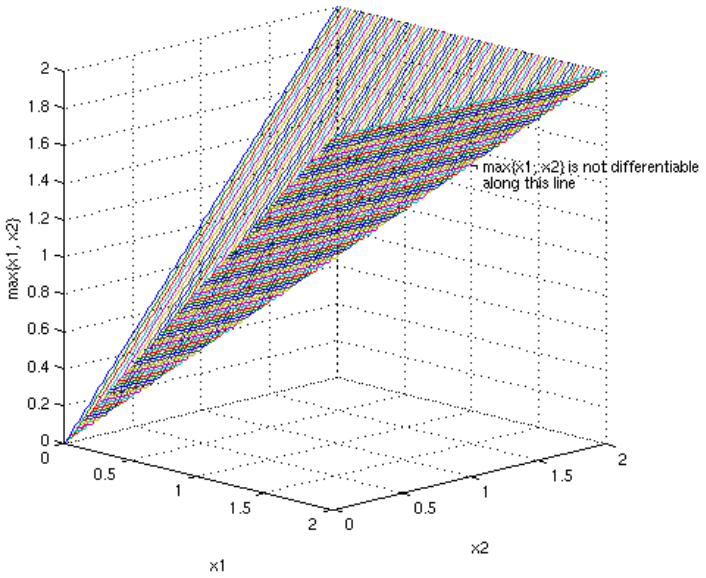


Figure 3.3: Shows the function $f(x_1, x_2) = \max\{x_1, x_2\}$ is differentiable at all the points and non-differentiable along the line of intersection of the planes $g_1(x_1, x_2) = x_1$ and $g_2(x_1, x_2) = x_2$

i.e. at the points along the line $x_1 = x_2$. Let $p = (p_1, p_2)$ be the sub-gradient of the functional f at the point $V = (x_1^{'}, x_2^{'})$ where $x_1^{'} = x_2^{'}$. At all other points the subgradients are unique valued and are equal to the gradients as (3.3.18) and (3.3.19). If p is sub-gradient at the point $v = (x_1^{'}, x_2^{'})$ then for any point $X = (x_1, x_2)$ near V it satisfy the condition of subgradient i.e.

$$\begin{aligned} f(X) - f(V) - \langle X - V, p\rangle &\geq 0 \\ \text{i.e. } \max\{x_1, x_2\} - \max\{x_1^{'}, x_2^{'}\} - \langle (x_1 - x_1^{'}, x_2 - x_2^{'}), p\rangle &\geq 0 \end{aligned}$$

While calculating p_1 along x_1, the other coordinate x_2 will remain unchanged. i.e $x_2 = x_2^{'}$. So $x_1 > x_2$ gives

$$x_1 - x_1^{'} - p_1(x_1 - x_1^{'}) \geq 0 \text{ i.e. } p_1 \leq 1$$

and similarly $x_1 < x_2$ gives

$$x_2 - x_2^{'} - p_1(x_1 - x_2^{'}) \geq 0 \text{ i.e. } p_1 \geq 0$$

Therefore, sub-gradient for the two-dimensional max operator along x_1 can be written as:

$$p_1 := \frac{\delta f}{\delta x_1} = \begin{cases} 1 & \text{if } x_1 > x_2 \\ 0 & \text{if } x_1 < x_2 \\ \in [0,1] & \text{elsewhere} \end{cases}$$

Similarly for subgradient of max operator along x_2 coordinate is as following:

$$p_2 := \frac{\delta f}{\delta x_2} = \begin{cases} 0 & \text{if } x_1 > x_2 \\ 1 & \text{if } x_1 < x_2 \\ \in [0,1] & \text{elsewhere} \end{cases}$$

3.3.5 Subgradients of morphologic regularization function

Here we derive the subgradients of the dilated (3.2.1) and eroded (3.2.2) image with respect to its pixel values. Let us denote the subgradient of a dilated image $D_s(X)$ by $\frac{\delta D_s}{\delta X}$. From the definition of $D_s(X)$, as given in (3.2.1), j^{th} element of i^{th} column of the subgradient $\frac{\delta D_s}{\delta X}$ is:

$$\frac{\delta D_{s,j}}{\delta x_i} = \begin{cases} 1 & \text{if } x_i = \max_{r \in (sB)_{(j)}}\{x_r\} \text{ and } \forall t \in (sB)_{(j)}, t \neq i, x_t < x_i \\ 0 & \text{if } x_i < \max_{r \in (sB)_{(j)}}\{x_r\} \\ \in [0,1] & \text{elsewhere} \end{cases} \tag{3.3.20}$$

Similarly subgradient of an eroded image $E_s(X)$ can be written as follows:

$$\frac{\delta E_{s,j}}{\delta x_i} = \begin{cases} 1 & \text{if } x_i = \min_{r \in (sB)_{(j)}}\{x_r\} \text{ and } \forall t \in (sB)_{(j)}, t \neq i, x_t > x_i \\ 0 & \text{if } x_i > \min_{r \in (sB)_{(j)}}\{x_r\} \\ \in [0,1] & \text{elsewhere} \end{cases} \tag{3.3.21}$$

Now in (3.3.20) and (3.3.21) if we choose subgradient equal to 1 out of the range $[0, 1]$. Then the subgradients become:

$$\frac{\delta D_{s,j}}{\delta x_i} = \begin{cases} 1 & \text{if } x_i = \max_{r\in(sB)_{(j)}}\{x_r\} \\ 0 & \text{if } x_i < \max_{r\in(sB)_{(j)}}\{x_r\} \end{cases} \tag{3.3.22}$$

$$\frac{\delta E_{s,j}}{\delta x_i} = \begin{cases} 1 & \text{if } x_i = \min_{r\in(sB)_{(j)}}\{x_r\} \\ 0 & \text{if } x_i > \min_{r\in(sB)_{(j)}}\{x_r\} \end{cases} \tag{3.3.23}$$

Note that s^{th} scale morphological closing and morphological opening are written in terms of dilation and erosion as:

$$C_s(\mathbf{X}) = E_s(D_s(\mathbf{X}))$$

$$O_s(\mathbf{X}) = D_s(E_s(\mathbf{X}))$$

Since analogous chain rule holds for sub-gradients, hence we can write down the sub-gradients of the regularization function $\Upsilon(\mathbf{X})$ (3.2.3) as

$$\begin{aligned} \tfrac{\delta\Upsilon(\mathbf{X})}{\delta\mathbf{X}} &:= \tfrac{\delta}{\delta\mathbf{X}}\textstyle\sum_{s=1}^{K}\alpha^s\mathbf{1}^t[C_s(\mathbf{X}) - O_s(\mathbf{X})] \\ &= \textstyle\sum_{s=1}^{K}\alpha^s[\tfrac{\delta C_s(\mathbf{X})}{\delta\mathbf{X}} - \tfrac{\delta O_s(\mathbf{X})}{\delta\mathbf{X}}]\mathbf{1} \\ &= \textstyle\sum_{s=1}^{K}\alpha^s[\tfrac{\delta}{\delta\mathbf{X}}E_s(D_s(\mathbf{X})) - \tfrac{\delta}{\delta\mathbf{X}}D_s(E_s(\mathbf{X}))]\mathbf{1} \\ &= \textstyle\sum_{s=1}^{K}\alpha^s[\tfrac{\delta}{\delta D_s(\mathbf{X})}E_s(D_s(\mathbf{X}))\tfrac{\delta}{\delta\mathbf{X}}D_s(\mathbf{X})) - \tfrac{\delta}{\delta E_s(\mathbf{X})}D_s(E_s(\mathbf{X}))\tfrac{\delta}{\delta\mathbf{X}}E_s(\mathbf{X})]\mathbf{1} \end{aligned} \tag{3.3.24}$$

Clearly, $\frac{\delta}{\delta D_s(\mathbf{X})}E_s(D_s(\mathbf{X}))$ and $\frac{\delta}{\delta E_s(\mathbf{X})}D_s(E_s(\mathbf{X}))$ can be calculated following (3.3.22) and (3.3.23) with respect to s^{th} scale dilation $D_s(\mathbf{X}) = [d_{s,1}, d_{s,2}, \ldots d_{s,mn}]$ and erosion $E_s(\mathbf{X}) = [e_{s,1}, e_{s,2}, \ldots e_{s,mn}]$ respectively as follows:

$$q_i^{E_{s,j}} := \frac{\delta E_{s,j}}{\delta d_{s,i}} = \begin{cases} 1 & \text{if } d_{s,i} = \min_{r\in(sB)_{(j)}}\{d_{s,r}\} \\ 0 & \text{if } d_{s,i} > \min_{r\in(sB)_{(j)}}\{d_{s,r}\} \end{cases} \tag{3.3.25}$$

$$q_i^{D_{s,j}} := \frac{\delta D_{s,j}}{\delta e_{s,i}} = \begin{cases} 1 & \text{if } e_{s,i} = \max_{r\in(sB)_{(j)}}\{e_{s,r}\} \\ 0 & \text{if } e_{s,i} < \max_{r\in(sB)_{(j)}}\{e_{s,r}\} \end{cases} \tag{3.3.26}$$

Substituting (3.3.23), (3.3.22), (3.3.25) and (3.3.26) in Eq. (3.3.24), we compute the subgradient $\frac{\delta \Upsilon(\mathbf{X})}{\delta \mathbf{X}}$ which is required in computing third step of (3.3.17). In Appendix A, we discuss an efficient method to compute subgradients of $D_s(\mathbf{X})$ and $E_s(\mathbf{X})$ using chain rule.

3.4 Experimental results

In this section we study and analyze the performance of the proposed method as well as that of some other existing SR reconstruction methods. In this discussion the proposed algorithm (3.3.17) is referred to as 'Breg + Morph', while the other methods are referred to as 'Grd + TV' [91], 'Grd + BTV' [33], 'Breg + TV' [66] and 'Grd + LABTV' [59] respectively. Here 'Grd' stands for iterative gradient descent technique and 'Breg' stands for Bregman iteration technique used for optimization. 'Morph', 'TV', 'BTV' and 'LABTV' stands for morphological, total variation, bi-lateral total variation and locally adaptive bi-lateral total variation regularization respectively. Note that the methods associated with 'Breg' solve constrained SR reconstruction formulation (3.1.8) whereas those with 'Grd' solve the unconstrained formulation (3.1.10).

Experimental setting

For performance evaluation a typical 512×512 gray-level HR image (e.g., Chart image) is chosen [see Fig. 3.4(a)]. We synthetically generate some LR images from this HR image and later reconstruct a HR image from these generated LR images only. Finally, we compute PSNR and SSIM as quantitative measure of quality of reconstructed HR image with respect to the original HR image. To generate the LR images we have considered only translational motion as sub-pixel shifts with known value. The blurring is chosen as 5×5 Gaussian smoothing kernel with scale parameter $\sigma = 2.5$ and the matrix H is formed accordingly. The down-sampling factor is chosen to be 5 and the matrix D is constructed to achieve this. We assume purely translational shifts by integer value and F_k is formed accordingly. There are 25 possible choices of integer shifts for resolution factor 5 to generate 25 LR images. In our experiment we have observed that only

10 LR images among 25 are sufficient to reconstruct HR image of acceptable quality with resolution factor 5. Hence, in our experiment we take 10 randomly chosen LR images and determine the index matrix R of Eq. (3.1.3) accordingly. The same setup is followed throughout this experimental section unless stated otherwise. Second, an interesting portion is marked on each resultant HR image and is displayed on the top-right/left corner after zooming it for careful study of the visual quality. The parameters for each algorithm are chosen to maximize the PSNR with respect to the ground truth. For fair evaluation purpose, we have applied each SR reconstruction algorithm on different sets of 10 LR images with different parameters and the best result of each algorithm is chosen as the output of the algorithm for this experiment. In our reconstruction algorithm (3.3.17), we have chosen the model parameters $\gamma = 1$ and $\mu^{'} = 0.5$ respectively and have got good results in most of the experiments. In our iterative algorithm (3.3.17), initial estimate $\mathbf{X}^{(0)}$ is taken as $\underline{\mathbf{Y}}$ with unknown pixels are filled by the corresponding nearest known pixels. For the experiments with noisy LR images after every 5 iterations of the first two steps of the algorithm (3.3.17) the third step is executed once and β is chosen as reciprocal of the noise variance. We have implemented the algorithms using MATLAB 7.6 and run on a regular Desktop with 3Gz Intel quad-core processor and 8GB of RAM.

In the first experiment, LR images contain small amount of noise ($\sigma = 2$) and the up-sampled-merged image obtained from 10 LR images is shown in Fig. 3.4(c). The results of different algorithms are shown in Fig. 3.4(d)-(h). Figures 3.4(d)-(f) present SR reconstructed images using gradient descent method for optimization with TV, BTV and LABTV regularizations respectively. Similarly Fig. 3.4(g)-(h) present the results using Bregman iteration method for optimization with TV and morphologic regularizations respectively. In each case algorithm is terminated if the residue is less than certain threshold (e.g. η in the algorithm (3.3.17)) or number of iterations is more than 100. In Fig. 3.5(a) we plot how different algorithms approach the terminating condition. Table 3.1 shows the number of iterations and corresponding numerical time comparison of the results shown in Fig. 3.4.

It is seen that the morphologic regularization yields SR reconstructed image of better quality compared to other regularization methods with less number of

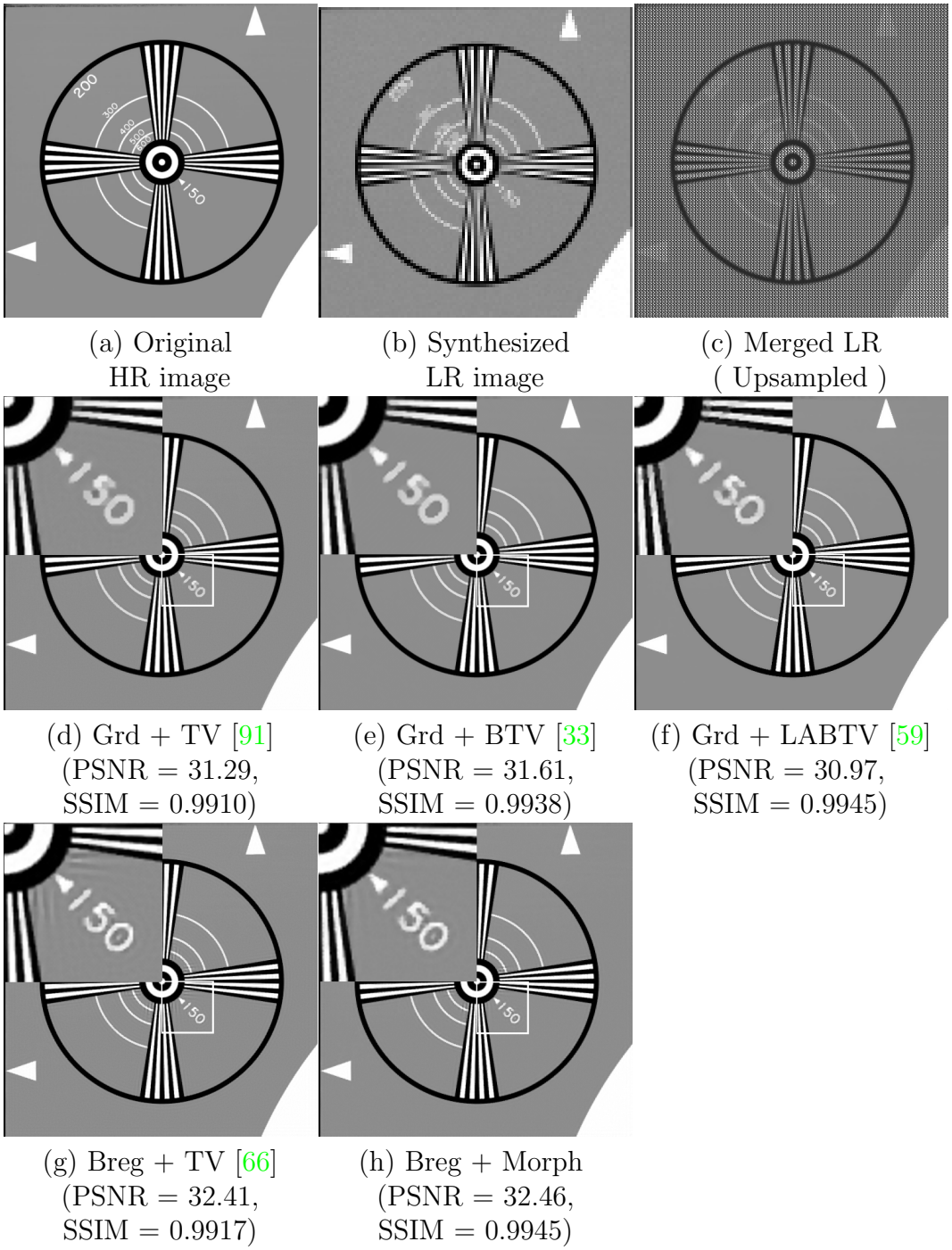
Table 3.1: Time comparison of different methods for the experiment in Fig. 3.4.

Method	Grd+TV	Grd+BTV	Grd+LABTV	Breg+TV	Breg+Morph
Iteration	221	211	308	68	63
Time	19.72s	32.47s	290.48s	6.83s	12.45s

iterations. In Section 3.3.1 it is mentioned that the convergence and stability of Bregman iteration based scheme are discussed in detail in [78, 39]. However, here we give an indication of the same based on experimental data. For this purpose in Fig. 3.5(b) we plot the objective function of 'Breg+TV' and 'Breg+Morph' upto the number of iterations as given in Table 3.1. The figure shows that after initial irregularities values of the objective function follows overall a decreasing in nature. Caption of Fig. 3.4 provides the quantitative measure of quality of the methods.

In Fig. 3.5(c), we plot improvement in PSNR of the estimated image $\mathbf{X}^{(n)}$ versus n for different SR reconstruction methods. These two plots, viz. Fig. 3.5(a) and Fig. 3.5(c), together show the reconstruction qualities of different methods with different number of iterations. Hence, with a chosen bound on residue as the stopping criterion for SR algorithms, the proposed method can achieve better quality with less number of iterations. Since different regularization method minimizes different functionals where regularization terms are different and data fidelity terms are same, this plot is useful to show the minimization of functionals. We observe that in addition to obtaining better quality image, the speed of convergence of iterative process is greatest in the proposed method. It is clearly seen that Bregman iteration based methods produce not only better results compared to gradient decent based method with all the regularization models using same or but also requires less number of iterations. It reveals two aspects: First, PSNR gradually approaches a stable value and second, performance of 'Breg+Morph' is best among the set. First observation implies that the Bregman iteration converges after sufficient iterations. And the second one suggests Bregman approach takes less iteration to achieve a certain quality. Thus, as expected, proposed method is superior to other competitive methods mentioned here.

We have applied the proposed SR reconstruction and also other methods over



(a) Original HR image

(b) Synthesized LR image

(c) Merged LR (Upsampled)

(d) Grd + TV [91] (PSNR = 31.29, SSIM = 0.9910)

(e) Grd + BTV [33] (PSNR = 31.61, SSIM = 0.9938)

(f) Grd + LABTV [59] (PSNR = 30.97, SSIM = 0.9945)

(g) Breg + TV [66] (PSNR = 32.41, SSIM = 0.9917)

(h) Breg + Morph (PSNR = 32.46, SSIM = 0.9945)

Figure 3.4: Illustrates results of various SR image reconstruction methods with small amount of noise ($\sigma = 2$): (a) Original HR image of a Chart, (b) One of the generated LR images, (c) Up-sampled and merged 10 LR images, (d)-(f) SR reconstructed image using gradient descent method with TV, BTV and LABTV regularization respectively, (g)-(h) SR reconstructed image using Bregman Iteration method with TV and morphologic regularization respectively. Number of iterations in each case are shown in Table 3.1.

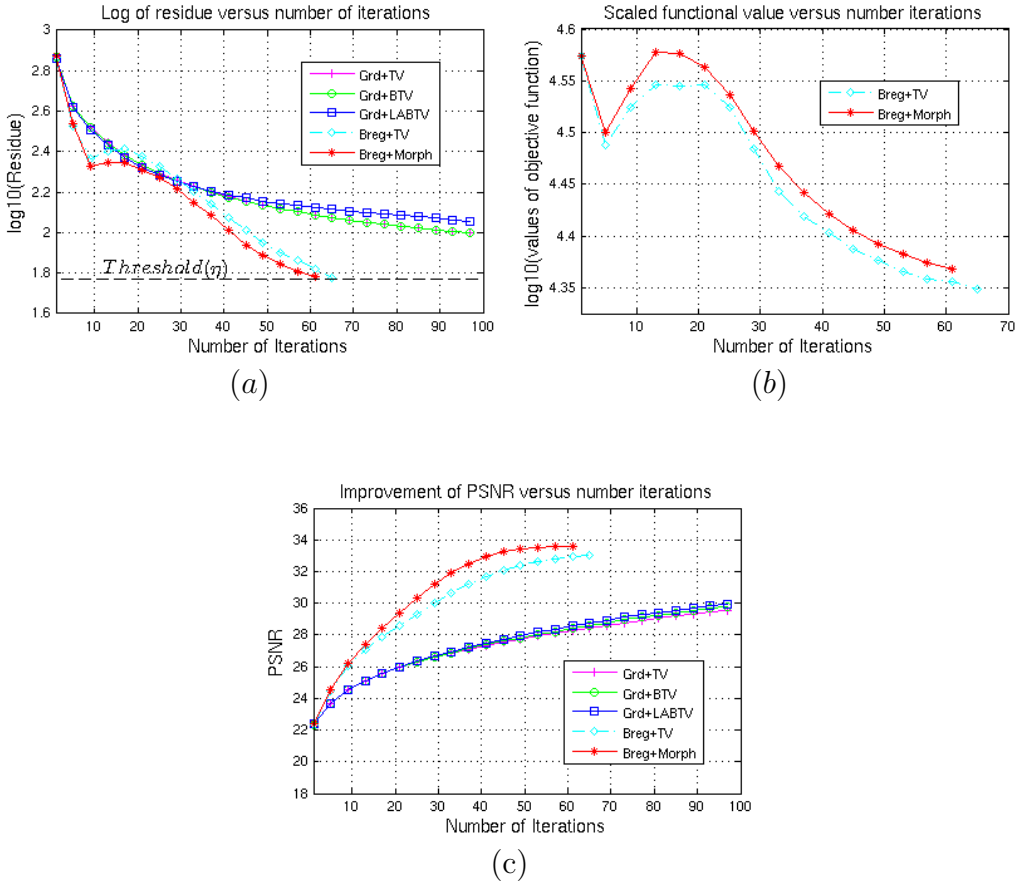


Figure 3.5: Comparison of reconstruction qualities of different methods versus number of iterations for the experiment in Fig. 3.4. (a) Illustrates how residue of data fidelity term approaches the threshold value to terminate the algorithm 3.3.1, (b) variation in objective function with iteration for 'Breg+TV' and 'Breg+Morph' [see text] and (c) PSNR up to 100 iterations for different algorithms as indicated in (a). Actual number of iterations and run time are shown in Table 3.1.

some more images and the results of only man, boat and Lena images are shown in Fig. 3.6. In each of the images the proposed method reconstructs more detail than other existing state-of-art reconstruction methods and we also achieve higher PSNR and SSIM in each case.

In the next experiment, we add high Gaussian noise with zero mean and standard deviation $\sigma = 12.0$ to the LR images. The results of different approaches are shown in Fig. 3.7. Here we see that proposed method is comparable to best performing methods 'Grd + BTV' [33] and 'Grd + LABTV' [59] in terms of both PSNR and SSIM.

In case of salt-and-pepper noise or impulse noise with random values, we can identify the position of noisy pixels by employing the concept of of Center-Weighted Median Filters (CWM) [53, 16] filters. So that such blur and noisy LR images may be handled by a two-phase algorithm. In the first phase, we find the locations of the pixels affected by noise, and in the second phase, those pixels are ignored for reconstruction of HR image. In the next experiment we consider LR images having 10% impulse noise with uniform distribution in the range $[-128, 128]$. Now CWM replaces the noisy pixel with the median value of the neighboring pixels and keeps the unaffected pixel as it is. The search for the noisy pixels using CWM filter depends on the neighborhood statistics of the pixels. Since the neighboring pixels of a pixel in a LR image does not remain neighboring ones when they are up-sampled and ensembled into an HR image $\underline{\mathbf{Y}}$ (see Fig. 3.1), we consider required number of closest known neighboring pixels in $\underline{\mathbf{Y}}$ of the candidate pixel. However, unlike [53, 16] in this work after detecting the noisy pixels we do not modify their values, rather we mark them as unknown pixels and modify the index matrix R accordingly, so that, those noisy pixels cannot effect the SR image reconstruction process. Figure 3.8 shows the results of the proposed method and other existing methods. Image shown in Fig. 3.8(f) present the result of proposed two-phase process.

In the next experiment we evaluate the performance of the proposed reconstruction method under miss-estimation of motion parameter and also the parameter of Gaussian blur. We corrupt the actual shifts of LR images by adding Gaussian distributed random values with zero mean and standard deviation 0.1 provided the changed values lie within the allowable range. For example, in case

(a)i Farsiu *et al.* [33] (PSNR = 29.80, SSIM = 0.9431)

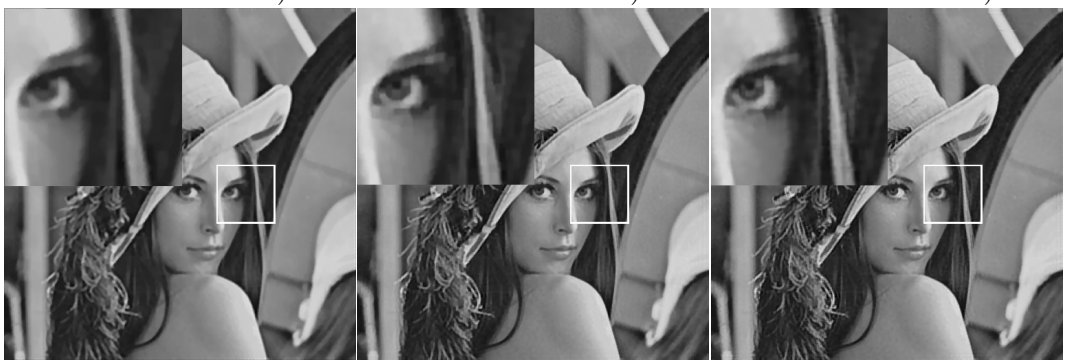
(a)ii Marquina *et al.* [66] (PSNR = 31.10, SSIM = 0.9656)

(a)iii Proposed method (PSNR = 31.45, SSIM = 0.9707)

(b)i Farsiu *et al.* [33] (PSNR = 29.86, SSIM = 0.9505)

(b)ii Marquina *et al.* [66] (PSNR = 31.60, SSIM = 0.9724)

(b)iii Proposed method (PSNR = 32.02, SSIM = 0.9755)

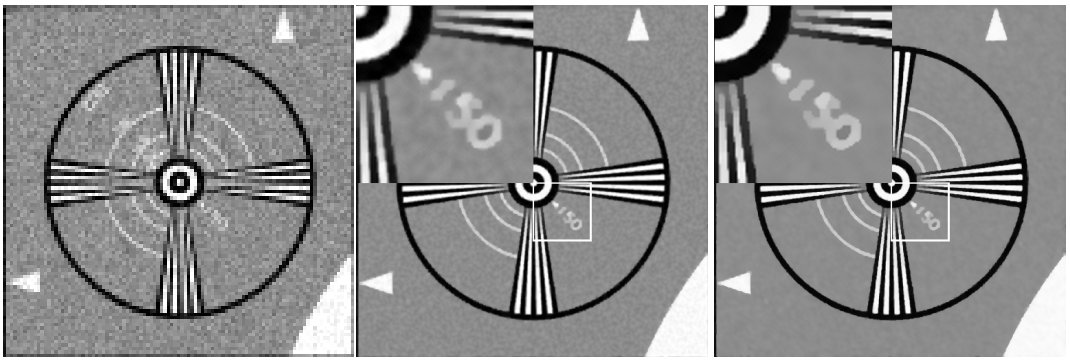


(c)i Farsiu *et al.* [33] (PSNR = 31.12, SSIM = 0.9562)

(c)ii Marquina *et al.* [66] (PSNR = 32.56, SSIM = 0.9714)

(c)iii Proposed method (PSNR = 32.81, SSIM = 0.9727)

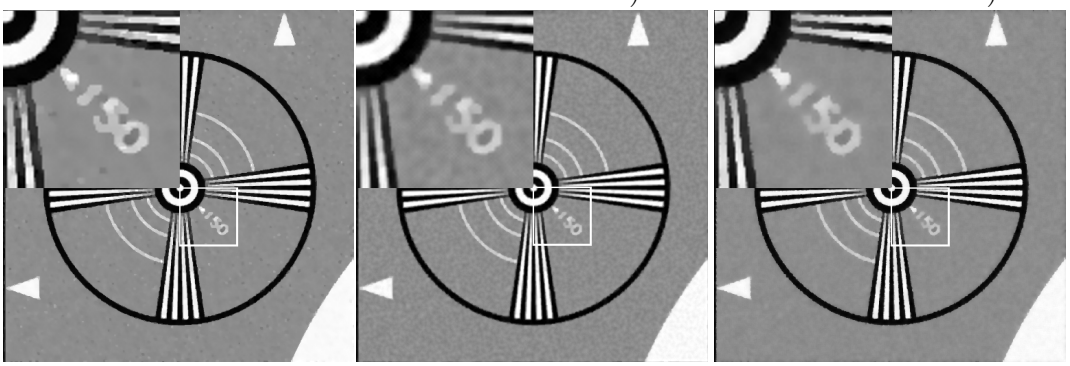
Figure 3.6: Comparison of reconstruction result of (a) man, (b) boat and (c) Lena images of proposed method over some existing methods: (i) SR image reconstructed using 'Grd + BTV' [33], (ii) SR image reconstructed using 'Breg + TV' [66], and (ii) SR image reconstructed using 'Breg + Morph' (Proposed method).



(a) LR image

(b) Grd + TV [91] (PSNR = 25.47, SSIM = 0.8786)

(c) Grd + BTV [33] (PSNR = 26.45, SSIM = 0.9523)

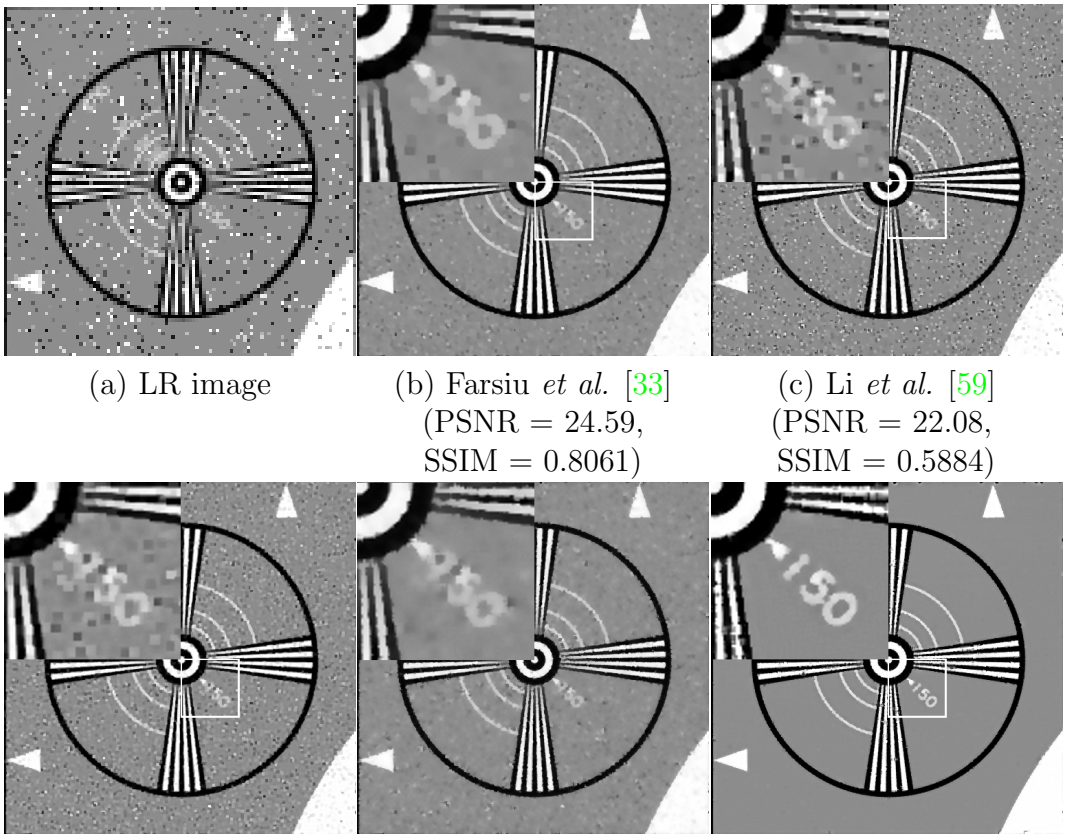


(d) Grd + LABTV [59] (PSNR = 26.17, SSIM = 0.9361)

(e) Breg + TV [66] (PSNR = 25.25, SSIM = 0.8745)

(f) Breg + Morph [prop.] (PSNR = 26.40, SSIM = 0.9384)

Figure 3.7: Illustrates results of various SR image reconstruction methods: (a) One of the generated LR images with high Gaussian noise (mean $\mu = 0$ and standard deviation $\sigma = 12$), (b)-(d) SR reconstructed image using gradient descent method with TV, BTV and LABTV regularization respectively, (e)-(f) SR reconstructed image using Bregman Iteration method with TV and morphologic regularization (proposed method) respectively.



(a) LR image

(b) Farsiu *et al.* [33] (PSNR = 24.59, SSIM = 0.8061)

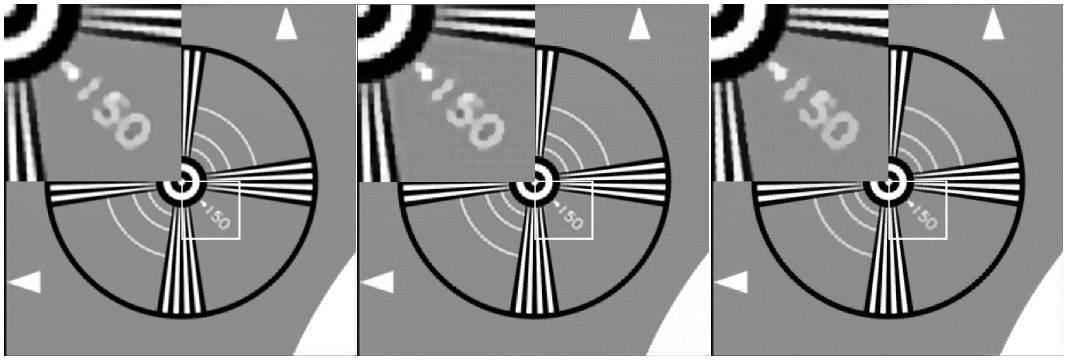
(c) Li *et al.* [59] (PSNR = 22.08, SSIM = 0.5884)

(d) Marquina *et al.* [66] (PSNR = 23.01, SSIM = 0.6052)

(e) Proposed Method (single phase) (PSNR = 24.10, SSIM = 0.8434)

(f) Proposed Method (two phase) (PSNR = 26.51, SSIM = 0.9827)

Figure 3.8: Illustrates results of various SR image reconstruction methods: (a) One of the generated LR images with impulse noise (Probability= 0.1) with uniform distribution in the range $\in$ [-128,128]), (b)-(d) SR image reconstructed using 'Grd + BTV' [33], 'Grd + LABTV' [59], 'Breg + TV' [66], and (e)-(f) reconstructed using 'Breg + Morph' (Proposed method) in single step and multi-step respectively.



(a) Farsiu *et al.* [33] (PSNR = 25.69, SSIM = 0.9827)

(b) Marquina *et al.* [66] (PSNR = 25.85, SSIM = 0.9646)

(c) Proposed method (PSNR = 26.30, SSIM = 0.9748)

Figure 3.9: Comparison of reconstruction result of Chart image for misestimated motion model and erroneous Gaussian blur parameter (a) SR image reconstruction using 'Grd + BTV' [33] (b) using 'Breg + BTV' [66] and (c) using 'Breg + Morph' (proposed method) respectively.

of resolution factor 5, the allowable range of shifts is [0, 5). If a noisy shift goes beyond this range it may be either truncated to lie within this range or a new random value is selected to corrupt the shift value. Here we adopt the later approach. However, this choice is not critical. Now the matrix F_k^{-1} is formed based on these noisy shifts and the corresponding matrix R of Eq. (3.1.3) would no longer be an index matrix rather an weight matrix. Moreover, though we generate the LR images using Gaussian blur with $\sigma = 2.5$, while reconstructing HR image we consider $\sigma = 2.35$. Figure 3.9 shows the results of this experiment. From the experiment it is seen that even under mispredicted motion parameter and scale of Gaussian blur the proposed SR reconstruction gives comparable result over the existing methods in terms of both PSNR and SSIM.

A more systematic study of performance of different algorithms for different amount of noise and different blurring parameter is conducted. In Fig. 3.10, we plot the average PSNR and SSIM of all the methods mentioned earlier applied on different images. In the experiment we have added different amount of Gaussian noise (standard deviation $\sigma = 0$ to 10) to LR images and applied various SR reconstruction algorithms. This is done on a set of images and then average

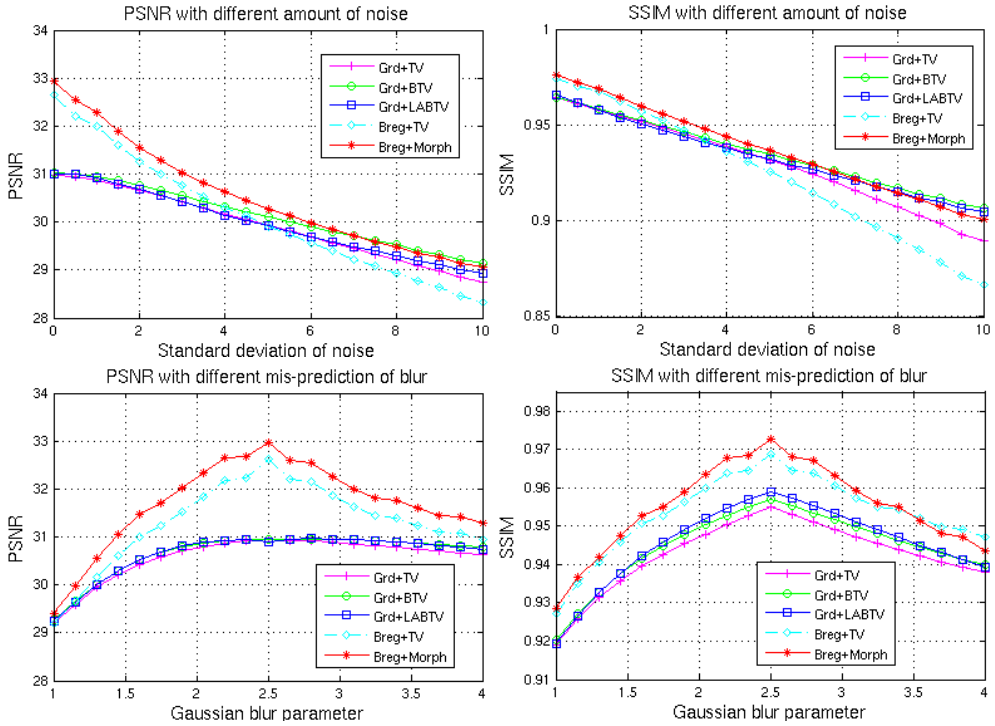


Figure 3.10: Analysis of the performance of SR image reconstruction algorithms applied on different gray images and then average PSNR and average SSIM are plotted. Top row : PSNR and SSIM of SR algorithms for noisy LR images with additive Gaussian noise. Bottom row : PSNR and SSIM for different amount of misprediction in blurring parameter.

PSNR and average SSIM are plotted. We observe that most of the cases proposed method is superior to the existing methods. Only exception is that with large amount of noise 'Grd + BTV' and 'Grd + LABTV' perform slightly better than the proposed method. Same experiment is done for varying blur parameter in reconstruction algorithms also. Here the blurring parameter in reconstruction model is varied from 1 to 4, while actual blurring parameter is 2.5. It is seen that the proposed method performs the best and the performance of all the methods is peaked at $\sigma = 2.5$ and falls off on either side.

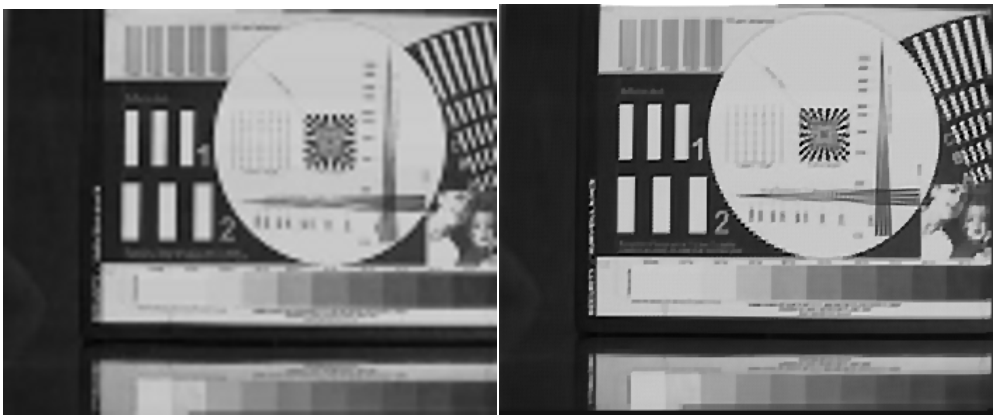
In the final experiment, we consider a sequence of 20 real compressed frames of size 288×352 (undergone translational motion approximately), called 'surveillance' frame, from a commercial video camera,(courtesy Farsiu and Robinson [68]).

We synthesize LR frames from HR ones by down-sampling it with resolution factor 5 followed by a blurring with 5×5 Gaussian kernel with blurring parameter $\sigma = 1.5$. Then we add Gaussian noise to each LR image with noise varience 5. We have used the method described by Vanderwalle *et al.* [114] to compute the sub-pixel shifts of LR frames. Then use SR reconstruction technique to Reconstruct those HR frames from LR ones. Figure 3.11(a) is reconstructed HR image using Bicubic interpolation and Fig. 3.11(b)-(f) are the reconstructed SR images with resolution factor 5 using existing method and proposed method in the same sequence as in Fig. 3.7. The first HR frame (taken as reference frame) is used for computing quantitative measures like PSNR and SSIM. However, visual comparison of the results in Fig. . 3.11 reveals that the artifacts occurred in Fig. 3.11(f), due to possible mis-prediction in motion vector as well as in parameter of Gaussian blur are comparable or less than that in Figs. 3.11(a)-(e).

3.5 Summary

In this chapter, we have presented an edge-preserving SR image reconstruction problem as de-blurring problem (3.1.3) with a new robust morphologic regularization method. Then we put forward two major contributions. First, we have proposed a morphologic regularization function based on multiscale opening and closing which can remove noise efficiently while preserving edge information. Secondly, we employ Bregman iteration method to solve the inverse problem for SR reconstruction with proposed morphologic regularization. It is well studied that multi-scale morphological filtering can reduce noise efficiently, so we have built up successfully a regularization method based on multiscale morphology and our experimental section shows that it works quite well, in fact better than existing methods. Non-linearity of the regularization function is handled in a linear fashion during optimization by means of sub-gradient and proximal map concept.

We also showed that if there are 'impulse noise with random values' or 'salt-and-pepper' noise in LR images, they can be handled efficiently using our two-step SR reconstruction algorithm. It first detects the noisy pixels (note: it does not substitute their values) and then considering those detected pixels as unknown



(a) Bicubic

(b) Rudin et al. [91]
($PSNR = 18.99$)
($SSIM = 0.6984$)

(c) Farsiu et al. [33]
($PSNR = 19.22$)
($SSIM = 0.6917$)

(d) Li et al. [59]
($PSNR = 19.08$)
($SSIM = 0.6915$)

(e) Marquina et al. [66]
($PSNR = 19.10$)
($SSIM = 0.6703$)

(f) Proposed Method
($PSNR = 19.12$)
($SSIM = 0.6970$)

Figure 3.11: Comparison of reconstruction result of real survillance frame sequence. (a) HR frame generated using Bicubic interpolation (c) SR frame reconstruction using TV regularization, (c) BTV regularization, (d) LABTV regularization, (e) TV regularization with Bregman iterations and (f) using proposed morphologic regularization method.

pixels, it reconstructs SR image using only those pixels which are not corrupted by noise.

The morphologic regularization method proposed here is tested only on SR reconstruction problem, but one can easily extend this regularization to other ill-posed problem as well. We also extend this regularization method to be adaptive by choosing SE of different shapes and sizes depending on the local statistics of neighboring pixels as discussed in the following chapter.

Chapter 4

Regularizations using Geodesic Transformation

In Section 1.3.1, we have discussed different types of state-of-the-art regularization methods. Most of those regularization term $\Upsilon(X)$ contains $\Gamma\mathbf{X}$ where Γ is an high-frequency operator. Moreover, most of the existing regularization technique uses fixed regularization parameter λ for all the smooth, edge and texture pixels. In Chapter 2, we proposed an adaptive regularization 2.2 based on morphologic gain map, which is able to tune the regularization parameter based on the structure at pixel location.

In Chapter 3, we have developed a regularization function (3.2.3) based on multi-scale morphology as following:

$$\Upsilon(X) = \sum_{s=1}^{S} \alpha^s \mathbf{1}^t [C_s(X) - O_s(X)] \tag{4.0.1}$$

where $\mathbf{1}$ is a column vector consisting of all 1's and α is the weighting coefficient. $C_s(X) = (X \oplus sB) \ominus sB$ and $O_s(X) = (X \ominus sB) \oplus sB$ are morphological closing and opening operators with disk structuring element sB of size s. Clearly, the regularization operators (1.9), (1.10) and (4.0.1) are based on some high-pass operators. An image, in general, may contain both detail and noise at different scales. Since both the noise and the edge contribute to high frequency components, they are usually removed simultaneously during the regularization process

enforcing smoothness. Ideal regularizer is expected to capture only noise present in the image so that minimizing $\Upsilon(X)$ leads to suppression of noise while keeping sharp edges unaltered. Among above regularizers, the morphologic regularizer (4.0.1) is the closest to the ideal, as the open and close filters remove noise smaller than SE and preserve edges upto the curvature of SE.

In this chapter, we develop a relationship between a non-linear low pass-filter and the existing regularization techniques. This leads to an efficient regularization technique based on the kernel defined in terms of geodesic distance. Here we also propose an adaptive morphological regularization, where the structuring element of the morphological operators varies from pixel to pixel, based on a geodesic kernel which performs more efficiently than non-adaptive regularization terms. In the next section we derive a relationship between smoothing filter and regularization operators, and show how an edge preserving smoother could lead to a good regularizer.

4.1 Regularizer and smoothing kernel

Smoothing a data-set mainly creates an approximating function that captures the underlying pattern in the data, while leaving out noise or other fine-scale redundant structures. Traditionally, people use symmetric linear filter (e.g., Gaussian filtering) or nonlinear filter (e.g., median filter). The regularization techniques described in the previous section may be viewed as a *prior* which suggests that the image is composed mostly of smooth regions. That means the reconstructed image would be mostly unaffected by low pass filtering with a constraint that it satisfies the forward image generating process (data error term of (1.7)). This constraint takes care of generating high frequency components due to edges whereas the smoothing term would prevent the inclusion of ringing artifacts. We usually decrease the emphasis on the regularizing term gradually as iteration proceeds, so that the image reconstruction process would end up with actual high frequency components, i.e., the edges in the image.

Let $K = (\alpha_{lm}|l, m = -w : w)$ be the kernel of linear/nonlinear low pass spatial filter. For average filter $\alpha_{lm} = 1/S$, and for Gaussian filter $\alpha_{lm} = \frac{1}{S}\exp^{-\frac{l^2+m^2}{2\sigma^2}}$, where S is the normalizing constant so that $\alpha_{lm} \geq 0$ and $\sum_{l,m} \alpha_{lm} = 1$. Suppose

W denotes an index set representing an window around a pixel. As discussed in the last section, an ideal image prior would be invariant to low pass filtered image KX that can remove noise and keep all the image details. Then the regularization term may be defined as

$$\begin{aligned}
\Upsilon(X) &= \|X - KX\|_p && (4.1.1)\\
&= \|X - \sum_{l,m\in W} \alpha_{lm} S_x^l S_y^m X\|_p && (4.1.2)\\
&= \|\sum_{l,m\in W} \alpha_{lm}(X - S_x^l S_y^m X)\|_p \text{ (since } \sum_{l,m\in W} \alpha_{lm} = 1) && (4.1.3)\\
&\leq \sum_{l,m\in W} \alpha_{lm}\|X - S_x^l S_y^m X\|_p \text{ (Jensen's Inequality), for } p > 1 && (4.1.4)
\end{aligned}$$

under minimization of $\Upsilon(X)$ [note that (1.7)], (4.1.4) is similar to (1.10). For average filtering ($\alpha_{lm} = 1/S$) with W as a 4-neighborhood, $\Upsilon(X)$ becomes

$$\Upsilon(X) = \sum_{l,m\in W} \alpha_{lm}\|X - S_x^l S_y^m X\|_p = 2\|(\nabla_x + \nabla_y)X\|_p = \|\nabla X\|_p \qquad (4.1.5)$$

which is similar to (1.9) for $p = 1$. An advantage of considering $\|\nabla X\|_1$ or $\sum_{l,m\in W} \alpha_{lm}\|X - S_x^l S_y^m X\|_1$ as regularization term is that it essentially minimizes $\|X - KX\|_p$ which is our smoothness prior. Now minimization of Eqs. (1.9), (1.10) and (4.0.1) tries to generate an image which is invariant to smoothing operators. The data term of (1.7) is responsible to generate high frequency components of the image and if it does not present then ideally regularization would produce a smooth flat image. Now, if $\Upsilon(X) = \|(\nabla_x + \nabla_y)X\|_p$, we end up with well-known BV (Bounded variation) for $p = 2$ and TV (Total variation) regularization for $p = 1$. Moreover, if we consider a larger window, $\Upsilon(X) = \sum_{l,m\in W} \alpha_{lm}\|X - S_x^l S_y^m X\|_p$, becomes BTV regularization for $p = 1$ and $\alpha_{lm} = \beta^{|m|+|l|}$ with $\beta \in (0, 1)$. Note that $\Upsilon(X)$ basically measures energy in high-frequency components. This suggests that, an attractive regularizer could be developed using spatial low pass filter.

Spatial image filtering techniques:

Most of the traditional filters use a symmetric non-adaptive kernel. Those filters are able to remove noise and artifacts present in the image, but also blur the edges and wipe out textures. Tomasi *et al.* [112] used adaptive kernel for edge preserving image smoothing, known as bilateral filtering, given by

$$\alpha_{ij,\,lm} \quad == \quad \exp^{-\{\frac{l^2+m^2+\lambda(X(i+l,j+m)-X(i,j))^2}{2\sigma_e^2}\}} = \exp^{-\{\frac{D^2_{ij,\,lm}}{2\sigma_e^2}\}} \tag{4.1.6}$$

where σ_e is the scale parameter and $X(i,j)$ is the intensity value at $(i,j)^{th}$ pixel and λ controls the emphasis between spatial and spectral differences. This kernel may also be viewed as the Gaussian of Euclidean distance $D_{ij,\,lm}$ between value of the pixels [Fig. 4.1]. It is still one of the state-of-the-art linear methods for edge

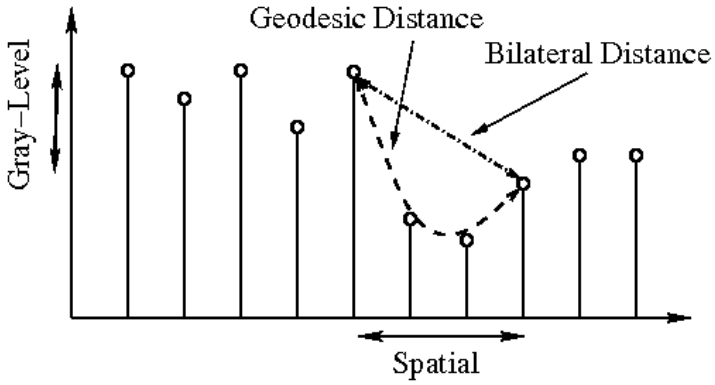


Figure 4.1: Illustrates the bilateral and geodesic distances between points.

preserving smoothing. Various nonlinear edge preserving morphological smoothing algorithms are also available in the literature [57, 17, 101, 23]. Lerallut *et al.* [57] define adaptive morphological operators, called amoeba filter, where shape of the structuring element varies from pixel to pixel, determined by thresholding the geodesic distance $D_{ij,\,lm}$ of the pixel $(i+l, j+m)$ from the center pixel (i,j) [Fig. 4.1]. Therefore, it is clear that regularization methods used in the inverse problems (4.1.4) evolve from low-pass filtering. Hence, a better regularizer can be developed from more efficient adaptive nonlinear low-pass filters. Based on these observations we develop two geodesic distance based regularization methods for

the SR image reconstruction problem.

4.2 Proposed morphologic regularization

We define general adaptive regularization as

$$\Upsilon(X) = \sum_{l,m \in W} \alpha_{lm} |\nabla^{lm}| \tag{4.2.1}$$

where α_{lm} is an adaptive kernel that varies from pixel to pixel. $|\nabla^{lm}| = \|X - S_x^l S_y^m X\|_p$ is the p^{th} norm of difference between the original image X and shifted one $S_x^l S_y^m X$ or $|\nabla^{lm}| = [C_s(X) - O_s(X)]$ is the L_1 norm of difference between morphologic closing and opening. In essence, a regularization term has two components: a derivative operator ∇^{lm} and a smoothing kernel α_{lm}. The elements of smoothing kernel α_{lm} usually decreases with distance from the candidate pixel (i, j). Morphologically viewed, an image may be considered as a topographic surface with (i, j) as spatial location and $X(i, j)$ as altitude. The geodesic distance in an image between pixels (i, j) and $(i + l, j + m)$ is defined as the length of the shortest path between them along the topographic surface [Fig. 4.1]. Note that BTV considers city-block distance (1.10) between pixel locations while bilateral filtering (4.1.6) considers Euclidean distance between two points on the topographic surface as discussed in the Section 4.1.

4.2.1 Proposed geodesic regularization

Based on the above study on the kernels, we propose an efficient adaptive kernel as follows:

$$\alpha_{ij,\,lm} = \frac{1}{S_{ij}} \exp\{-\frac{D^p_{ij,\,lm}}{2\omega^p}\} \tag{4.2.2}$$

where $D^p_{ij,\,lm}$ is the geodesic distance of the pixels (i, j) and $(i + l, j + m)$ within W. S_{ij} is the normalizing factor of the kernel over W at location (i, j). Based on the above geodesic kernel $\alpha_{ij,\,lm}$, We define a geodesic regularizer as follows:

$$\Upsilon(X) = \sum_{l,m \in W} \alpha_{lm} |X - S_x^l S_y^m X| \tag{4.2.3}$$

4.2.2 Proposed morphologic regularization

Moreover, we use adaptive morphological operators (4.0.1) by choosing adaptive structuring element A instead of a flat one and is defined as

$$a_{ij,\,lm} = c(\alpha_{ij,\,lm} - 1), \tag{4.2.4}$$

where c is a constant for the image. Accordingly, the adaptive morphological dilation is defined as:

$$(X \oplus A)(i,j) = \sup_{lm \in W} \{X(i-l, j-m) + a_{ij,\,lm}\}. \tag{4.2.5}$$

and adaptive morphological erosion is defined as:

$$(X \ominus A)(i,j) = \inf_{lm \in W} \{X(i+l, j+m) - a_{ij,\,lm}\} \tag{4.2.6}$$

Based on the above adaptive morphological operators, we define an adaptive morphologic regularizer as:

$$\Upsilon(X) = \sum_{l,m \in W} \alpha_{lm} |C_a(X) - O_a(X)| \tag{4.2.7}$$

where $C_a(X)$ and $O_a(X)$ are adaptive closing and opening operators.

4.2.3 Interpretation

The pixel-wise value due to different regularization terms (proposed as well as state-of-the-arts) for a noisy image are displayed in Fig. 4.2. Final regularization term is nothing but the aggregate of their magnitude, in short, energy. Image restoration algorithms try to minimize this energy along with the data error term (1.7). In TV and BTV regularization, edge information is quite significant. Thus minimization of these two regularizer leads to weakening or blurring the edges. It is clear from the image that proposed geodesic and adaptive morphologic regularization term extract more of noise than edges of the image. So in these two regularization terms noise energy is far more dominant than edge energy. In these cases, minimization of regularizers leads to suppression of noise leaving edge in-

formation more or less intact. This characteristics are also reflected in the quality of the restored image as will be seen in the experimental results. Finally, implementation of such regularizers require efficient computation of geodesic distance.

Figure 4.2: Illustrates different regularization methods. Top row: noisy input image with salt and paper noise (2%), TV regularization term (1.9) and BTV regularization (1.10). Bottom row: geodesic regularization (4.2.3), morphologic regularization (4.0.1), and adaptive morphologic regularization (4.2.7) terms. (Quantitative comparisons are in text).

4.3 Geodesic distance computation

Both of the proposed regularizers are based on the adaptive kernel $a_{ij,\,lm}$ as in (4.2.2) which is defined in terms of geodesic distance $D_{ij,\,lm}$ between the pixels (i, j) and $(i+l, j+m)$. There are two main kinds of algorithms that exist in the literature for computing Geodesic distance: raster-scan and wave-front propaga-

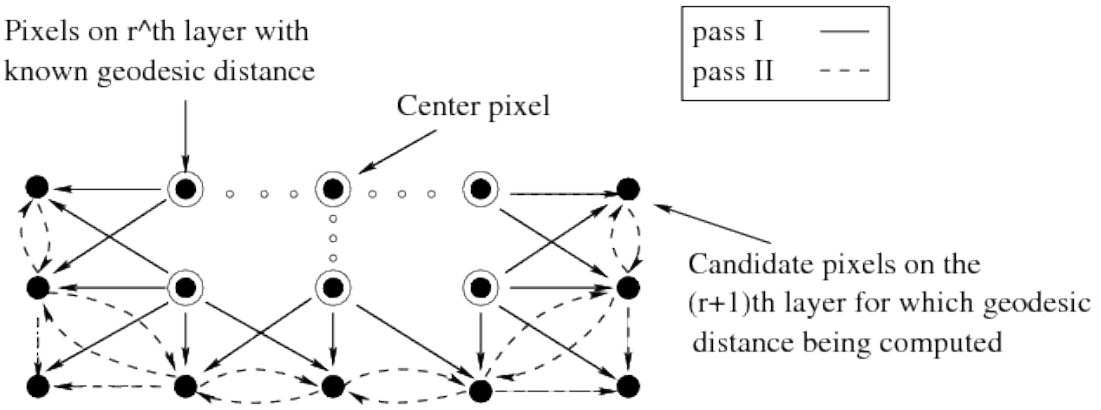


Figure 4.3: Flow chart of computation of geodesic distance of a pixel to its neighboring pixels

tion. Raster-scan algorithms are based on kernel operations applied sequentially over the image in multiple passes [2]. Those methods are the extensions of Dijkstra algorithm for two dimensional grids. Wave-front algorithms such as Fast Marching Method (FMM) [96], on the other hand, are based on the iterative propagation of a pixel front with certain velocity. Geodesic versions of both kinds of algorithms may be found in [47] and [122].

Most of the existing geodesic distance computation techniques are developed for generating a distance transformation map for a gray-level image over a binary image. Whereas, in the present work, we need to compute the geodesic distance of every pixel within W from the central pixel. So here, we develop an efficient two-pass iterative raster-scan algorithm which is an extension of Dijkstra algorithm. Let $\Pi(ij, lm) = \{(i, j) = x_1, x_2, \ldots, (i+l, j+m) = x_n\}$ be a path in a connected domain between the pixels x_k and x_{k+1}. Then the geodesic distance between them is defined as:

$$D_{ij,\,lm} = \min_{\Pi} \left[\sum_{k=1}^{n-1} \{ s_k + \gamma |X(x_k) - X(x_{k+1})| \} \right] \tag{4.3.1}$$

The factor γ weights the contribution of the intensity difference against the spatial distances, and s_k is spatial step size from x_k to x_{k+1}. Equation (4.3.1) generalizes the conventional spatial distance which may be obtained by setting $\gamma = 0$.

We define derivative operator

$$\nabla_{ij}^{lm} = |X(i,j) - X(i+l, j+m)|_1, \text{ where } l, m \in \{-1, 0, 1\} \tag{4.3.2}$$

Then $\gamma\nabla_{ij}^{lm}+s$ is the geodesic distance of the $(i,j)^{th}$ pixel from $(l,m)^{th}$ neighboring pixel, where s represents the cost of unit step in spatial direction. We take $s = 1$ for horizontal and vertical neighboring pixels and $s = \sqrt{2}$ for diagonal ones. The parameter γ controls the emphasis between spatial and spectral distances. Let Geo_{lm} represents geodesic distance from current $(i,j)^{th}$ pixel to $(i+l, j+m)^{th}$ pixel. We describe the computation of geodesic distance in concise algorithmic form as follows. Propagation of geodesic distance information is also illustrated in Fig. 4.3.

Algorithm

For central pixel (i, j), initialize $r = 0$, $Geo_{ij} = 0$
While r < w iterate Pass I and Pass II.

- **Pass I** : Compute interim geodesic distances of the pixel (l,m) on the $(r+1)^{th}$ layer from the central pixel along the possible shortest paths through the inner r^{th} layer.
  - Let $N_{lm}^{(r)}$ be the neighborhood of (l,m) consisting of pixels on the r^{th} layer as shown in Fig. 4.3, and consider it as a flat adaptive structuring element at (l,m).
  - Compute interim geodesic distance Geo_{lm}^{tmp} of pixel (l,m) from the center of the window by eroding with $N_{lm}^{(r)}$ as

$$Geo_{lm}^{tmp} = \min_{(u,v)\in N_{lm}^{r}} \{Geo_{uv} + \gamma\nabla_{lm}^{uv} + s\}$$

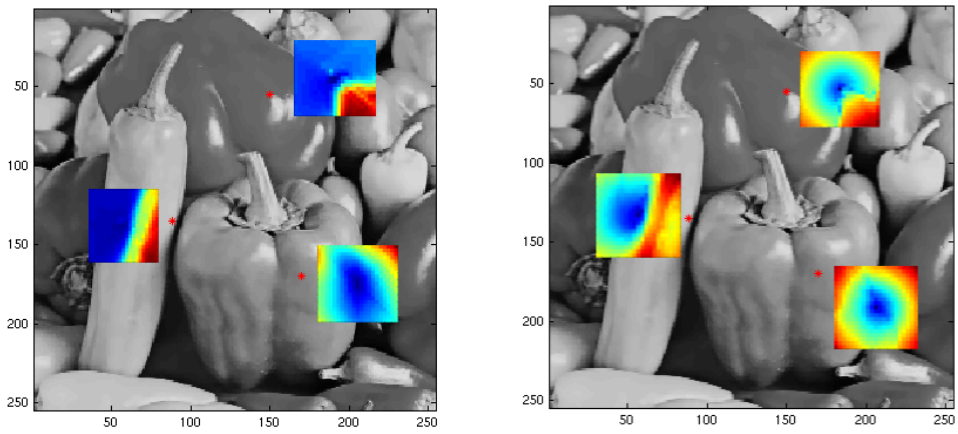
- **Pass II** : Compute actual geodesic distance of the pixel (l,m) on the $(r+1)^{th}$ layer using interim geodesic distances from Pass I.
  - Let $N_{lm}^{(r+1)}$ be the neighborhood of (l,m) pixel consist of pixels in the

$(r+1)^{th}$ layer as shown in Fig. 4.3 and consider it as a flat adaptive SE at (l, m).

– Compute actual geodesic distance of pixel (l, m) from the center of the window by eroding with N_{lm}^{r+1} as

$$Geo_{lm} = \min\left[Geo_{lm}^{tmp}, \min_{(u,v)\in N_{lm}^{(r+1)}}\{Geo_{uv}^{tmp} + \gamma\nabla_{lm}^{uv} + s\}\right]$$

Now, we can compute kernel matrix $K = [\alpha_{ij,\,lm}]$ as in (4.2.2) with $D_{ij,\,lm} = Geo_{lm}$. In Fig. 4.4, we have shown the computed kernels over different regions of an image. It is clear from the figure that the kernel is less sensitive to noise and more sensitive to edges.



(a) *Geodesic distance* (b) *Bilateral distance*

Figure 4.4: kernels based on geodesic and bilateral distances are shown at different points.

To make our geodesic kernel invariant to noise, we use patch-based derivatives ∇_{ij}^{lm}, instead of pixel-based derivative defined in (4.3.2). We convolute original image with the weight matrices shown in Fig. 4.5 to get the said patch-based derivatives.

We solve (1.7) with the proposed adaptive regularization (4.2.1) by Bregman iterations discussed in the last chapter. The non-differentiability and nonlinearity

$$\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & \frac{1}{6} \\ -\frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{bmatrix}$$

Figure 4.5: Different weight matrices for patch-based derivatives

are handled by the same subgradient operators used in the Section 3.3.

4.4 Experimental results and discussion

In this section we study the performance of the proposed regularization along with some state-of-the-arts regularization methods for multi-frame Super Resolution image reconstruction.

Experimental setting

For quantitative evaluation purpose, we have applied each algorithm on different sets of images with different parameters and the best result of each algorithm is chosen as the output of the algorithm. To compute geodesic kernel as described in (4.2.2), we have chosen $\omega = 5$ and $\gamma = 0.5$. For adaptive structuring element computation, we have chosen $c = 50$ for gray-level range [0-255]. We choose 512×512 gray-scale Lena image as test image with Gaussian blurring kernel ($\sigma_b = 2.5$) plus the Gaussian noise ($\sigma_a = 5$).

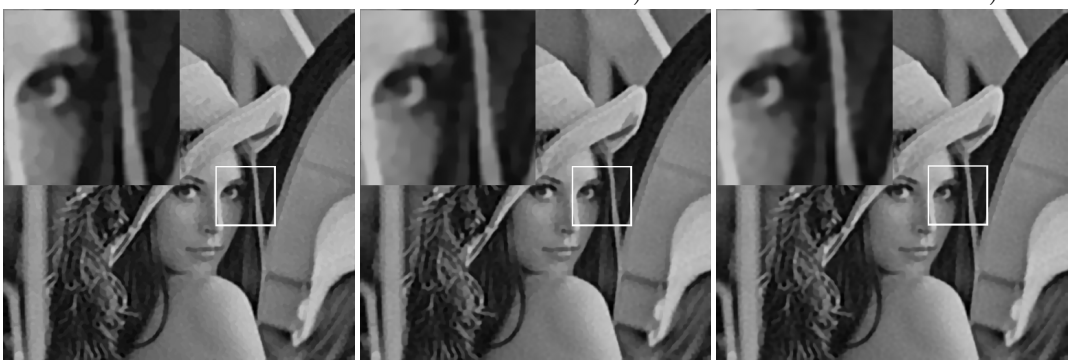
We synthesize 10 LR images down-sampled by resolution factor 5, and having different sub-pixel shifts with high Gaussian noise (mean $\mu = 0$ and standard deviation $\sigma = 5$). Then reconstruct HR image from these 10 LR images using the proposed regularization method as well as some existing regularization methods. Experimental results on Lena image are shown in Fig. 4.6. It is seen that the proposed morphological regularization yields better quality SR reconstructed images, though not significantly very high, compared to other regularization methods. Experimental result on Fig. 4.6 also provides the quantitative measure of quality of proposed geodesic regularization (4.2.3) (PSNR = 29.48dB, SSIM = 0.9082), adaptive morphologic regularization (4.2.7) method (PSNR = 29.59dB, SSIM = 0.9107), Rudin *et al.* [91] (PSNR = 29.43dB, SSIM = 0.9077), Farsiu

et al. [33] (PSNR = 29.47dB, SSIM = 0.9080) and morphological regularization (PSNR = 29.55dB, SSIM = 0.9106). They all perform equally well. However, though marginally, the proposed adaptive morphologic regularization methods perform better than others.

(a) LR Image

(b) TV (PSNR = 29.43, SSIM = 0.9077)

(c) BTV (PSNR = 29.47, SSIM = 0.9080)



(d) Geo (PSNR = 29.48, SSIM = 0.9082)

(e) Morph (PSNR = 29.55, SSIM = 0.9106)

(f) Amorph (PSNR = 29.59, SSIM = 0.9107)

Figure 4.6: Illustrates results of various SR image reconstruction techniques on Lena image. Top row : one of LR images, SR image using TV regularization [91] and BTV regularization [33]. Bottom row : SR reconstructed image using geodesic regularization, morphologic regularization and adaptive morphologic regularization method.

We have applied the proposed SR reconstruction and also other methods over some more images and the results of only man and boat images are shown in Figures 4.7 and 4.8 respectively. In each of the images the proposed method

reconstructs more detail than other existing state-of-art reconstruction methods and we also achieve higher PSNR and SSIM in each case.

(a) LR Image

(b) TV (PSNR = 27.39, SSIM = 0.8671)

(c) BTV (PSNR = 27.43, SSIM = 0.8678)

(d) Geo (PSNR = 27.43, SSIM = 0.8679)

(e) Morph (PSNR = 27.49, SSIM = 0.8698)

(f) Amorph (PSNR = 27.52, SSIM = 0.8698)

Figure 4.7: Illustrates results of various SR image reconstruction techniques on man image. Top row : one of LR images, SR image using TV regularization [91] and BTV regularization [33]. Bottom row : SR reconstructed image using geodesic regularization, morphologic regularization and adaptive morphologic regularization method.

4.5 Summary

In this chapter, we have proposed two geodesic kernel based regularization for multi-frame Super Resolution image reconstruction. We analyze different reg-

(a) LR Image

(b) TV (PSNR = 27.10, SSIM = 0.8794)

(c) BTV (PSNR = 27.12, SSIM = 0.8796)

(d) Geo (PSNR = 27.14, SSIM = 0.8789)

(e) Morph (PSNR = 27.17, SSIM = 0.8795)

(f) Amorph (PSNR = 27.19, SSIM = 0.8809)

Figure 4.8: Illustrates results of various SR image reconstruction techniques on boat image. Top row : one of LR images, SR image using TV regularization [91] and BTV regularization [33]. Bottom row : SR reconstructed image using geodesic regularization, morphologic regularization and adaptive morphologic regularization method.

ularization methods used in this domain and found that they come from the same concept of edge preserving smoother. We propose a new robust adaptive geodesic regularization and adaptive morphologic regularization methods that can suppress the noise more efficiently while preserving the edges. Our experimental section shows that it works quite well, in fact better than existing techniques.

Chapter 5

Embedding Natural Image Prior for Single-frame SR

The multi-frame super resolution is an active research area in last three decades. It sounds very attractive due to excellent mathematical formulations and justification to combine the information of multiple LR images to get a HR one. Moreover it has direct application for spatial video Super Resolution. Theoretically and sometimes for real images under global transformation those algorithm produces excellent results. However, in real scenario most of the cases it fails to achieve image quality up to that mark. Moreover, sometimes number of LR frame available is very low, or in most of the cases only one LR frame is available. Hence, in recent years single frame super resolution is attracting the attention of more researchers.

In Section 1.2.2, we have given a survey on Single frame SR methods and different state-of-the-arts techniques in this field. These algorithms are roughly categorized into two major approaches: (i) Interpolation-based methods and (ii) Patch-based methods. Traditionally people use interpolation based methods; however, they are lacking in modeling the visual complexity of the real images. Patch-based methods are most popular approach for this task. Those methods learn LR-HR patch correspondence from a database for natural LR-HR patch examples. Sometimes these methods are called 'Example-based SR'. Among popular approaches, Freeman [37] proposed an example-based learning strategy that

applies to generic images where the LR to HR prediction is learned via a Markov Random Field (MRF) solved by belief propagation. Yang *et al.* [121] explored the sparsity priors for this task, with the assumption that a LR patch and corresponding HR patch can be represented by the same sparse code given a dictionary of LR and HR patch pairs.

But for natural images, different HR patches can result in identical LR patch when blurred and down-sampled. In essence, there exists a many-to-one mapping while LR patches are generated from HR ones. So during the inverse process for SR reconstruction, various HR patches may be predicted from a single LR patch and the HR patch which is most coherent with the neighborhood is chosen to generate HR patch. Therefore, in this chapter we use multiple dictionary to predict multiple possibilities. Here, we try to discover a semantic relationship of different possible HR patches to a given LR patch using a probabilistic model called *probabilistic Latent Semantic Analysis (pLSA)*. Using the same model we also predict the higher level semantic in which the HR patches possibly belong to. We learn a separate predictor for each of the semantic. For this purpose we adopt and extend the model of Yang *et al.* [121] for sparse representation.

The rest of this chapter is organized as follows. In Section 5.1 We first discus about patch-based techniques for single frame SR and then give a brief overview of our framework and summarize the proposed algorithm in Section 5.2. In Section 5.3, we describe our algorithm in detail. Section 5.4 presents the experimental results and the comparison with those produced by other methods. Finally, we conclude this chapter in Section 5.5.

5.1 Patch-based technique for Single frame SR

The richness of real-world images is difficult to capture analytically. For the past several years, researchers have been exploring a learning-based approach for enlarging images [37, 30, 38, 121]. From a training set, the algorithm learns the correspondence of fine details that are seen in different regions of LR and HR images, and then uses those learned relationships to predict fine details in HR image from a given LR image.

To understand why this approach should work at all consider that a collection

of image pixels are special signals that have much less variability than a corresponding set of completely random variables. Researchers have studied these regularities to in different single frame SR algorithms. Patch-based SR exploit these regularities in small image portions called patches (typical size 5×5 to 10×10). This aim for the more attainable goal of synthesizing visually plausible image details, such as sharp edges and texture.

To generate a training set for patch-based technique, it starts from a collection of HR images and degrade each of them in a manner similar to 1.3 (assuming $F_k = I$) so that during the degradation it plan to undo in the images it later process. Typically, HR images are blurred and subsampled to create a LR image of fixed resolution factor s (say). The number of pixels reduces to $\frac{1}{s^2}$ of original number of pixels. To change resolution by higher factors, it typically use the single octave algorithm recursively. Usually an initial analytic interpolation is applied, such as bicubic interpolation, to the LR image. This produces an image of size same as the HR image but lacks HR detail. In the training set, only the differences between this interpolated image, considering it as LR image, and the original HR image need to store. The training set has to store the HR patch corresponding to every possible LR image patch. Even by restricting ourselves to plausible image information (small LR patches), this is a huge amount of information to store. So a preprocessing step is required for the image patches to remove variability and make the training sets as generally applicable as possible. The relationship between HR and LR image patches is essentially independent of local image contrast. Therefore, it is expected to apply a local contrast normalization to all the images. Even after normalization, huge size of training set makes simple nearest neighbor search so slow that the time required for SR method becomes prohibitively large. One possible solution is to apply k-means algorithm over this huge number of patch pairs and then discard all other patch pairs than the cluster centers to reduce the training set to a feasible one. However, arbitrary reduction of the cluster number and discarding variability among patch-pairs may results in serious visible artifacts.

Moreover, if local image information alone were used to predict the missing HR details, using the training set patch pairs directly, that may not end up with a good HR image. For a given LR input image, want to enlarge, would apply the

preprocessing steps, break the image into patches, and look-up the missing HR detail from the training dataset. Unfortunately, that approach does not work. The resulting HR would have discontinuity over the adjacent HR patches. Hence the local patch alone is not sufficient to estimate plausible HR detail we are looking for. Freeman [37] use a MRF model to overcome the problem. They have shown experimentally that for a LR patch of input image, k nearest neighbor LR patches from the database has no visible difference, however, corresponding HR patches has lot of differences and individually each of them could be a good candidate patch. They used MRF model to choose the best HR patch that agrees with the neighboring HR patches. Yang *et al.* [121] has chosen patch feature vector (usually the derivatives along both the directions) instead of raw LR patches that certainly incorporate the information of neighboring patches in a better way. They then represent each LR patch by sparse combination of LR component of the dictionary elements, and use the same sparse combination of HR component of the corresponding dictionary elements to generate the desired HR patch. This dictionary is learned from a database of natural LR and HR patch pairs. We adopt this model for our local patch prediction, and details of the technique employed is given in Section 5.2. However, instead of learning single dictionary suggested by them, we learn multiple dictionaries to capture more natural variability and the detail description and maintaining smaller search space as described in the following sections.

5.2 Proposed framework

It is mentioned earlier that a LR image patch can be generated from any of the several possible HR image patches. Here we propose is to build separate predictor corresponding to each of these possibilities to generate appropriate HR patch from a given LR patch. We analyze each patch in local context, and accordingly choose a suitable predictor to estimate most relevant HR patch. To achieve this we proceed as follows. Given a collection of natural images, our intention is to discover the inherent topics that group the collection of this natural image documents according to the possible different HR patches for a LR patch. Here we have used the general terms, viz. topics, documents and words that are

mostly used in text literature. In the context of image SR, we define "Document" as cropped portion of an image consisting of smaller patches called "Words". "Topic", on the other hand, is an even higher level concept that group different documents according to the co-occurrences of different words within and across the documents. In the subsequent section we will elaborate those terms.

Here, we briefly investigate the main frameworks that are used for this work:
(i) topic discovery - where document categories are discovered by pLSA clustering on all cropped regions of the available images (and topic classification/inference for the test image),
(ii) encoding sparse prior - prediction of HR patches from LR ones using sparse dual dictionary, and
(iii) back-projection of HR image - use gradient descent steps to back project HR image predicted by sparse dual dictionaries on LR image space to satisfy image reconstruction constraint.

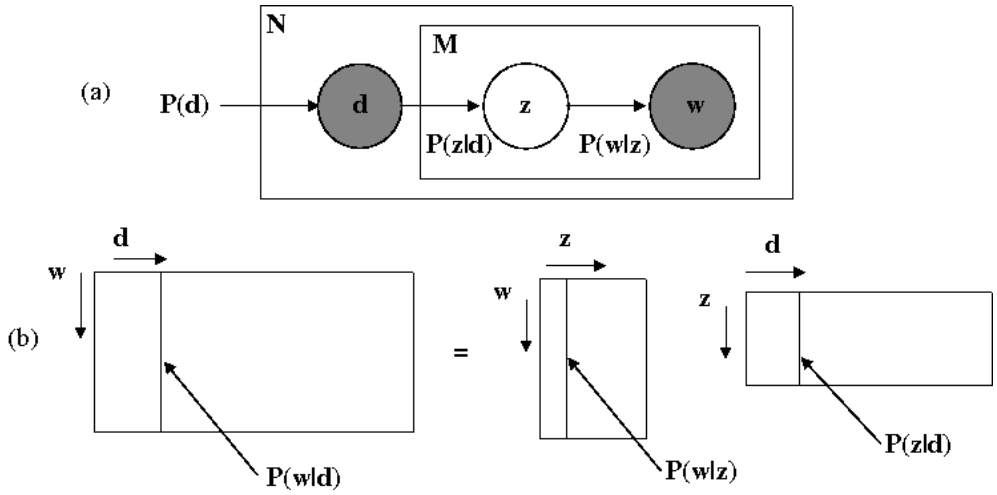


Figure 5.1: (a) Illustrates the pLSA graphical model: Nodes inside a given box indicate that they are replicated the number of times indicated in the top left corner. Filled circles indicate observed random variables and unfilled are unobserved. (b) Describes the document specific word distributions as described by Eq. (5.2.1). In pLSA the goal is to find the topic specific word distributions $P(w_i|z_k)$ and corresponding document specific mixing proportions $P(z_k|d_j)$ which make up the document specific word distribution $P(w_i|d_j)$.

The separation of documents based on its inherent topic may be accomplished

by the probabilistic Latent Semantic Analysis (pLSA) [44] or by the Latent Dirichlet Allocation (LDA) [1]. In the context of computer vision, this model have been used for object recognition [100] and natural scene categorization [54] successfully. Both the models usually give similar results. In this work we focus on the simpler pLSA model. These models use the 'bag-of-words' for document representation. Now in our case, the error due to quantization of feature vector using 'bag-of-words' model may lead to a serious error. Here, we propose a new feature descriptor to represent a document by the convex sparse code, and then by average pooling of all words (patches) inside the document (details are in sec. 5.3.2).

In the following section we briefly describe how the pLSA model discovers the inherent/latent topics in an unsupervised way.

5.2.1 pLSA: the topic discovery model

We describe the topic discovery model briefly in this section. Suppose we have N documents containing words from a vocabulary of size M. The corpus (collection of documents) can be represented in a $M \times N$ word-document co-occurrence table, where each entry in the table $n(w_i, d_j)$ stores the frequency of occurrences of the word w_i in the document d_j. Suppose there are K hidden (latent) topics z_k associated with the corpus and there are certain topic proportions corresponding to each document in the corpus. Then the joint probability density function $P(w_i, d_j, z_k)$ of words, documents and topics may be represented by the graphical model shown in Fig. 5.1(a). Marginalizing over the topic z_k determines the conditional probability $P(w_i|d_j)$:

$$P(w_i|d_j) = \sum_{k=1}^{K} P(z_k|d_j)P(w_i|z_k), \tag{5.2.1}$$

where $P(z_k|d_j)$ is the probability of topic z_k occurring in document d_j, and $P(w_i|z_k)$ is the probability of word w_i occurring in a particular topic z_k. The prior probability $P(d_j)$ of the documents are assumed to follow an uniform distribution. The model (5.2.1) expresses each document as a convex combination of K topic as shown in Fig. 5.1(b). Now the normalized co-occurrence table

$n(w_i, d_j)$ can approximate the conditional distribution $P(w_i|d_j)$. This leads to a matrix decomposition as shown in Fig. 5.1(b) with the constraint that both the topic vectors ($P(w_i|z_k)$) and the mixture coefficients ($P(z_k|d_j)$) are normalized to make them probability distributions. We could have obtained approximate values of those distributions by applying singular value decomposition (SVD) on document specific word distributions $P(z_k|d_j)$. However, this would then suffer from the well-known ambiguities Polysemy and Synonymy. pLSA [44] and LDA [1] handle those issues in alternate way very efficiently. In pLSA/Topic model, essentially, each document is modeled as a mixture of topics, where the histogram for a particular document is composed of a mixture of histogram corresponding to each topic. Fitting the model involves determining (i) the topic vectors $P(w_i|z_k)$ which are common to all documents, and (ii) the document specific mixture coefficients $P(z_k|d_j)$. A maximum likelihood estimation of the parameters is obtained by maximizing the objective function:

$$E = \prod_{i=1}^{M} \prod_{j=1}^{N} P(w_i|d_j)^{n(w_i,d_j)}, \tag{5.2.2}$$

where $P(w_i|d_j)$ is given by (5.2.1) and $n(w_i, d_j)$ is the co-occurrence matrix corresponding to the corpus. The model is fitted using the Expectation Maximization (EM) algorithm [44]. This is equivalent to minimizing the Kullback-Leibler divergence between the measured empirical distribution $P(w|d)$ and the fitted model. The maximum mixture coefficient $z_0 = \max_{z_k} P(z_k|d_j)$ would assign the topic z_0 to document d_j. The topic corresponding to each word (patch) inside a document is determined accordingly.

5.2.2 Encoding sparse prior on natural image patches

Unlike the traditional patch-based methods, in this work we generate the HR patch for each input LR patch based on the topic assigned to it. Let a topic z_k is assigned to LR patch y in document d_j by the topic model pLSA as described above. For this local model, we have two dictionaries D_k^l and D_k^h, where D_k^h is composed of HR patches and D_k^l is composed of corresponding LR patches from the topic z_k. For an input LR patch y and corresponding HR patch x, we find a

sparse representation α_k^* with respect to dual dictionary $[D_k^l; D_k^h]$. The problem of finding the sparsest representation of $[y; x]$ can be formulated as [121]:

$$\alpha_k^* = \arg\min_{\alpha_k}\{\|\alpha_k\|_0 : \|D_k^l\alpha_k - Fy\|_2^2 \leq \epsilon\}, \tag{5.2.3}$$

where F is a (linear) feature extraction operator (e.g., first and second order derivative operators) and $\|\alpha_k\|_0$ corresponds to l_0 norm that counts number of non-zero entries in α_k. Then the corresponding HR patch x is generated according to these coefficients α_k using D_k^h, i.e, $x = D_k^h\alpha_k^*$.

Although the optimization problem (5.2.3) is NP-hard in general, recent results [28, 61] indicate that as long as the desired coefficient vector α_k is sufficiently sparse, they can be efficiently recovered by minimizing the l_1 norm instead, as follows:

$$\alpha_k^* = \arg\min_{\alpha_k}\{\|\alpha_k\|_1 : \|D_k^l\alpha_k - Fy\|_2^2 \leq \epsilon\}, \tag{5.2.4}$$

where $\|\alpha_k\|_1$ corresponds to l_1 norm that introduce sparsity.

Solving Eq. (5.2.4) individually for each patch does not guarantee compatibility between adjacent patches. So we use the same formulation as in [121] to enforce compatibility between adjacent patches in the following way. Yang *et al.* [121] have modified Eq. (5.2.4) so that the super resolution reconstruction $D_k^h\alpha_k^*$ of patch x is constrained to closely agree with the previously computed HR patches w. The resulting optimization problem is represented as

$$\alpha_k^* = \arg\min_{\alpha_k}\{\|\alpha_k\|_1 : \|D_k^l\alpha_k - Fy\|_2^2 \leq \epsilon_1 \text{ and } \|PD_k^h\alpha_k - x\|_2^2 \leq \epsilon_2\}, \tag{5.2.5}$$

where the matrix P extracts the region of overlap between the current target HR patch and the previously reconstructed HR image. The constraint optimization problem (5.2.5) can be reformulated as unconstrained optimization problem as:

$$\alpha_k^* = \arg\min_{\alpha_k}\{\lambda\|\alpha_k\|_1 + (1/2)\|\hat{D}_k\alpha_k - \hat{y}\|_2^2\}, \tag{5.2.6}$$

where $\hat{D}_k = [D_k^l; \beta PD_k^h]$ and $\hat{y} = [Fy; \beta x]$. The parameter β controls the trade-off between matching the LR input and finding a HR patch that is compatible with its neighbors, and λ corresponds to the degree of sparsity in sparse code repre-

sentation. Large value of λ implies more sparsity which results more smoothness to the texture in the output. In all our experiments, we set $\beta = 1$ and $\lambda = 0.15$. Given the optimal solution α_k^*, the HR patch can be reconstructed as $x = D_k^h \alpha_k^*$.

In the training phase, the coupled dictionary $\hat{D}_k$ are learned from the known patch-pairs x and y of the documents belong to topic z_k. This is done by jointly optimizing $\hat{D}_k$ and α_k in Eq. (5.2.6). During SR reconstruction the estimation of HR patch x from LR patch y that belong to the document under topic z_k, we optimize α_k and x in Eq. (5.2.6) iteratively, and choose predicted HR patch as $x^* = D_k^h \alpha_k^*$ when it converges.

5.2.3 Summary of proposed algorithm

The proposed single-frame SR method may be described as a two phase algorithm: (1) Learning phase - Inherent topics comprising the natural patches are learned from an image database using pLSA model and also the dual dictionary for each topic is learned, and (ii) Testing phase - HR image is predicted from the test LR image via topic assignment to each patch and its sparse representation using the dual dictionary of that topic learned in the first phase. The entire super-resolution process is summarized in Algorithm-1 and Algorithm-2.

5.3 Detail algorithmic steps

5.3.1 Data-set Generation

To develop a system for predicting HR image, a large number of training pairs of HR-LR images are required. We have downloaded a number of sharp natural images from '*flickr*'[1]. Some of those images are shown in Fig. 5.2. The color images are first converted from RGB to YCbCr format. Then only the Y (intensity) channel is considered as HR image in our image data-set and is used for further processing. The other two components, which contain chromaticity information, are ignored at this stage. We generate LR versions corresponding to each of these HR images in the following way:

[1]Online photo management and sharing application, http://www.flickr.com/

Algorithm 1 : Learning inherent topics and corresponding dual dictionaries

Input : A collection of sharp natural images, number of topics K, size of the dictionary L, patch size n and document size m.

1. From all the natural HR images generate corresponding LR images by some resolution factor s.
2. Divide the natural HR images and the corresponding LR ones into overlapping documents of size $m \times m$ and the documents into overlapping patches of size $n \times n$.
3. Collect all the patch pairs - one HR patch and another corresponding LR patch, and learn a convex sparse dictionary called as topic-trigger dictionary.
4. Represent each document by a feature vector obtained by average pooling of the sparse codes of each of the patches over the document.
5. Run the Expectation Maximization (EM) algorithm on the document feature vectors to find out the inherent topic depicted by the documents.
6. Calculate maximum mixture coefficients (of assigned topics) for each of the documents and partition the document database accordingly.
7. On the documents under each of the topic, apply dual dictionary learning technique to generate individual dictionary for each topic.
8. Save the word-topic distributions and the dual dictionaries for subsequent SR image reconstruction from the given LR image.

- Apply a Gaussian blur operator on each of these intensity (Y) images with a small kernel (size $[3 \times 3]$ and σ =0.8).

- Down-sample the blurred image with a resolution factor s (=1.6). This is the synthesized LR version of a given intensity image.

- Then up-sample the LR image with the same resolution factor s by bicubic interpolation. Thus original HR and corresponding up-sampled LR images are of same size. The main difference is that the LR image contains blurring effects and the ringing noise, whereas the HR one has sharper edges.

Algorithm 2 : Super Resolution using Topic Model and sparse representation

Input: Dual dictionary for each topic, the topic-trigger dictionary, a LR test image and the resolution factor.

1. Divide the image into overlapping documents and the documents into overlapping patches and then represent each document in the same manner as was during training.
2. Assign topics to the documents using EM algorithm.
3. Using the dual dictionary $[D_k^l;\ D_k^h]$ of the inferred topic of a document predict the HR patches in the document by sparse representation.
4. Make an average of all the predicted HR patches in overlapping region of the overlapping documents.
5. Then to satisfy the LR formulation constraint apply back-projection technique (5.3.6) to get the estimated HR image.
6. Apply steps $2-6$ repeatedly until the SR image with desired resolution factor is achieved.

In Figure 5.2 we show the formation of LR image in the same order as stated above.

The recent works [121, 105, 35] suggest that the derivative features can represent a patch more efficiently than the pixel color or intensities. In this work, we use four 1-D derivative operators $\delta_1 = [-1, 0, 1]$, $\delta_2 = \delta_1^T$, $\delta_3 = [1, 0, -2, 0, 1]$ and $\delta_4 = \delta_3^T$ to extract the LR patch features. We apply these operators to all the up-sampled LR images. Then we select some patches of size $(n \times n)$ randomly from HR image. The four derivatives of each LR patch are reshaped and then are concatenated into a vector. We remove the vectors for which the patch does not contain any texture or structure. This is determined by the variance over the patch. In other words, we remove the smooth patches having very low variance. Then each LR patch feature vector and corresponding HR patch vector are independently normalized to mean zero and standard deviation one. It is assumed that LR and HR patches have the same mean. Moreover, in [121], it is shown that the standard deviation of HR patches is approximately 1.2 times the standard deviation of the corresponding LR patch. We also incorporate this

Figure 5.2: Top row: Some downloaded HR natural images. These color images are mapped into YCbCr space and only the intensity Y plane is stored. Bottom row (left to right): LR image formation process in the order as in the text. Bottom-right image is the LR version of the bottom-left HR image.

statistics of natural images into our model, and is utilized during prediction so that we can maintain the mean and standard deviation of the LR patch in the corresponding HR patch. We divide the HR images and corresponding filtered LR images into overlapping documents of size $(m \times m)$ and the documents into overlapping patches of size $(n \times n)$. These patches are concatenated into a vector comprising HR and LR feature patch-pair for learning the dual dictionary.

5.3.2 Topic discovery

From the given set of patch-pairs, our target is to discover the inherent topics within the corpus, so that it would be possible to semantically group possible HR patches corresponding to LR patches. To apply topic model, we need to represent each document by a collection of words from a vocabulary. The vocabulary may be built by applying k-means clustering algorithm over all patch-pairs available in the database. The cluster centers then comprise the vocabulary and each patch pair vector is represented by the closest element in the vocabulary. This leads to well-known 'bag-of-words' model in computer vision. As we stated earlier, in our case 'bag-of-words' model may not be a good choice due to quantization error, instead we use sparse coding of the patches. Then the average pooling of

these codes over a document represents the document. Furthermore, we require a model that can analyze a LR patch locally and then predict the HR patch based on semantically meaningful topic assignment. Topic model does this job very efficiently.

5.3.2.1 Trigger dictionary learning:

A convex sparse dictionary is learned over the all patch-pairs in the corpus. Here the term convex specifies that each patch pair vector can be represented by a convex combination of dictionary elements. To ensure this, we introduce two more constraints in the optimization problem (5.2.6), viz, $\alpha_k(m) \geq 0$ and $\sum_m \alpha_k(m) = 1$ and solve for $\hat{D}_k$ and α_k jointly. We call this as a 'Trigger Dictionary' because it represents each document by a feature vector that can discover the topics in the corpus and also assigns topic to the test documents. In the following subsection we elaborate these in detail.

5.3.2.2 Document representation:

Here we propose a new descriptor to summarize a document. This is formed in two steps: (i) patch coding followed by (ii) average pooling. Let the document d_j has L patches p_l. In the patch coding step, we approximate each patch p_l by a convex sparse representation α_l of the trigger dictionary of size M. Then each patch can be represented by a vector α_l of size M. In the average pooling step, codes associated with the patches are pooled over the document and represent each document d_j by v_{d_j}, where $v_{d_{(j,m)}} = \sum_{l=1}^{L} \alpha_{(l,m)}$ for $m = 1, 2, \ldots, M$. This document feature vector v_{d_j} is used for topic learning and also to group the documents based on its inherent topics.

5.3.2.3 EM algorithm for topic learning:

In the Section 5.2.1, we have seen that for pLSA model, only observed variable is $n(w_i, d_j)$. We consider this document-word co-occurrence matrix as $[v_{d_1}, v_{d_2}, \ldots, v_{d_\mathbf{N}}]$, where N is the total number of documents in the corpus. We minimize the log likelihood in Eq. (5.2.2) using an Expectation Maximization (EM) algorithm. EM alternates two coupled steps: (i) an expectation (E) step where the posterior

probabilities are computed for the latent variables, and (ii) a maximization (M) step, where the parameters are updated. Standard calculations yield the E-step equation

$$P(z_k|w_i, d_j) = \frac{P(z_k)P(d_j|z_k)P(w_i|z_k)}{\sum_{k'=1}^{K} P(z_{k'})P(d_j|z_{k'})P(w_i|z_{k'})} \tag{5.3.1}$$

as well as the following M-step formula

$$P(w_i|z_k) \propto \sum_{j=1}^{N} n(w_i, d_j)P(z_k|w_i, d_j) \tag{5.3.2}$$

$$P(d_j|z_k) \propto \sum_{i=1}^{M} n(w_i, d_j)P(z_k|w_i, d_j), \tag{5.3.3}$$

$$P(z_k) \propto \sum_{j=1}^{N}\sum_{i=1}^{M} n(w_i, d_j)P(z_k|w_i, d_j). \tag{5.3.4}$$

At each iteration the right-hand side of equations in M-step are normalized to follow a probability distribution. After Learning, $P(z_k|d_j)$ ($\propto P(z_k, d_j) = P(d_j|z_k)P(z_k)$), represents the mixture proportions of each document, and its maximum value for each of the document may be assumed as the assigned document topic. In Figs. 5.3 and 5.4, we show the documents under different topics. It is clear from the figure that the some topic consists of documents of overall homogeneous regions with sharp edges, while some other consists of mostly texture regions. $P(w_i|z_k)$ and $P(z_k)$ are stored for future EM algorithm that would be used for test document topic prediction.

For each document, we calculate the corresponding topic assignment. Hence this probabilistic model based clustering method separates out all the documents corresponding to the assigned topic. Now we learn a sparse dual dictionary for each of the topic cluster over the documents assigned to it.

5.3.3 Sparse dual dictionary learning

For each topic z_k we learn a sparse dual dictionary $[D_k^l; D_k^h]$ (as in Eq. (5.2.6)) over the words under the topic z_k. During learning, we have known HR patches x and corresponding LR patches y, and we solve the optimization problem (5.2.6) for $\hat{D}_k$

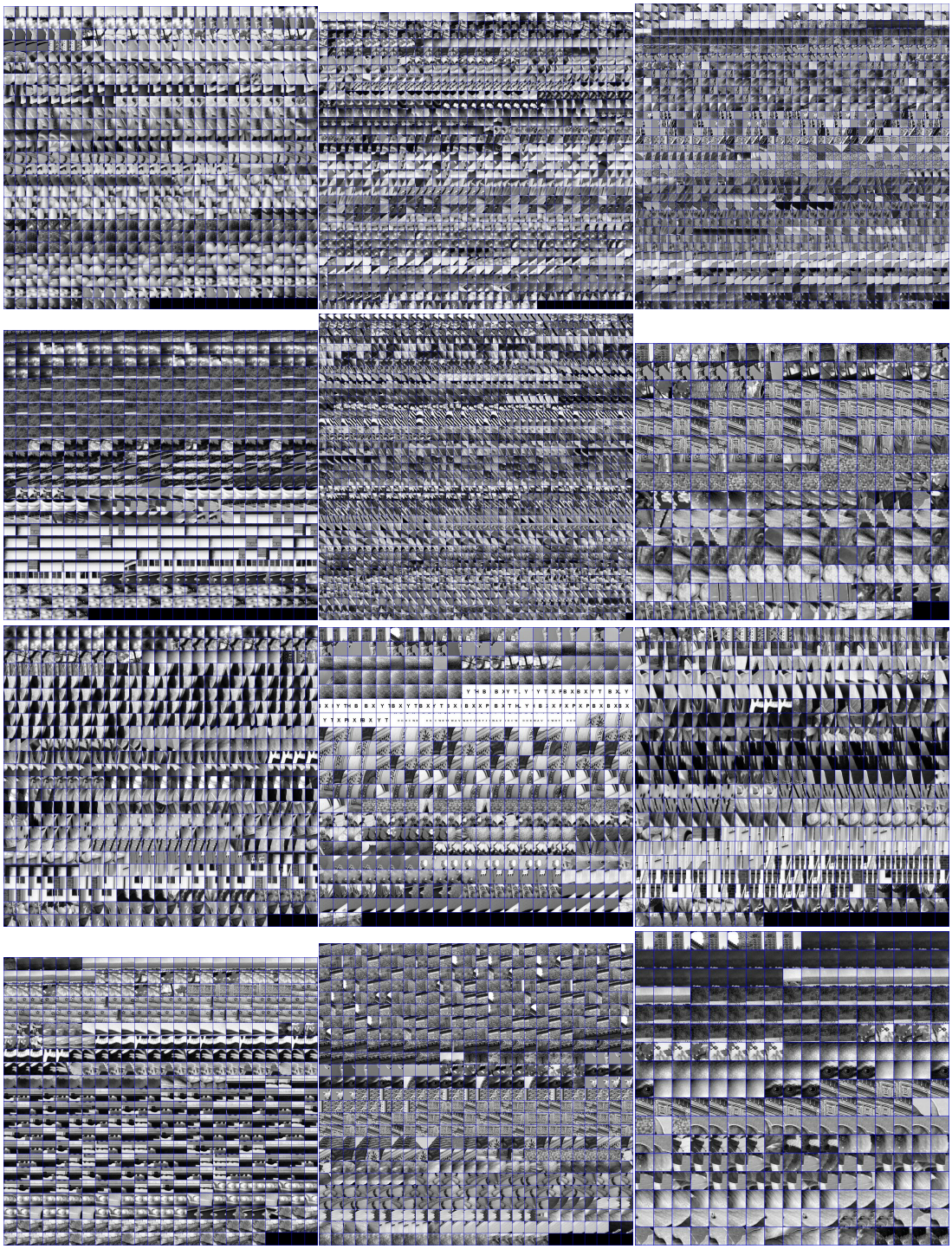


Figure 5.3: Documents under different typical topics are displayed. Documents under topics 1 - topic 12.

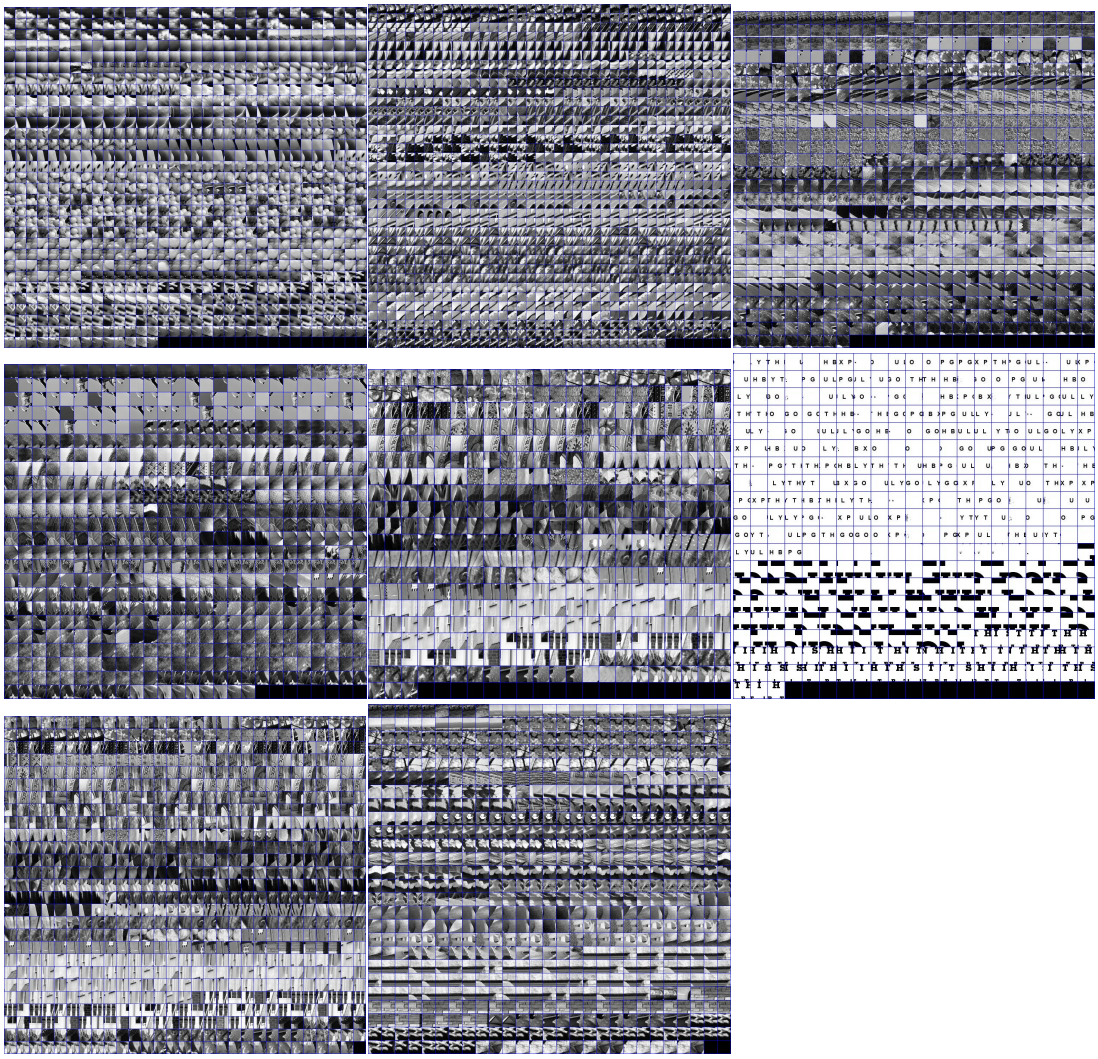


Figure 5.4: Documents under different typical topics are displayed. Documents under topics 13 - topic 20.

and α_k. It is a time consuming process as we have to learn individual dictionary for each topic z_k. However, it is done offline prior to SR reconstruction for test images. After learning each of them can predict a HR patch x^* corresponding to a LR patch y from a test LR image.

5.3.4 SR image reconstruction

For SR reconstruction, the input LR image is first up-sampled using bicubic interpolation with resolution factor s, and then is divided into overlapping documents and the documents into overlapping patches in the same way as was done during training. We represent each document by the feature vector obtained by average pooling over the convex sparse code of the comprising patches. In this case we use only the LR components of the vectors of the trigger dictionary.

5.3.4.1 Topic inference for each document:

Documents of the LR image are analyzed for possible topic assignment for each document. This is also done by EM algorithm [Eqs. (5.3.1)-(5.3.4)]. In this case, the topic specific word distributions $P(w_i|z_k)$ and the topic prior probability $P(z_k)$ are known apriori (calculated during the training phase). Therefore we need to execute only Eq. (5.3.1) and Eq. (5.3.3) alternatively. After convergence the topic corresponding to the maximum value of $P(z_k|d_j)$ is assigned as the topic associated to the document d_j.

5.3.4.2 Predicting HR patches:

Each LR patch y of a document belonging to the topic z_k is represented by a sparse code α_k in the dual dictionary $[D_k^l; D_k^h]$. First, we initialize HR patch x with the bicubic interpolated patch and then optimize Eq. (5.2.6) repeatedly until it converges. For optimal solution α_k^*, the predicted HR patch becomes $x^* = D_k^h \alpha_k^*$. We do this for all the patches in the test image, and then take pixel-wise average of overlapping portions of the patches to get the estimated HR image. A post refinement is done on the estimated HR image using back-projection technique to satisfy LR image formation constraint.

5.3.4.3 Back-projection to satisfy the constraints:

The fundamental constraint of SR image reconstruction is that the reconstructed HR image X should be consistent with the input LR image Y. The observed LR image Y should be a blurred and down-sampled version of the solution X. Mathematically,

$$Y = \downarrow_s HX. \tag{5.3.5}$$

Here, H represents a blurring filter and $\downarrow_s$ is the down-sampling operator with factor s. Therefore our estimated HR image is made to satisfy this constraint by updating it using an iterative method as

$$X^{(t+1)} = X^{(t)} + \gamma H^T \uparrow_s (Y - \downarrow_s HX^{(t)}), \tag{5.3.6}$$

where $X^{(t)}$ is the estimated HR image after the t-th iteration, γ is an iterative constant, and $\uparrow_s$ denotes the up-sampling by a factor s. We take the result X^* from back-projection as the final estimate of HR image. This image is as close as possible to the initial super-resolution X_0 given by sparsity, while satisfying the reconstruction constraint. In all our experiments we have chosen H as Gaussian kernel (size $= 3 \times 3$ and $\sigma = 0.2$) and $\gamma = 0.2$.

5.4 Experimental results

We have implemented our algorithm using MATLAB 7.6 on a Linux OS with 8GB of RAM and a 3.6-GHz Intel processor. We have used mex code for some of the subroutines to speed up the process. We have chosen a small up-sampling factor ($s = 1.25$) for the proposed algorithm. HR image with higher up-sampling factor are generated by applying the algorithm required number of times. It takes usually less than 10 seconds for a typical 200×200 LR image with resolution factor 4. To learn the sparse dictionaries and the sparse representation of patches [Eq. (5.2.6)], we have used the optimization toolbox SPAMS (SPArse Modeling Software) version 2.1.

When working with color images, the image is first converted from RGB to YCbCr. Then the SR algorithm is applied to Y (intensity) channel only. For

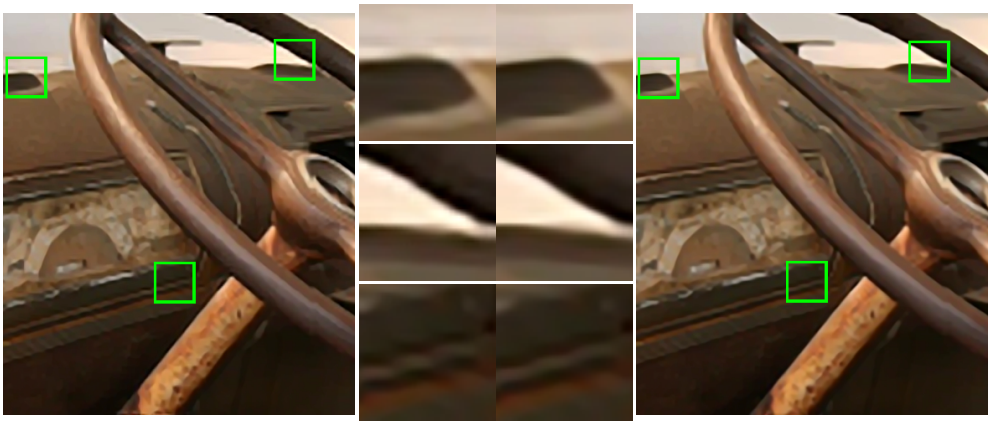


Figure 5.5: Comparison of reconstruction results of wheel image, generated using single global dictionary proposed by Yang *et al.* [121] and proposed method with multiple dictionaries. The marked portions of each of the images are cropped and displayed at bottom for better view.

rest of the components (Cb and Cr which are usually low frequency channels and stores the chromaticity information only) we use simply bicubic interpolation. The three modified components are then combined and converted back to RGB to get the estimated color SR image.

In our experiments we have chosen $20,000$ documents of size 50×50, with patch size 7×7, and then have applied pLSA model with 20 topics. The typical size for dual dictionaries as well as for trigger dictionary is chosen 500. In Fig. 5.5, we compare our result with Yang *et al.* [121]. It is clear from the image that the result of our method contains less artifacts and more sharp boundaries. We also compare our method with the state-of-art techniques and got better or equivalent result. Some results are depicted in Figs. 5.6-5.10.

The a quantitative comparison is shown in the following table 5.1.

5.5 Summary

In this chapter we have adopted pLSA based technique for single-frame super resolution via sparse representation. Since a LR patch can possibly be generated from any of multiple possible HR patches, we needed multiple predictors corre-

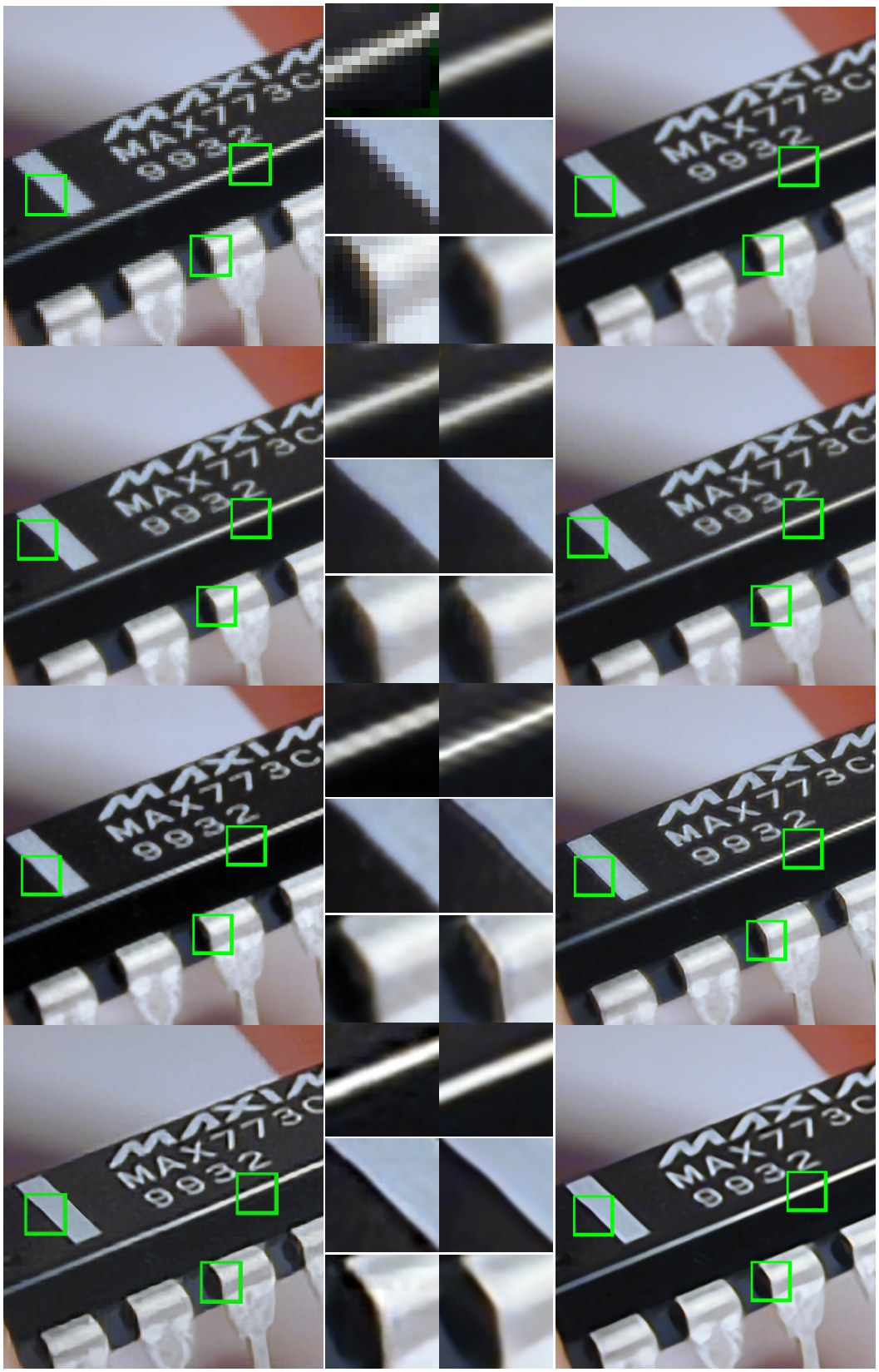


Figure 5.6: Comparison of reconstruction results of chip image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Perfect Resize (commercial software for image zooming), Sun *et al.* [105], Fattal *et al.* [34], Glasner *et al.* [38], Yang *et al.* [121] and proposed method.

Table 5.1: PSNR(dB) and SSIM results of reconstructed HR images (noise level $\sigma_n = 0$)

Methods / Images	**Bicubic**		**Yang *et al.*[121]**		**Proposed**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Lenna	28.83	0.9160	28.84	0.9242	**29.08**	**0.9280**
Airplane	26.19	0.9163	26.63	0.9260	**26.66**	**0.9297**
Peppers	27.92	0.9376	29.05	0.9388	**29.07**	**0.9423**
Barbara	**23.84**	0.8155	23.67	0.8238	23.75	**0.8269**
Boat	26.31	0.8675	26.49	0.8876	**26.68**	**0.8920**
Couple	27.66	0.8431	27.82	0.8671	**27.94**	**0.8696**
Tiffany	**30.38**	0.9240	30.17	0.9277	30.24	**0.9304**
Splash	30.70	0.9465	32.13	0.9470	**32.61**	**0.9499**

sponding to each of the possibilities and then select the most relevant predictor to produce the HR patch. So we do semantic analysis of each of the patches locally to decide which topic it belongs to, and the corresponding predictor is used. We adopt the sparse dual dictionary learning scheme corresponding to every topic that can predict a possible HR patch. Our experimental results show that proposed method performs superior to the existing methods in most of the cases.

Figure 5.7: Comparison results of SR for koala image using proposed method over other state-of-arts methods. Output images along with some cropped portions using different methods are displayed. Top to bottom and left to right images are: input image, upsampled image using bicubic interpolation, Perfect Resize, Glasner *et al.* [38], Freedman *et al.* [36], and proposed method. It is clear from the images that proposed technique can produce comparable results with the state-of-art methods.

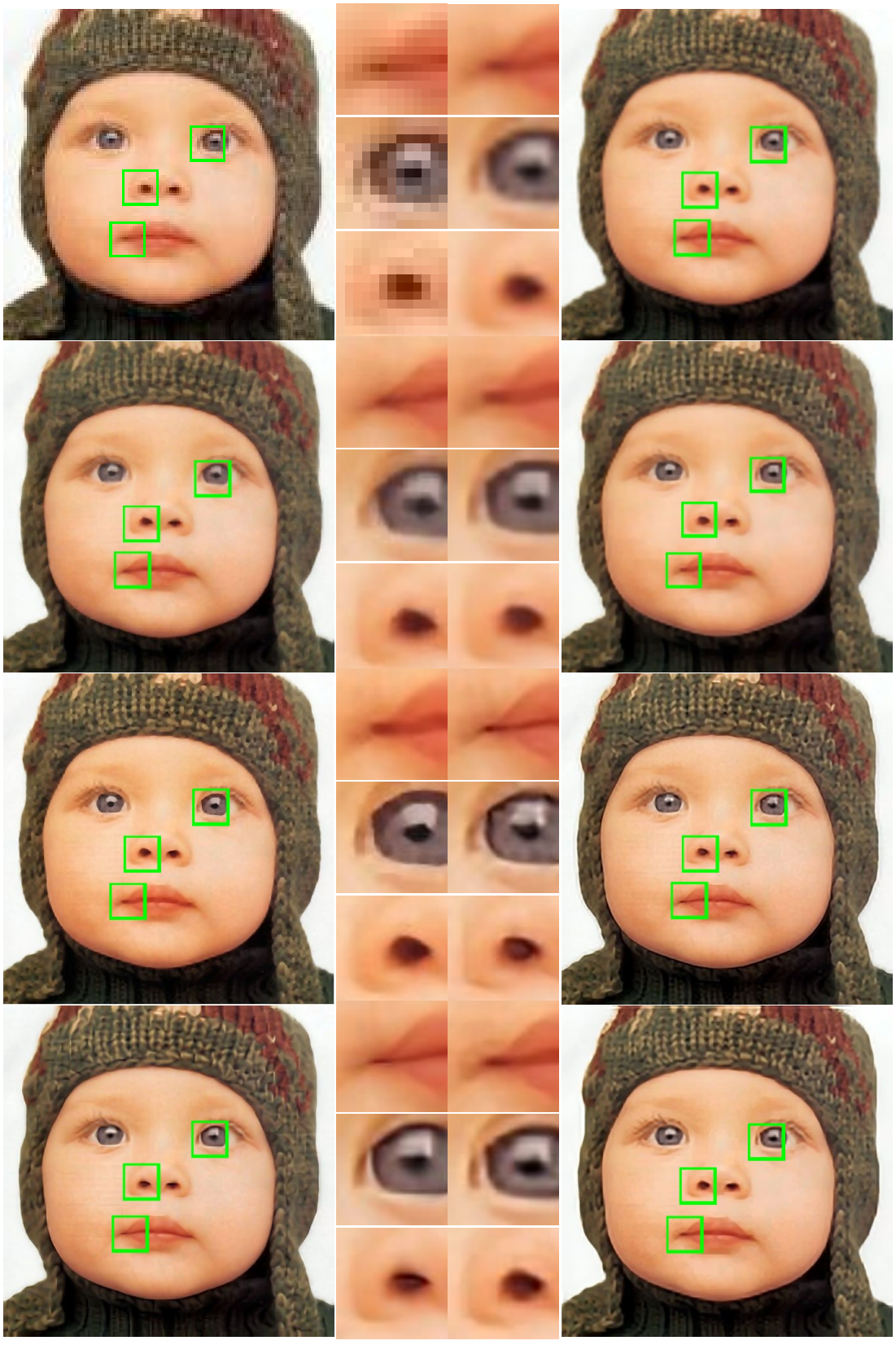


Figure 5.8: Comparison of reconstruction results of child image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Perfect Resize, Sun *et al.* [105], Fattal *et al.* [34], Glasner *et al.* [38], Freedman *et al.* [36] and proposed method.

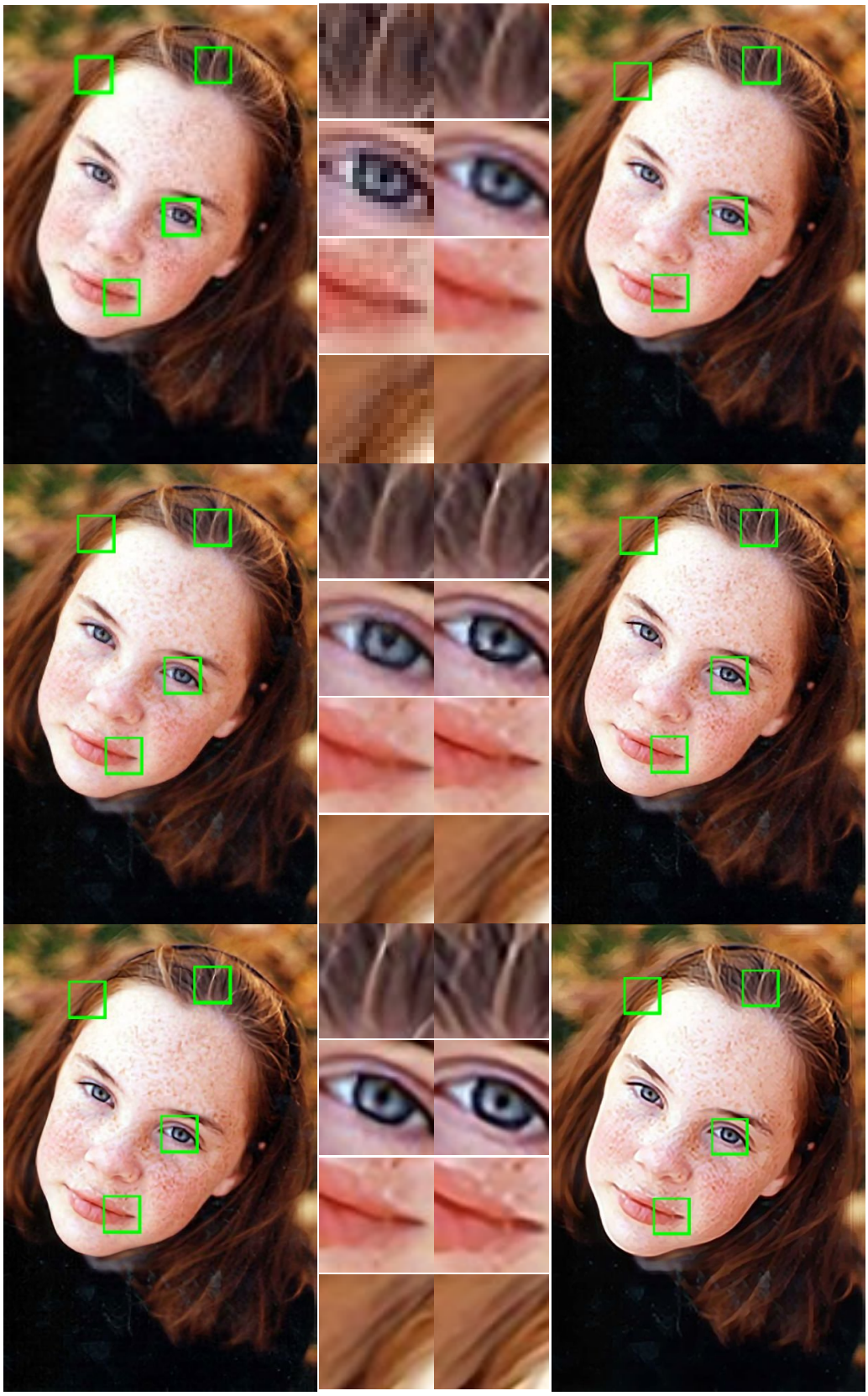


Figure 5.9: Comparison of reconstruction results of girl image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Perfect Resize, Glasner *et al.* [38], Freedman *et al.* [36] and proposed method.

Figure 5.10: Comparison of reconstruction results of sclpture image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Fattal *et al.* [34], Glasner *et al.* [38], Freedman *et al.* [36] and proposed method.

Chapter 6

Fuzzy-rule Based Approach for Single-frame SR

In the last chapter, we have discussed effectiveness of sparsity prior for single-frame SR image reconstruction [50, 24, 65]. In these works it is assumed that LR patches and corresponding HR patches lie on the same Manifold in l_1 space. Hence if a given LR patch is a sparse convex combination of LR patches in a LR component of dual dictionary, the corresponding HR patch may be estimated using the same convex combination of HR patches of the HR component of the same dual dictionary. We extended this algorithm by developing different dictionaries, and thus, different predictors for different regions depending on the local statistics of patches.

On the other context, in the last few decades, there has been a growing research interest in the applications of soft computing techniques, such as neural networks and fuzzy systems, to the problems in digital image processing. That includes image filtering [93, 119, 25, 92, 13], image segmentation [85, 4, 55], image classification [60, 104, 74] and image interpretation [56, 45]. Several fuzzy filters such as the iterative fuzzy control based filters [32], and the GOA filter [115] for noise reduction have already been developed. Most of these state-of-the-art filtering methods are mainly developed for the reduction of fat-tailed noise like impulse noise. Also there are fuzzy techniques that produce convincing results for additive noise, which is illustrated in [93] and [94]. Neuro-fuzzy Systems (NFS) offer

the ability of neural networks to learn from examples. Fuzzy systems also have capabilities to model the uncertainty which is inevitably encountered in noisy environments [124]. They can also manage imprecision embedded in the experts knowledge of the concerned domain. They constitute an adequate framework for knowledge representation and reasoning, and thus, reducing the semantic gap between symbolic concepts and numerical data [45]. Fuzzy rule-base system allows representing the imprecision which is inherent to the definition of certain concepts. For instance, the concept "close to" is intrinsically vague and imprecise, and its semantics depends (i) on the context in which objects are embedded, (ii) on the scale of the objects and (iii) of their environment. They are suitable for the development of the learning algorithms for computer vision/image understanding because they, being a nonlinear knowledge-based methods, are able to deal with ambiguities in a robust way.

Now, for natural images, different HR patches can result in identical LR patch. Hence, it would be more effective if we model each of the HR possibilities corresponding to a given LR patch with different fuzzy-rules. Then those rules are combined according to their strengths to preserve spatial coherence. In this work, we learn these fuzzy-rules from a collection of LR-HR patch pairs. Here also we assume that LR patch feature space and HR patch feature space lie on the same manifold in Reimann space and use similar kind of mechanism for learning nonlinear mapping between input and output space as suggested by Pal *et al.* [80] for structure preserving dimensionality reduction.

The rest of the chapter is organized as follows. In the next section we describe the main intuition behind the proposed algorithm, followed by a detail description of learning mechanism. The SR image reconstruction using fuzzy-rules and the required time comparison with other methods are illustrated in Sections 6.2 and 6.3 respectively. In Section 6.4, we show some experimental results of the proposed methods as well as of some state-of-the-art methods. Finally, we give a summary in Section 6.5.

6.1 Proposed fuzzy model for SR

In this section we describe the framework of the proposed SR technique. We propose a fuzzy-rule based system for predicting HR patches corresponding to each of the LR patch of an input LR image. It has two phases (i) Offline phase: where we generate a rule based system, and learn the corresponding parameters from a collection of LR-HR patch pairs and (ii) Online phase: where we apply trained rule based system on the patches of a test LR image to generate its HR version. In the following subsections we discuss these two phases in detail.

6.1.1 Data-set Generation

We use the same technique for generating LR-HR patch pairs as discussed in Section 5.3.1. Those LR and HR patch pairs are concatenated into a single vector comprising LR and HR patch-pair $\mathbf{x}^* = (\mathbf{x};\ \mathbf{y})$ to learn the fuzzy-rules for prediction. For $(n \times n)$ size patch the length of the LR patch feature vector is $4 * n^2$ and that for HR patch vector is n^2. Instead of taking all patch pairs, some patch pairs are randomly selected from each of the image and build our LR-HR patch pair data-set. This selection is done to reduce the training data.

6.1.2 Fuzzy rule learning

Let us denote our training data-set consisting of a set of LR patches denoted by $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \mid \mathbf{x}_i \in \mathbb{R}^p\}$ and the corresponding HR patches by $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N \mid \mathbf{y}_i \in \mathbb{R}^q\}$. Suppose, $p = 4{*}n^2$, $q = n^2$ for $n \times n$ patch and N is the number of patches in the data-set. We concatenate corresponding patch pairs, so the resultant LR-HR patch-pair data-set is represented as

$$X^* = \{\mathbf{x}_i^* = (\mathbf{x}_i;\ \mathbf{y}_i) \in \mathbb{R}^{p+q},\ i = 1, 2, \ldots, N\} \tag{6.1.1}$$

where $\mathbf{x}_i^*$ is the i^{th} LR-HR patch-pair in the data-set.

We partition our patch-pair data-set X^* into C clusters with cluster centroids

$$V^* = \{\mathbf{v}_c^* = (\mathbf{v}_c^x;\ \mathbf{v}_c^y) \in \mathbb{R}^{p+q},\ c = 1, 2, \ldots, C\}. \tag{6.1.2}$$

where $\mathbf{v}_c^* = (\mathbf{v}_c^x;\ \mathbf{v}_c^y)$ is the c^{th} cluster center.

Here, we assume that input LR patch feature space and output HR patch space lie on the same manifold. Therefore, after clustering the LR and HR patch features simultaneously, if there is a cluster in the LR patch feature space with centroid $\mathbf{v}_c^x$, then the corresponding points in the HR patch space are likely to form a cluster around $\mathbf{v}_c^y$. Suppose there is a good cluster around $\mathbf{v}_c^*$ in patch-pair space, then for any patch-pair $(\mathbf{x}_i;\ \mathbf{y}_i)$ in that cluster: if $\|\mathbf{x}_i - \mathbf{v}_c^x\|$ is small then $\|\mathbf{y}_i - \mathbf{v}_c^y\|$ would also be small. In other words, if $\mathbf{v}_c^x$ be an rough estimate of an LR patch feature $\mathbf{x}_i$ then $\mathbf{v}_c^y$ would be a rough estimate of the corresponding HR patch $\mathbf{y}_i$. Thus, for each cluster a rule can be written expressing this concept of prototype. Combining all those rules we get a locally continuous and smooth predictor that maps input LR patch to an appropriate HR patch. In Fig. 6.1, we have shown some cluster centers (only normalized $\mathbf{v}_c^y$'s are displayed).

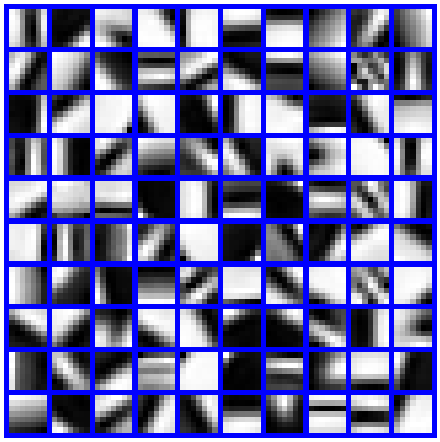


Figure 6.1: Representing the HR part of the cluster center of LR-HR patch pairs after clustering.

Note that we could have applied clustering technique, on LR patch feature space and HR patch space separately. However, in that case establishing correspondence between LR and HR cluster centers would have been hard to achieve.

6.1.3 Fuzzy rule Identification scheme

We formulate the rule in respect of the c^{th} cluster following Takagi-Sugeno (TS) model [106] of the form

$$R_c^{TS}: \text{ If } \mathbf{x} \text{ is CLOSE to } \mathbf{v}_c^x \text{ then } \mathbf{y} = u_c(\mathbf{x}, \mathbf{v}_c^y). \tag{6.1.3}$$

The behavior of the function $u_c(.)$ from input LR patch feature space to HR patch space relies on the neighborhood of $\mathbf{v}_c^y$. This set of C rules forms an initial rule base for HR patch prediction. Now as the LR patch feature space is p-dimensional, i.e, $\mathbf{x} \in \mathbb{R}^p$, the antecedent part of R_c^{TS} can be viewed as a conjunction of p atomic clauses:

$$\text{If } \mathbf{x} \text{ is CLOSE to } \mathbf{v}_c^x \approx \begin{pmatrix} \text{If } x_1 \text{ is CLOSE to } v_{c1}^x \\ \text{AND If } x_2 \text{ is CLOSE to } v_{c2}^x \\ \dots \\ \text{AND If } x_p \text{ is CLOSE to } v_{cp}^x \end{pmatrix} \tag{6.1.4}$$

Note that, the right-hand-side of (6.1.4) is one of the plausible interpretations of the left-hand-side and they are not necessarily identical. We compute the response $\pi_{cj}(x_j)$ of each individual clauses "If x_j is CLOSE to v_{cj}^x" and then choose a T-norm for the intersection of these p clauses to get the firing strength of the c^{th} rule.

Moreover, since consequent part $\mathbf{y} \in \mathbb{R}^q$ is a q-dimensional vector, the c^{th} rule R_c^{TS} corresponds to q different rules as

$$R_{cj}^{TS}: \text{ If } \mathbf{x} \text{ is CLOSE to } \mathbf{v}_c^x \text{ then } y_j = u_{cj}(\mathbf{x}, v_{cj}^y), \qquad j = 1, 2, \dots, q. \tag{6.1.5}$$

where we use the following nonlinear mapping for each component:

$$u_{cj}(\mathbf{x}, v_{cj}^y) = d_{cj0}.v_{cj}^y + d_{cj1}.x_1 + \dots + d_{cjp}.x_p, \qquad c = 1, 2, \dots, C;\ j = 1, 2, \dots, q \tag{6.1.6}$$

where d_{cji} $(i = 0, 1, \dots, p)$ are the parameters to be identified. Thus the output HR patch pixels are linear combination of centroids v_{cj}^y and the input LR patch feature vector $\mathbf{x}$. Now, u_{cj} is the response of the j^{th} components of c^{th} rule. Then

we combine the C different responses to get a combined response y_j. We use the conventional way [106] of combining the rules as:

$$y_j = \frac{\sum_{c=1}^{C} \alpha_c . u_{cj}(\mathbf{x}, v_{cj}^y)}{\sum_{c=1}^{C} \alpha_c}, \qquad j = 1, 2, \ldots, q. \tag{6.1.7}$$

where, α_c is the firing strength of the c^{th} rule R_c^{TS} computed using a T-norm over p atomic clauses (6.1.4). Since v_{cj}^y is given (a constant), without loss of generality we denote $d_{cj0}.v_{cj}^y$ as d_{cj0} and identify it efficiently along with the other parameters d_{cji} $(i = 1, 2, \ldots, p)$.

6.1.4 Choice of antecedent membership functions

In order to get the firing strength of the rule base, we need to define the membership function of closeness for each of the p atomic clauses: "x_j is CLOSE to v_{cj}^x" and a T-norm to combine those clauses. Different choice of antecedent membership functions such as triangular, trapezoidal and Gaussian are possible, and the choice of membership function affects the result of our algorithm in its own way though not much. Here we use a popular membership function, that is, Gaussian function because it is differentiable everywhere and in this case has only two parameters (v_{cj}, σ_{cj}) which can be initialized using information from the clusters and then tuned further very easily. To be more explicit, "x_j is CLOSE to v_{cj}^x" is modeled by:

$$\pi_{cj}(x_j;\ v_{cj}, \sigma_{cj}) = \exp(-\frac{(x_j - v_{cj})^2}{\sigma_{cj}^2}). \tag{6.1.8}$$

The spread of the membership function σ_{cj} can be initialized by the standard deviation of the j^{th} component of the LR feature vectors in the training data that are included in the c^{th} cluster.

For the intersection of the p clauses to get the firing strength of c^{th} rule, we may choose any of the T-norm from [52]. The most common choice of operator for T-norm is the product or min. As min is not differentiable, we could use a softer version of min that is differentiable. A way of defining the $softmin$ is as

follows:

$$\alpha_c := softmin(\pi_{c1}, \pi_{c2}, \ldots, \pi_{cp}, t) = \left(\frac{\pi_{c1}^t + \pi_{c2}^t + \ldots \pi_{cp}^t}{p} \right)^{\frac{1}{t}}.$$

Note that the $softmin$ approaches to the true minimum as $t \to -\infty$. However, here the issue of computational cost is involved. For a typical 7×7 patch, the number of clauses (i.e., dimension of feature vector) is $4 * 7^2 = 196$. So, for one rule we need to compute $(196 + 1)$ exponentials and there are 100 such rules (for typical choice of $C = 100$). And we need to do this computation for every LR patch during training as well as testing. This results in a very slow SR technique. Hence it is more natural to use computationally less expensive product operator defined as: $\alpha_c = \Pi_{j=1}^p \pi_{cj}$. Then the rule strength combining p clauses with Gaussian membership reduces to

$$\begin{aligned} \alpha_c &= \Pi_{j=1}^p \pi_{cj}(x_j;\ v_{cj}, \sigma_{cj}). \\ &= \Pi_{j=1}^p \exp\left(-\tfrac{(x_j - v_{cj})^2}{\sigma_{cj}^2} \right) \end{aligned}$$

i.e.

$$\alpha_c = \exp\left(-\sum_{j=1}^{p} \frac{(x_j - v_{cj})^2}{\sigma_{cj}^2} \right). \tag{6.1.9}$$

Here instead of computing $4 * n^2 + 1$ exponentials, we need to compute only one exponential and we could not find any noticeable difference in the output. The main concern here is that the product of those numbers (each of them is less than one) or exponential of large negative number results a number close to zero. Then how we can use this fast operator? The answer lies in the defuzzification process of combining the consequents of those C rules. As in Eq. (6.1.7), we are normalizing the rule strength $\mu_c = \sum_{c=1}^{C} \alpha_c$ during the combination of rules, the final output is, thus, computed as a convex combination of the consequents of C rules. Therefore, even if the product is close to zero, due to normalization process it takes a moderate value.

Since there is a strong correlation between a pixel and its neighborhood, the feature vectors associated with a pixel and its neighborhood are also related. In the above settings of the proposed model, the way we have divided each rule R_c^{TS}

in (6.1.4) into p independent clauses and then combined them using a T-norm (6.1.9), we are not exploiting the dependency. This strong dependency can be embedded (exploited) if we use Mahalanobis distance $\sqrt{(\mathbf{x}-\mathbf{v}_c)^T{\Sigma_c}^{-1}(\mathbf{x}-\mathbf{v}_c)}$ in the exponential of (6.1.9) instead of normalized Euclidean distance $\sqrt{\sum_{j=1}^{p}\frac{(x_j-v_{cj})^2}{\sigma_{cj}^2}}$. The parameters $(\mathbf{v}_c, \Sigma_c^{-1})$ could be obtained by clustering the data set into C different clusters as (6.1.2) where $\mathbf{v}_c$ and Σ_c^{-1} are the cluster center and the inverse of variance covariance matrix of the c^{th} cluster. K-means algorithm is widely used for this purpose, but those parameters can be calculated more efficiently if we use Gaussian Mixture Models (GMM) for clustering and the parameter estimation. We do not claim here that GMM would do better clustering than K-means rather the parameters $(\mathbf{v}_c, \Sigma_c^{-1})$ can be computed more efficiently using GMM. Another potential choice would be to use the Gustafson-kessel clustering algorithm [41].

6.1.5 Clustering Using Gaussian Mixture Models (GMM)

In this section we discuss partitioning our training data-set (6.1.1) into C clusters (6.1.2). A Gaussian mixture density is a weighted sum of C component densities given by the equation

$$e(\mathbf{x}^*;\ \lambda) = \sum_{c=1}^{C} \tau_c(\mathbf{x}^*) f_c(\mathbf{x}^*;\ \lambda_c) \tag{6.1.10}$$

where $\mathbf{x}^*$ is a point in the data-set, $f_c(\mathbf{x}^*;\ \lambda_c)$, $c = 1, 2, \ldots, C$ is the density of c^{th} component parameterized by λ_c, $c = 1, 2, \ldots, C$; and $\tau_c(\mathbf{x}^*)$, $c = 1, 2, \ldots, C$ is the mixture weight. Each component density is a $(p+q)$-variate Gaussian function of the form

$$f_c(\mathbf{x}^*;\ \lambda_c) = \frac{1}{(2\pi)^{(p+q)/2}|\Sigma_c^*|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{x}^* - \mathbf{v}_c^*)^T \Sigma_c^{*-1}(\mathbf{x}^* - \mathbf{v}_c^*)\} \tag{6.1.11}$$

with mean vector $\mathbf{v}_c^*$ and the covariance matrix Σ_c^*. The mixture weights $\tau_c(\mathbf{x}^*)$ satisfy the constraint that $0 < \tau_c(\mathbf{x}^*) < 1$ and $\sum_{c=1}^{C} \tau_c(\mathbf{x}^*) = 1$. The Gaussian mixture density is completely parameterized by the mean vectors $\mathbf{v}_c^*$, covariance matrices Σ_c^* and mixture weights $\tau_c(\mathbf{x}^*)$ from all component densities. These

parameters are collectively represented by the notation

$$\lambda = \{\lambda_c | \lambda_c = (\tau_c, \mathbf{v}_c^*, \Sigma_c^*), \qquad c = 1, 2, \ldots, C\} \tag{6.1.12}$$

While clustering using GMM, we eventually seek for the parameters λ that maximize the log-likelihood of the mixture densities (6.1.10) with constraints $0 < \tau_c(\mathbf{x}^*) < 1$ and $\sum_{c=1}^{C} \tau_c(\mathbf{x}^*) = 1$ over all points $\mathbf{x}^*$ in the data-set X^*, i.e.,

$$\begin{aligned} \hat{\lambda} &= \arg\max_\lambda \left[\sum_{\mathbf{x}^* \in X^*} \log(e(\mathbf{x}^*;\ \lambda))\right] \\ &= \arg\max_\lambda \left[\sum_{\mathbf{x}^* \in X^*} \log(\sum_{c=1}^{C} \tau_c f_c(\mathbf{x}^*;\ \lambda_c))\right] \end{aligned} \tag{6.1.13}$$

This optimization is done using Expectation-Maximization algorithm [67]. A nonnegative regularization parameter is added to the diagonal of covariance matrices Σ_c^* to make them positive-definite. Then we extract the components of $(\mathbf{v}_c^*, \Sigma_c^{*-1})$ corresponding to first p dimensions and denote it by $(\mathbf{v}_c, \Sigma_c^{-1})$ which is nothing but the mean and inverse of variance covariance matrix of only LR patch feature vectors of the c^{th} cluster. Then the rule strength (6.1.9) of the rule in (6.1.5) may be rewritten as:

$$\alpha_c = \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{v}_c)^T {\Sigma_c}^{-1} (\mathbf{x} - \mathbf{v}_c)\right). \tag{6.1.14}$$

Note that, $\sqrt{(\mathbf{x} - \mathbf{v}_c)^T {\Sigma_c}^{-1} (\mathbf{x} - \mathbf{v}_c)}$ is the Mahalanobis distance between a point $\mathbf{x}$ and c^{th} cluster center $\mathbf{v}_c$. If the covariance matrix Σ_c is diagonal so is Σ_c^{-1}, then the distance becomes a Euclidean distance and in that case rule strength (6.1.14) is identical with (6.1.9). Now the choice of Mahalanobis distance over Euclidean is crucial because within a patch the pixel values are strongly correlated with each other, so their relation can be modeled better if we consider non-diagonal elements of covariance matrices.

As pointed out by Takagi and Sugeno [106], given a set of rules with fixed antecedents, optimizing the parameters of the consequent equations with respect to training data reduces to a linear lest square estimation. This problem can be solved very easily and the details are described in the following section.

6.1.6 Estimation of Consequent parameters

To estimate the appropriate set of consequent parameters d_{cji}, we formulate the problem as a linear least square optimization problem. Let, $\mu_{ic} = \frac{\alpha_{ic}}{\sum_{c=1}^{C} \alpha_{ic}}$ is the normalized strength of c^{th} rule R_c^{TS} at $\mathbf{x}_i$, accordingly we rewrite (6.1.7) as

$$y_{ij} = \sum_{c=1}^{C} \mu_{ic}.(d_{cj0} + d_{cj1}.x_{i1} + \ldots + d_{cjp}.x_{ip}) \tag{6.1.15}$$

reshuffling the parameters the above equation may also be written in the following form:

$$y_{ij} = \sum_{c=1}^{C} (d_{cj0}.\mu_{ic} + d_{cj1}.x_{i1}\mu_{ic} + \ldots + d_{cjp}.x_{ip}.\mu_{ic}). \tag{6.1.16}$$

We fix the parameters of the antecedent membership functions of the initial rules base system, then estimate the rest of the consequent parameters:

$$D^j = (d_{1j0}\ d_{2j0}\ \ldots d_{Cj0}\ d_{1j1}\ d_{2j1}\ \ldots\ d_{Cj1}\ \ldots\ d_{1jp}\ d_{2jp}\ \ldots\ d_{Cjp})^T$$

using our patch-pair data-set (6.1.16). Since our training data-set X^* consist of a set of N LR-HR patch pairs, $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)^T$ and $Y = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)^T$, (6.1.16) gives us a system of linear equations

$$Y = A * D \tag{6.1.17}$$

where $D = (D^1 D^2 \ldots D^q)_{(p+1)C \times q}$ and

$$A = \begin{pmatrix} \mu_{11} \ldots \mu_{1C} & \mu_{11}.x_{11} \ldots \mu_{1C}.x_{11} & \ldots & \mu_{11}.x_{1p} \ldots \mu_{1C}.x_{1p} \\ \mu_{21} \ldots \mu_{2C} & \mu_{21}.x_{21} \ldots \mu_{2C}.x_{21} & \ldots & \mu_{21}.x_{21} \ldots \mu_{2C}.x_{2p} \\ \ddots & \ddots & & \ddots \\ \mu_{N1} \ldots \mu_{NC} & \mu_{N1}.x_{N1} \ldots \mu_{NC}.x_{N1} & \ldots & \mu_{N1}.x_{N1} \ldots \mu_{NC}.x_{Np} \end{pmatrix} \tag{6.1.18}$$

Here $Y_{N \times q}$ is a matrix of output values and $A_{N \times C(p+1)}$ is a matrix with predefined x_{ij} and corresponding membership values μ_{cj}. The matrix $D_{C(p+1) \times q}$ contains all the parameters to be estimated. These parameters are estimated using least

square estimation (LSE) method, which minimizes the following error

$$\hat{D} = \min_{D} \|Y - AD\|_2^2 \tag{6.1.19}$$

this leads to direct solution $\hat{D} = (A^T A)^{-1} A^T Y$. However, for a moderate data-set, the method may end up with a unstable solution due to large number $[(p+1)\times C \times q]$ of unknown parameters d_{cji}. So we regularize over the solution space by adding a constraint in LSE. In other words, we minimize the following expression instead

$$\hat{D} = \min_{D} \|Y - AD\|_2^2 + \gamma \|D\|_2^2 \tag{6.1.20}$$

which is popularly known as Ridge Regression and gives the solution of the form

$$\hat{D} = \left(A^T A + \gamma I\right)^{-1} A^T Y \tag{6.1.21}$$

where I is an identity matrix and γ is the regularization parameter that controls the smoothness of the solution.

Now if the dimension of $A^T A$ is very large (which is proportional to patch-size and number of cluster), computation of $\left(A^T A + \gamma I\right)^{-1}$ requires a lot of memory and is also time consuming. In that case we may use gradient descent to find solution of the regularized minimization problem (6.1.20). This is what we do and get a quite satisfactory results.

6.1.7 Further tuning of the parameters

In Section 6.1.6, we have estimated the consequent parameters D of the TS model by minimizing the regularized least square error assuming that the antecedent parameters are fixed. We have also discussed about the antecedent parameters $(\mathbf{v}_c, \Sigma_c^{-1})$ of the initial rule base system in Section 6.1.5. We take those as an initial choice for the parameters and then use an optimization method for further refinement all the parameters including the antecedent membership function that minimizes regularized least square error on our training data-set as

$$\{\hat{D},\ \hat{\lambda}\} = \min_{D,\ \lambda} \left[\|Y - AD\|_2^2 + \gamma \|D\|_2^2\right] \tag{6.1.22}$$

where γ is the regularization parameter that gives a stable solution. When the training data are large, even less emphasis to regularization terms may yield a good solution. This hybrid learning can improve the performance of our model because the objective function minimized by the clustering algorithm and the objective function that a rule base should minimize are not same. Here, we use an EM (expectation-maximization) algorithm for finding an optimal antecedent membership parameters $\lambda = \{(\mathbf{v}_c, \Sigma_c^{-1}),\ c = 1, 2, \ldots, C\}$ as well as the optimal consequent parameters D^j. In this hybrid algorithm, first we estimate the antecedent parameters keeping the consequent parameters fixed, and in the next stage estimate the consequent parameters with the modified antecedent parameters. These two steps are iterated until stabilization.

E-step

In this step, it is assumed that consequent parameters $D^{(t)}$ are fixed. Based on that the antecedents parameters λ of the rule base are estimated. Then the optimization problem (6.1.22) becomes

$$\hat{\lambda} = \min_{\lambda} \left[\|Y - AD^{(t)}\|_2^2 \right] \tag{6.1.23}$$

The parameters $\lambda_c = (\mathbf{v}_c, \Sigma_c^{-1})$ are tuned using gradient descent method to minimize the error $E = \|Y - AD^{(t)}\|_2^2$ as

$$\mathbf{v}_c^{(t+1)} = \mathbf{v}_c^{(t)} - \alpha_v \left[\frac{\partial E}{\partial \mathbf{v}_c} \right]_{(t)} \tag{6.1.24}$$

$$\Sigma_c^{-1(t+1)} = \Sigma_c^{-1(t)} - \alpha_{\Sigma^{-1}} \left[\frac{\partial E}{\partial \Sigma_c^{-1}} \right]_{(t)} \tag{6.1.25}$$

where α_v and $\alpha_{\Sigma^{-1}}$ are gradient descent step sizes along v and Σ^{-1} respectively and $[.]_{(t)}$ are the refinements obtained from the gradients by substituting the parameters $\{v_c^{(t)}, \Sigma_c^{-1(t)}, D^{(t)}\}$ after t^{th} iteration. Usually both the step sizes are taken very small, and $\alpha_{\Sigma^{-1}}$ is much smaller than α_v. The gradient are calculated

as follows:

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{v}_c} &= \sum_i (Y_i - A_i D)\, D^T \left[\frac{\partial A_i}{\partial \mu_{ic}}\right]^T \frac{\partial \mu_{ic}}{\partial \mathbf{v}_c} \\ \frac{\partial E}{\partial \Sigma_c^{-1}} &= \sum_i (Y_i - A_i D)\, D^T \left[\frac{\partial A_i}{\partial \mu_{ic}}\right]^T \frac{\partial \mu_{ic}}{\partial \Sigma_c^{-1}} \end{aligned} \tag{6.1.26}$$

$$\begin{aligned} \text{where } \frac{\partial \mu_{ic}}{\partial \mathbf{v}_c} &= \frac{\partial \mu_{ic}}{\partial \alpha_{ic}} \left[\frac{\partial \alpha_{ic}}{\partial \mathbf{v}_c}\right] \\ &= \frac{1}{(\sum_c \alpha_{ic})^2} \left[\frac{\partial \alpha_{ic}}{\partial \mathbf{v}_c}\right] \\ &= \frac{\alpha_{ic}}{(\sum_c \alpha_{ic})^2} \Sigma_c^{-1} (\mathbf{x}_i - \mathbf{v}_c) \end{aligned} \tag{6.1.27}$$

$$\begin{aligned} \text{and } \frac{\partial \mu_{ic}}{\partial \Sigma_c^{-1}} &= \frac{\partial \mu_{ic}}{\partial \alpha_{ic}} \left[\frac{\partial \alpha_{ic}}{\partial \Sigma_c^{-1}}\right] \\ &= \frac{1}{(\sum_c \alpha_{ic})^2} \left[\frac{\partial \alpha_{ic}}{\partial \Sigma_c^{-1}}\right] \\ &= -\frac{\alpha_{ic}}{(\sum_c \alpha_{ic})^2} (\mathbf{x}_i - \mathbf{v}_c)^T (\mathbf{x}_i - \mathbf{v}_c) \hat{I} \end{aligned} \tag{6.1.28}$$

where Y_i and A_i are the i^{th} row of Y and A respectively that corresponds to i^{th} point in the data-set as in (6.1.17) and $\hat{I}$ is the matrix whose all elements are one.

M-step

In this step, the modified antecedent parameters $\lambda_c = (\mathbf{v}_c^{(t+1)},\ \Sigma_c^{-1^{(t+1)}})$ are kept fixed and, the consequent parameters D are estimated. Then the optimization problem (6.1.22) becomes

$$\hat{D} = \min_D \left[\|Y - A^{(t+1)} D\|_2^2 + \gamma_1 \|D\|_2^2\right]. \tag{6.1.29}$$

where the matrix $A^{(t+1)}$ in (6.1.16) is recomputed after modifying the antecedent membership parameters. Consequently, the consequent parameters are updated as:

$$D^{(t+1)} = \left(A^{T^{(t+1)}} A^{(t+1)} + \gamma_1 I\right)^{-1} A^{T^{(t+1)}} Y \tag{6.1.30}$$

where t indicates the iteration number.

We do the above two steps iteratively and the stopping criterion could be either when the global error value (6.1.22) be less than a predefined threshold or when the number of iterations reaches a predefined number. Moreover, such a two stage hybrid scheme for tuning the consequents along with the membership parameters is justified, because when the membership parameters are altered, the

LSE based estimate of the consequents may not remain optimal.

However, in this hybrid scheme, since the initial antecedent memberships are judiciously chosen based on cluster analysis, they are likely to form a reasonably good set of linguistic values. The LSE estimate of consequent parameters keeping the antecedent membership fixed results in a fairly good rule based system.

At the end of this offline phase, we have C clusters, antecedent parameters $(\mathbf{v}_c, \Sigma_c^{-1})$ of the rule based system and the consequent parameters D. These parameters are saved for future up-sampling process.

6.2 Up-scaling of the test document

After tuning the antecedent and consequent parameters, our rule based system is ready to do up-scaling of test images. As our model is trained to estimate HR patch from bicubic interpolated patch (Section 6.1.1), we up-sample the input LR image using bicubic interpolation with resolution factor s. We divide the input image into overlapping patches and extract LR feature vector $\mathbf{x}_i$ corresponding to each overlapping patch as discussed in the Section 6.1.1. During extraction of the patches, we also store the mean and variance of each of the LR patch. It is assumed that a LR and HR patch has the same mean intensity. It has already been observed [121] that the standard deviation of HR patch is approximately 1.2 times the standard deviation of corresponding LR patch. Accordingly, we get mean and standard deviation of the HR patch from that of the LR patch itself. We compute the membership value of each LR patch feature vector $\mathbf{x}_i$ and then compute the normalized rule strength μ_c for each of the rule R_c^{TS}. We use the tuned fuzzy-rule to generate the corresponding normalized HR patch. For each LR patch feature vector, calculate the matrix A as in (6.1.7) and then multiply with the trained antecedent parameters D to get all the HR patches of the corresponding LR one, i.e., $Y = A * D$. The estimated HR patches are again normalized to mean and 1.2 times the variance of the corresponding LR patches. We rearrange each estimated HR patch vector $\mathbf{y}_i$ into $n \times n$ array and place it in the corresponding location of $\mathbf{x}_i$. We simply take the average of the overlapping portions of the estimated HR patches. Thus, the fuzzy-rule based method up-sample a LR image by a fixed resolution factor s. Higher resolution image may

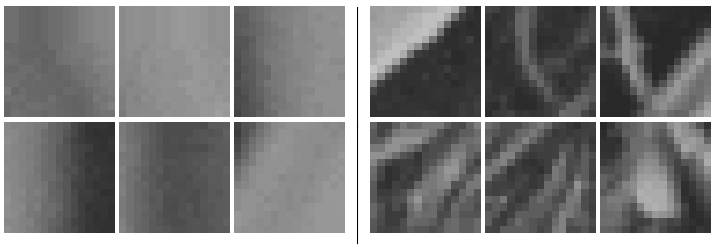


Figure 6.2: The six image patches on the left portions are some of the patches where bicubic interpolation and fuzzy-rule based prediction gives comparable results in terms of RMSE. The six patches on the right portions are some of the patches where fuzzy-rule based prediction gives a definite better results in terms of RMSE.

be generated by repeated application of this process. More detail description of the parameters used in the proposed algorithm are given in Section 6.4.

6.2.1 Selective patch processing

Natural images typically contain large smooth regions as well as strong discontinuities, such as edges and corners. Although simple interpolation methods for image up-scaling, e.g., bilinear or bicubic interpolation, results in noticeable artifacts, e.g. ringing, jaggies and blurring effects along the edges and corners, they perform reasonably well on smooth regions. Figure 6.2 illustrates this fact, where we upscale ($x1.6$) to Lena image, where each image patch of a small image (having ground truth) by (i) bicubic interpolation and (ii) fuzzy-rule based system for root-mean-square error (RMSE) comparison. The patches on the left part shows the patches where the results of both the methods are comparable, while the right part reveal that the fuzzy-rule based SR method outperforms the bicubic interpolation. Similar observation has also been reported by Yang *et al.* [121]. Based on these observations, we may selectively process those highly textured regions or roughness regions using our fuzzy-rule based technique and simply apply cheaper bicubic interpolation on the rest smooth regions. This saves a significant amount of computation. We compute the roughness of a patch by the variance of the LR patch feature vector $\mathbf{x}$. If it is high then the corresponding patch probably belongs to texture region; otherwise to smooth region. We choose a threshold judiciously,

to choose the appropriate class for the patch and consequently the up-scaling method. The detail description of the threshold parameter is mentioned in the experiment section.

We summarize the proposed algorithm in the following section.

6.2.2 Summary of proposed algorithm

Here we concisely present the proposed SR method, as discussed in last section comprising two phases: (1) Offline phase - where a rule base is developed and the parameters are learned from a data-set of LR-HR patch pairs, and (ii) Online phase - where HR image is predicted from a test LR image by the rule based developed in the offline phase. Accordingly, the entire SR process is summarized in Algorithm-3 and Algorithm-4.

Algorithm 3 : Developing a fuzzy-rule based system for single-frame SR

Input: A collection of sharp natural images, number of clusters C, patch size n and resolution factor s.

1. Build a data-set of LR-HR patch pairs from the collection of HR images as discussed in Section 6.1.1.
2. Discard patches with low variance.
3. Do clustering on the data-set of remaining patch pairs using GMM (as stated in Section 6.1.5), which produces C cluster centers $\mathbf{v}_c$ and corresponding variance-covariance matrices Σ_c^{-1} of the clusters.
4. Generate fuzzy-rules corresponding to each of the cluster (as described in Section 6.1.3) and use TS model for HR patch prediction (6.1.7).
5. Estimate consequent parameters D from the initial rule base using the training data-set (see Section 6.1.6).
6. The antecedent and consequent parameters are further tuned using an EM algorithm (as discussed in Section 6.1.7).
7. The fuzzy-rules and corresponding estimated antecedent and consequent parameters are retained for future up-sampling process.

Algorithm 4 : SR using the fuzzy-rule based system

Input: Fuzzy rule based system, the estimated antecedent and consequent parameters, a LR test image and the resolution factor.

1. Divide the test image into overlapping patches and then extract LR patch feature vector for each patch as was done during training data-set generation 6.1.1.
2. Compute and store the mean and variance of each LR patch.
3. Calculate rule firing strength for each LR patch feature vector and then compute rule strength matrix A for all patches (6.1.18).
4. The predicted patches are obtained by multiplying rule strength matrix A with consequent parameters D as in (6.1.17).
5. Adjust the intensity of the predicted HR patch by the mean and 1.2 times the variance of corresponding LR patch.
6. Make an average of all the predicted HR patches in overlapping region of the patches.
7. Apply above $(2) - (6)$ steps repeatedly until the SR image with desired resolution factor is achieved.

6.3 Time comparison

Glasner *et al.* [38] have down-scaled the input LR image to a further coarser scale and then learned the correspondence between the LR and the HR patches. They have also estimated multiple HR frames from a bunch of LR frames and then combined those HR frames to produce the final output. This two-step method requires a huge number of nearest neighborhood patch comparison which is computationally very expensive. Raanan Fattal [34] has extracted some spatial statistical features from the input LR image and estimated corresponding HR edge map from Edge-Frame Continuity Moduli tables, which are obtained by minimizing an error function corresponding to Gauss-Markov Random Field model using conjugate descent optimization techniques. This optimization this system is poorly conditioned, due to weak interactions across edges and strong ones in flat regions, takes a significant amount of time. Sun *et al.* [105] have learned

the relationship between gradient profile sharpness of HR image and LR images. They have embedded gradient field correspondence during the back-projection as the HR image prior. This is somewhat cheaper but still requires a quite significant amount of time. Yang *et al.* [121] have assumed that if a LR patch is a sparse convex combination of some LR patches in a Dictionary then the HR patch may be estimated as the same sparse convex combination of corresponding HR patches in the dictionary. In those cases for each patch they need to estimate corresponding sparse convex combination of LR patches in the dictionary which is also computationally quite expensive.

In contrast to the existing methods mentioned above, the proposed fuzzy-rule based technique is much cheaper as we need to compute mainly the rule strength of each patch and the product of two matrices, and only the latter one requires some computation. This method is significantly faster than the existing methods. The computation time is further reduced by selective patch processing. However, time required in the fuzzy-rule based method is comparable with Freedman *et al.* [36] where they learn the patch correspondence from the image itself by some nearly bi-orthogonal filters without solving the implicit back-projection equations.

6.4 Experimental Results

We have implemented our algorithm using MATLAB 7.6 on a Linux Server with 64GB of RAM and 80 Intel processors of 3.6-GHz. We have used mex code for some of the subroutines to speed up the process. For training (off-line) purpose we have used this machine, however during testing a normal desktop can zoom a sufficiently large image within a reasonable amount of time. It typically takes less than 5 seconds for a typical 200×200 LR image to upscale with resolution factor 4 on an average machine. This quantity is totally dependent on the size of the LR image, patch-size, number of clusters, resolution factor and percentage of smooth patches. In the following section we provide details of the parameters taken for this model-based zooming method.

6.4.1 Detail description of parameters

We have chosen a small up-scaling factor ($s = 1.6$ and $s = 1.45$) in our experiment. HR images with higher up-scaling factors can be generated by applying the algorithm required number of times. For example, for image up-scaling with resolution factor 4 and 3, we need to apply the proposed algorithm 3-times with resolution factor 1.6 and 1.45 respectively. When working with color images, the input image is first converted from RGB to YCbCr. Then the proposed SR algorithm is applied on Y (intensity) channel only and for the rest of the components (i.e., Cb and Cr which are usually low frequency channels and store the chromatic information only) we use bicubic interpolation. The three modified components are then combined and converted back to RGB to get the estimated color SR image displayed.

In patch based techniques, it is observed that usually 5×5 to 10×10 patch sizes are good choice for processing. In our experiment we have chosen $200,000$ patches of size 7×7 and then cluster them into $C = 100$ clusters. Thus the size of each LR feature vector is $4 * 7^2 = 196$ and size of A for test data is $200,000 \times 100 * (196 + 1)$. It is observed that if we increase the number of clusters then the quality of the image gets improved at the cost of increased computation time and memory requirement. For large number of clusters it is difficult to do SR on a limited hardware. We have experimentally seen that use of more than 100 clusters does not improve quality of the images significantly and 100 clusters work well on a regular desktop. It is also observed that, with 100 clusters the proposed fuzzy-rule based system is good enough to produce the results comparable to those of state-of-the-art techniques. For selective patch processing, we use a threshold on the variance of LR image patches in such a way that only 30% of all the patches are predicted using the proposed fuzzy-rule based system and the rest of the patches are considered to be smooth and are zoomed by bicubic interpolation.

In Fig. 6.3, we up-scale a LR old man image with scaling factor 4 with the above parameter settings. We compare our result with some recently reported techniques, due to Fattal *et al.* [34], Glasner *et al.* [38] and Topic modeling introduced in the last chapter. It is clear from the images that the results of the

proposed fuzzy-rule based method produces comparable result with the state-of-the-arts in terms of reduced artifacts, sharp edges and textural details.

We have carried out our experiment with different number of clusters keeping other settings unaltered. The results are shown in the Fig. 6.4. The image quality is improved as we increase the number of clusters. However, the quality does not improve significantly by increasing the number of cluster beyond 100. So we fix $C = 100$ and perform all the remaining experiments.

We also compare the results of the proposed fuzzy-rule based method with that of the other state-of-the-art techniques over different images. The proposed method is found to produce better or equivalent result. Some results are shown in Figs. 6.5 to 6.9. In this experiment we have found that the proposed fuzzy-rule based method is faster than most of the recent algorithms (viz. sparse coding) because in our case we need to compute only some membership values, and the de-fuzzification step requires only multiplication of two matrices (A and D).

6.5 Summary

In this chapter, we have proposed a fuzzy-rule based scheme for single-frame SR. We have used the Takagi-Sugeno (TS) model with the consequents expressed as a linear combination of the input variables. An initial rule base is extracted using a multi-variate cluster analysis based on Gaussian mixture model (GMM). The consequent parameters of these rules are then estimated using LSE technique. The antecedent as well as the consequent parameters of the rule base thus obtained are further refined using an EM type hybrid algorithm. We have tested the proposed scheme over some benchmark images for single-frame SR and obtained better or similar results compared to the state-of-the-art techniques. The proposed method is also faster than most of the competitive SR methods. As far as our knowledge goes, this attempt is the first of this kind of applications using fuzzy-rules in generation of HR images as well as an application in high dimension data. Certainly, there is a huge space to explore for more image processing applications using the proposed approach.

Figure 6.3: Comparison of results of SR techniques for oldman image using the proposed fuzzy-rule based method over some of the state-of-arts methods. Output images along with some cropped portions using different methods are displayed. Top to bottom and left to right images are original image, upsampled image using bicubic interpolation, Fattal *et al.* [34], Glasner *et al.* [38], topic modelling and proposed Fuzzy-Rule based method. It is clear from the images that proposed fuzzy-rule based technique can produce comparable results with the state-of-art methods.

Figure 6.4: Rresults of SR for oldman image with different number of clusters. Top left to bottom right are the results with cluster number 25, 50, 75, 100, 125 and 150 respectively.

Figure 6.5: Comparison results of SR for koala image using proposed method over other state-of-arts methods. Output images along with some cropped portions using different methods are displayed. Top to bottom and left to right images are upsampled image using bicubic interpolation, Perfect Resize (commercial software for image zooming), Glasner *et al.* [38], Freedman *et al.* [36], topic modelling and proposed Fuzzy-Rule based method. It is clear from the images that proposed fuzzy-rule based technique can produce comparable results with the state-of-art methods.

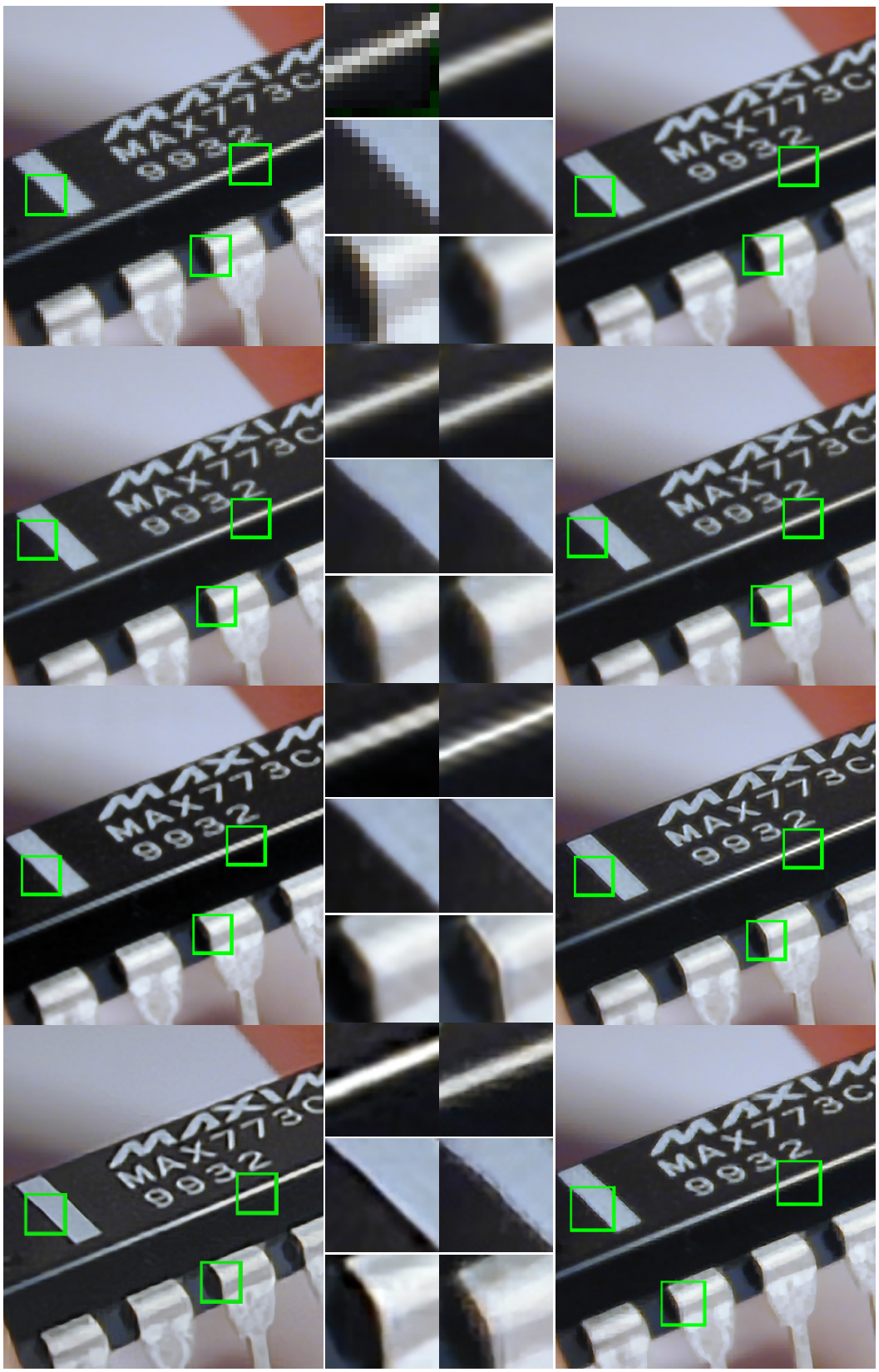


Figure 6.6: Comparison of reconstruction results of chip image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Perfect Resize, Sun *et al.* [105], Fattal *et al.* [34], Glasner *et al.* [38], Topic Modelling and proposed Fuzzy-Rule based method.

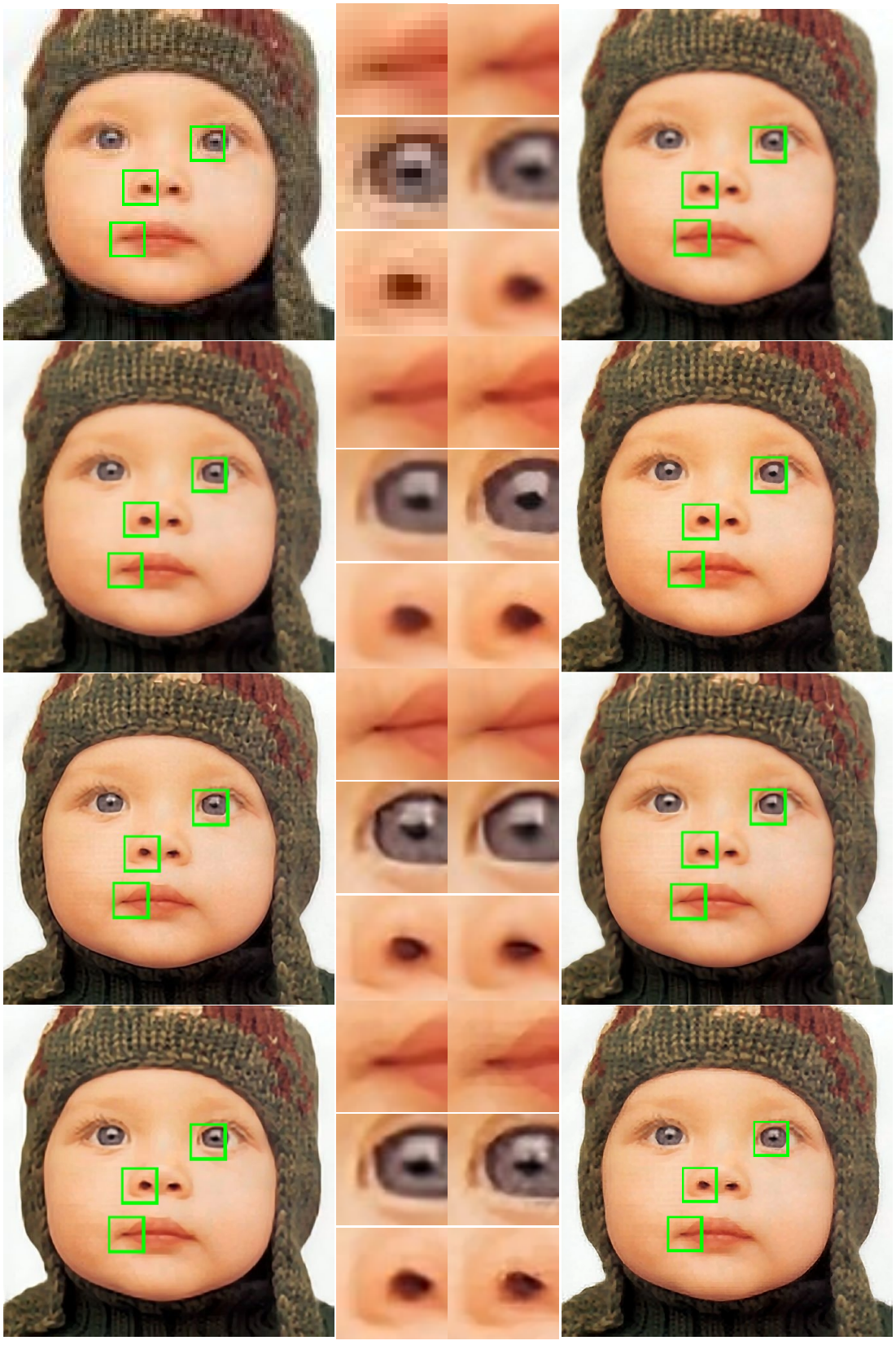


Figure 6.7: Comparison of reconstruction results of child image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Sun *et al.* [105], Fattal *et al.* [34], Glasner *et al.* [38], Freedman *et al.* [36], Topic modelling and proposed Fuzzy-Rule based method.

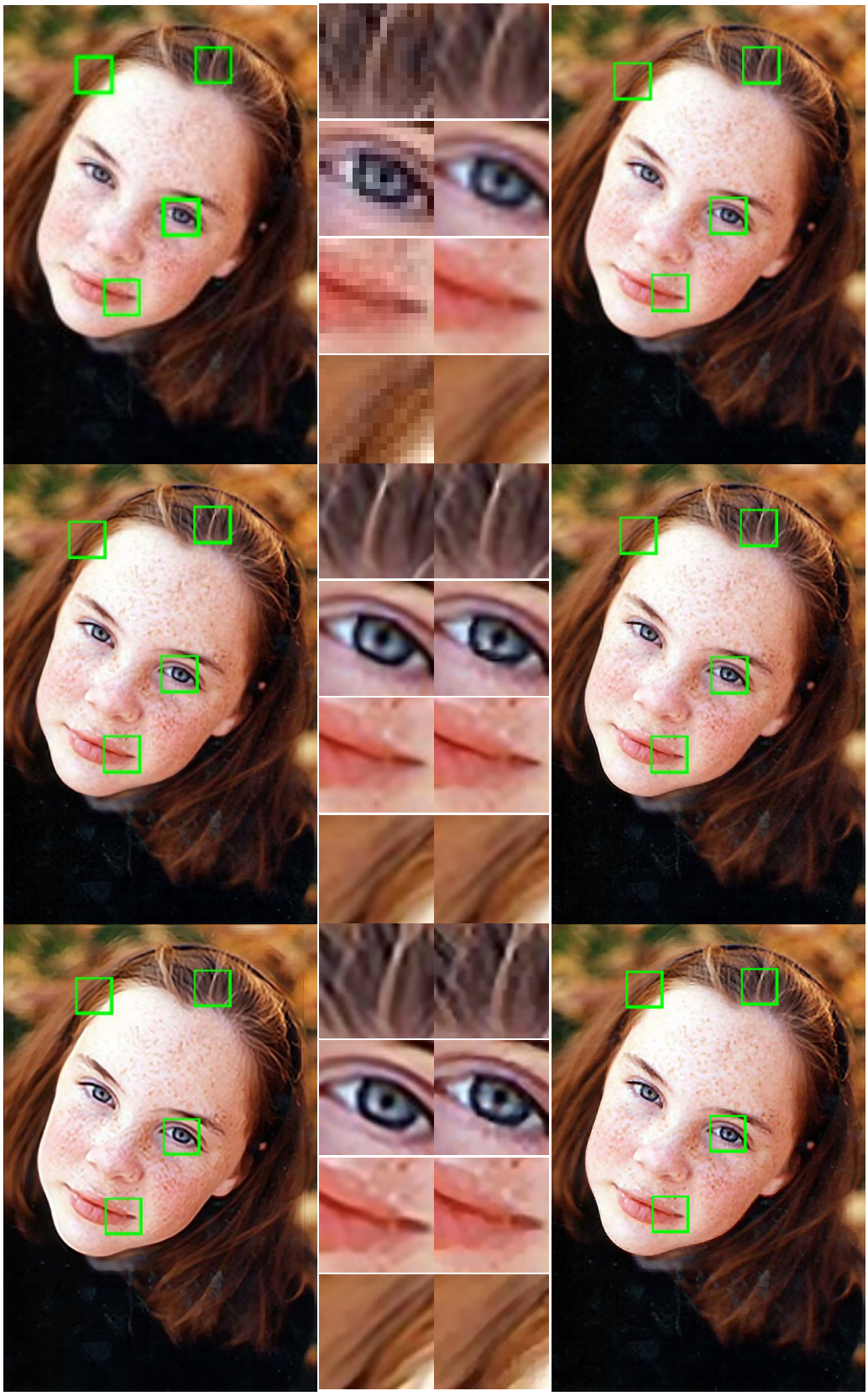


Figure 6.8: Comparison of reconstruction results of girl image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Freedman *et al.* [36], Glasner *et al.* [38], topic modelling and proposed Fuzzy-Rule based method.

Figure 6.9: Comparison of reconstruction results of sclpture image. Top to bottom and left to right images are : Input image (Scaled for display), bicubic interpolation, Freedman *et al.* [36], Glasner *et al.* [38], topic modelling and proposed Fuzzy-Rule based method.

Chapter 7

Conclusions

In this thesis, we have explored different kind of priors/regularization methods and corresponding optimization techniques for multi-frame and single-frame super resolution image reconstruction. In the first three chapter we have proposed some excellent morphologic regularization methods and corresponding efficient optimization techniques for multi-frame SR reconstruction and in the last two chapter we have developed two novel techniques for SR from a single image.

In Chapter 2, we have proposed a gain controlled adaptive regularization for SR image reconstruction. We have devised an iterative algorithm based on algebraic reconstruction technique (ART) for two-dimensional SR image reconstruction and also modify the algorithm for L_1 norm as the dissimilarity measure. This makes the convergence of the algorithm much faster than the conventional iterative methods employed for SR image reconstruction. In Chapter 1, we have formally defined the problem and have discussed different types of state-of-the-art regularization methods that exist in the literature. Most of those regularization term $\Upsilon(X)$ contains $\Gamma\mathbf{X}$ where Γ is an high-frequency operator. An image, in general, may contain both detail and noise at different scales. Since both the noise and edges contribute to high frequency component, they both are usually removed simultaneously during the regularization process enforcing smoothness. Therefore these sharp edges would be removed while trying to remove random noise during the smoothness process using traditional regularizations. In this chapter, we have developed an adaptive regularization technique, where we have assigned a gain value at each pixel depending upon the smoothness of neighbor-

hood of the pixel. During the iterative process of reconstruction, every pixel is gets smoothed depending on its gain value. We have proposed a mathematical morphology based technique for computing the gain map. Thus the edge preserving property of the morphological filters is incorporated in the regularization term leading to sharp edges in the reconstructed image while suppressing noise.

We have presented an edge-preserving SR image reconstruction problem as deblurring problem (3.1.3) with a new robust morphologic regularization method in Chapter 3. Chapter 2 suggest that regularization using multi-scale morphology would be a good choice for regularization, here we have explored that an efficient combination of morphologic operators can replace traditional high-frequency operators Γ in regularization term $\Upsilon(X)$. It is well studied that multi-scale morphological filtering can reduce noise efficiently, so we have built up successfully a regularization method based on multiscale morphology and our experimental section shows that it works quite well, in fact, better than many existing methods. Also, we have employed Bregman iteration method to solve the inverse problem for SR reconstruction with proposed morphologic regularization. Non-linearity of the regularization function is handled in a linear fashion during optimization by means of sub-gradient and proximal map concept. We have also showed that if there are 'impulse noise with random values' or 'salt-and-pepper' noise in LR image, they can be handled efficiently using our two-step SR reconstruction algorithm. It first detects the noisy pixels (note: it does not substitute their values), and then considering those detected pixels as unknown pixels it reconstructs high resolution image using only those pixels which are not corrupted by noise.

We have upgraded this regularization method making it making it adaptive by choosing SE of different shapes and sizes depending on the local statistics of neighboring pixels as discussed in the Chapter 4. We have analyzed different existing regularization terms, and found that they are basically the energy of residual between the actual pixel value and the estimated pixel value. Usually estimate of pixel value are obtained as the weighted average of neighboring pixels, and algorithm differ because of size of the neighborhood and the value of the weights. We find out that estimate would be more representative if weights are calculated based on geodesic distance. Here, we have proposed two geodesic kernel based regularization for multi-frame Super Resolution image reconstruc-

tion. We have analyzed different regularization methods used in this domain and found that they come from the same concept of edge preserving smoother. We have proposed a new robust adaptive geodesic regularization and adaptive morphologic regularization methods that can suppress the noise more efficiently while preserving the edges. The experimental section shows that those adaptive regularization works quite well, in fact better than existing techniques. Above three proposed regularization methods try to impose edge or structure preserving smoothness in the reconstructed high resolution image, and implement this prior by minimizing the gradient energy in case of multi-frame SR. Similar prior may be incorporated other way, i.e., by means of imposing noise free sharp HR patch in place of noisy and blur LR patches. We have used this idea to develop algorithms for single frame SR. The proposed methods are presented in Chapters 5 and 6. In the context of single-frame SR, we propose two novel single-frame SR technique using (i) probabilistic latent semantic analysis (pLSA) and a (ii) fuzzy-rule based prediction framework.

In Chapter 5, we have adopted a model based technique pLSA for single-frame super resolution via sparse representation. Since, a natural image usually consists of smooth regions, textures, sharp boundaries and their combination. We considered them as a different 'topics', and modeled each of the topics separately for prediction mechanism from LR to HR. From a collection of natural images we learned those topics in unsupervised way and developed separate predictors for individual topics. An Expectation Maximization (EM) algorithm is used on the document feature vectors to find out the inherent topic depicted by the documents. On the documents under each of the topic, separate sparse dual dictionaries are learned. The proposed mechanism certainly produce output better than the single sparse dictionary based methods in terms of qualitatively measure. However, it takes slight time for document topic assignment during up-sampling process, results SR method slower than usual sparse reconstruction based technique. In the following chapter 6, we have developed a SR method which is faster than most of the existing SR techniques and also produce better or comparable results in terms of qualitative measure.

We also have developed a fuzzy rule based system for SR as discussed in Chapter 6. Here the underlying assumption is that a small LR patch could by generated

from one of many possible HR patches. We used distinct fuzzy rules for each of these possible patches and combined them using a 'T-norm' to get the estimated HR patch corresponding to a LR one. The fuzzy-rules and the rule parameters are learned from a database of LR and HR patch pairs. Our experimental results show that proposed Single-frame SR methods also performs superior to the existing methods in most of the cases. We have tested the proposed scheme over some benchmark images for single frame super resolution and obtained better or similar results in the respect of qualitatively evaluation compared to that of very recent techniques. We also show experimentally that proposed method is computationally less expensive than the existing ones including the method proposed in Chapter 6.

7.1 Future Scope of the thesis

In this thesis we have proposed some prior based novel regularization methods for multi-frame SR, and have also proposed some learning mechanism for patch-based single-frame SR based on natural image prior. This is an attempt to explore the powerful mathematical morphology tools in the field of ill-posed inverse reconstruction problems. We derive subgradients of those highly non-linear operators and have shown an way to optimize the inverse problems using those regularization methods. The morphologic regularization methods proposed here are tested only on multi-frame SR reconstruction problem. However, one can easily extend these regularizations to other classical ill-posed inverse problems of image processing as well (e.g., deblurring). Though our methods are relatively faster, they are still far away from being real-time. Effort may be directed towards expedite the processing either by efficient implementation or by upgrading the regularization technique.

In the context of single-frame SR, it has huge commercial demand and still is not commercialized in broader way. *Prefect Resize* (an Adobe Photoshop plug-in, http://www.ononesoftware.com/products/perfect-resize/) is one of the commercially available software, which works fine but not upto that mark. Even we compare proposed method with this software and got similar or better re-

sult. Therefore, we surely enhance the progress of model based image zooming research towards publicly available high-quality image zooming software. Moreover, we have uploaded the source code of the proposed method publicly under general GNU public license and hope it would surely help the researchers on this domain. The proposed methods are the learning mechanisms for prior of ill-posed Single-frame SR problem from natural images database. Those methods can be straightway extended to Video SR technique, where either we can apply SR technique on individual frame to get SR video, or we can divide the video into overlapping 3D blocks (instead of 2D patches for a single-frame) and use a similar algorithm to increase the spatial as well as temporal resolution. Modeling the prediction system using nonlinear neuro-fuzzy framework, where the parameters can be learned by using neural networks, could also be another direction of future work. One may also be interested in applying the proposed methods to the other image reconstruction applications, viz. image mosaicing, texture synthesis. As far as our knowledge goes, this attempt is the first of this kind of applications using fuzzy rules in generation of HR images as well as an application in high dimension data. Certainly, there is a huge space to explore other image processing applications using the proposed fuzzy rule-based approach.

Appendix A

1 Subgradients of morphological operators

Let us consider the error $\Upsilon(X)$ due to regularization as mentioned in (3.2.3). Now use the following properties of Morphological dilation and erosion

$$\begin{aligned} D_s(X) &= (X \oplus B) \oplus \text{, s times } \cdots \oplus B \\ &= D(D(\ldots \text{ s times } \ldots D(X)\ldots)) \end{aligned}$$

$$\begin{aligned} E_s(X) &= (X \ominus B) \ominus \text{, s times} \cdots \ominus B \\ &= E(E(\ldots \text{ s times } \ldots E(X)\ldots)) \end{aligned}$$

where we use simply the operator $D(X)$ for morphological dilation instead of $D_1(X)$ and $E(X)$ instead of $E_1(X)$.

Since an analogous chain rule holds, subgradient of the regularizing error $\Upsilon(X)$ described in (3.2.3) can be obtained using the chain rule of subgradient and the above properties of morphological operators as follows :

$$\begin{aligned} \tfrac{\delta}{\delta X} D_s(X) &= \tfrac{\delta}{\delta X}[D(D_{s-1}(X))] \\ &= \tfrac{\delta D_s}{\delta D_{s-1}} \tfrac{\delta D_{s-1}}{\delta D_{s-2}} \cdots \tfrac{\delta D_1(X)}{\delta X} \end{aligned} \tag{1.1}$$

$$\begin{aligned} \tfrac{\delta}{\delta X} E_s(X) &= \tfrac{\delta}{\delta X}[E(E_{s-1}(X))] \\ &= \tfrac{\delta E_s}{\delta E_{s-1}} \tfrac{\delta E_{s-1}}{\delta E_{s-2}} \cdots \tfrac{\delta E_1(X)}{\delta X} \end{aligned} \tag{1.2}$$

Since computing morphological filters [(3.3.22), (3.3.23)] are just computing

maximum and minimum in a neighborhood defined by structuring element sB, as scale s increases, the size of morphological operator sB also increases and as a result computational time for maximum and minimum values also increases. We use the above chain rule (1.1) and (1.2) to reduce computation searching for large neighborhood. So we search in unit scale and perform searching $s-t$ times on t^{th} subgradient of dilated or eroded image. Let us denote $Z^{d_1} = \frac{\delta D}{\delta X}\mathbf{1}$ and $Z^{e_1} = \frac{\delta E}{\delta X}\mathbf{1}$ are column vectors obtained from (3.3.22) and (3.3.23) as follows :

$$\begin{aligned} Z_i^{d_1} &= [\tfrac{\delta D_1}{\delta x_i}, \tfrac{\delta D_1}{\delta x_i}, \dots \tfrac{\delta D_{mn}}{\delta x_i}]\mathbf{1} \\ &= \sharp j \in B_{(i)} \text{ Such that } x_i = \max_{s\in B_{(j)}}\{x_s\} \end{aligned}$$

$$\begin{aligned} Z_i^{e_1} &= [\tfrac{\delta E_1}{\delta x_i}, \tfrac{\delta E_1}{\delta x_i}, \dots \tfrac{\delta E_{mn}}{\delta x_i}]\mathbf{1} \\ &= \sharp j \in B_{(i)} \text{ Such that } x_i = \min_{s\in B_{(j)}}\{x_s\} \end{aligned}$$

Therefore following the above chain rule (1.1) and (1.2) we get :

$$\begin{aligned} \tfrac{\delta}{\delta X}[D_2(X)]\mathbf{1} &= \tfrac{\delta D_2}{\delta D_1} Z^{d_1} \\ \tfrac{\delta}{\delta X}[E_2(X)]\mathbf{1} &= \tfrac{\delta E_2}{\delta E_1} Z^{e_1} \end{aligned}$$

Thus we get recursive expression

$$\begin{aligned} Z^{d_s} := \tfrac{\delta}{\delta X}[D_s(X)]\mathbf{1} &= \tfrac{\delta D_s}{\delta D_{s-1}} Z^{d_{s-1}} \\ &= \tfrac{\delta D(D_{s-1})}{\delta D_{s-1}} Z^{d_{s-1}} \\ Z^{e_s} := \tfrac{\delta}{\delta X}[E_s(X)]\mathbf{1} &= \tfrac{\delta E_s}{\delta E_{s-1}} Z^{e_{s-1}} \\ &= \tfrac{\delta E(E_{s-1})}{\delta E_{s-1}} Z^{e_{s-1}} \end{aligned}$$

Where $\frac{\delta D(D_{s-1})}{\delta D_{s-1}}$ and $\frac{\delta E(E_{s-1})}{\delta E_{s-1}}$ are computed in the same way as in equations (3.3.22) and (3.3.23) with respect to dilation D_{s-1} and erosion E_{s-1} respectively with unit scale.

2 Computational complexity for morphologic regularization term

We use the above method for efficient computation of subgradients of morphological operators. Now computing $\frac{\delta}{\delta X} D_s(X)$ in (1.1) and $\frac{\delta}{\delta X} E_s(X)$ in (1.2) as above would take the same number of operations as derivatives in 'BTV' regularization term (1.10) for $w = s$ and $2/(2s+1)^2$th comparison that 'TV' regularization. Since in 'TV' regularization term we need to compute only first order derivatives and for 'BTV' regularization higher order derivatives for all neighborhood pixels. Now computing $\frac{\delta}{\delta D_s(\mathbf{X})} E_s(D_s(\mathbf{X}))$ and $\frac{\delta}{\delta E_s(\mathbf{X})} D_s(E_s(\mathbf{X}))$ as shown in Eq. (3.3.25) and Eq. (3.3.26) would take twice as much as comparison for computing $\frac{\delta}{\delta X} D_s(X)$ and $\frac{\delta}{\delta X} E_s(X)$. As a result computation of subgradients of our morphological operators $\Upsilon(\mathbf{X})$ as described in Eq. (3.2.3) takes twice as much as computation of derivatives of BTV regularization operator (1.10) and $(2s+1)^2$ as much as computation of derivatives of TV regularization operator (1.9).

References

[1] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, **3**:993–1022, July 2003. 92, 93

[2] Gunilla Borgefors. Distance transformations in digital images. *Comput. Vision Graph. Image Process.*, **34**[3]:344–371, June 1986. 80

[3] N. K. Bose, S. Lertrattanapanich, and M. B. Chappalli. Superresolution with second generation wavelets. *Signal Processing: Image Communication*, **19**[5]:387–391, 2004. 4

[4] V. Boskovitz and H. Guterman. An adaptive neuro-fuzzy system for automatic image segmentation and edge detection. *Fuzzy Systems, IEEE Transactions on*, **10**[2]:247–262, April 2002. 113

[5] L. M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, **7**[3]:200–217, 1967. 49

[6] R. W. Brockett and P. Maragos. Evolution equations for continuous-scale morphological filtering. *Signal Processing, IEEE Transactions on*, **42**[12]:3377–3386, December 1994. 47

[7] L. Brown. A survey of image registration techniques. *ACM Computing Surveys*, **24**[4]:325–376, 1992. 20

[8] J. Cai, S. Osher, and Z. Shen. Linearized Bregman iterations for compressed sensing. *Mathematics of Computational*, **78**[267]:1515–1536, September 2009. 48

[9] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, **52**[2]:489–509, February 2006. 48

[10] Ayan Chakrabarti, A. N. Rajagopalan, and Rama Chellappa. Super-resolution of face images using kernel pca-based prior. *Multimedia, IEEE Transactions on*, **9**[4]:888–892, June 2007. 5

[11] Tony F. Chan and Chiu-Kwong T. Wong. Multichannel image deconvolution by total variation regularization. In *Proceedings of the SPIE*, pages 358–366, San Diego, CA, USA, July 1997. 5, 11, 44, 45, 46, 48

[12] Bhabatosh Chanda and Dwijesh Dutta Majumder. *Digital Image Processing and Analysis*. PHI Learning Pvt. Ltd., New Delhi, 2012. 2, 7

[13] Jyh-Yeong Chang and Jia-Lin Chen. Classifier-augmented median filters for image restoration. *Instrumentation and Measurement, IEEE Transactions on*, **53**[2]:351–356, 2004. 113

[14] T-C. Chang, L. He, and T. Fang. MR image reconstruction from sparse radial samples using Bregman iteration. In *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, May 2005. 48

[15] Subhasis Chaudhuri. *Super-resolution imaging*, **632**. Springer, 2001. 2

[16] Tao Chen and Hong Ren Wu. Adaptive impulse detection using center-weighted median filters. *IEEE Signal Processing Letters*, **8**[1]:1–3, January 2001. 64

[17] F. Cheng and A.N. Venetsanopoulos. Adaptive morphological operators, fast algorithms and their applications. *Pattern Recognition*, **33**[6]:917–933, 2000. 76

[18] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Journal on Multiscale Modeling and Simulation*, **4**[4]:1164–1200, April 2005. 49, 52, 55

[19] Patrick L. Combettes and Jean-Christophe Pesquet. Proximal splitting methods in signal processing. *Technical report (2010), Preprint available at http://arxiv.org/abs/0912.3522*, May 2010. 55

[20] Shengyang Dai, Mei Han, Wei Xu, Ying Wu, and Yihong Gong. Soft edge smoothness prior for alpha channel super resolution. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, June 2007. 5, 6

[21] J. Darbon and S. Osher. Fast discrete optimizations for sparse approximations and deconvolutions. *preprint*, 2007. 48

[22] Ishita De, Bhabatosh Chanda, and Buddhajyoti Chattopadhyay. Enhancing effective depth-of-field by image fusion using mathematical morphology. *Image and Vision Computing, Elsevier*, **24**[12]:1278–1287, April 2006. 40

[23] Johan Debayle and Jean-Charles Pinoli. General adaptive neighborhood image processing. *Journal of Mathematical Imaging and Vision*, **25**[2]:267–284, September 2006. 76

[24] Weisheng Dong, Lei Zhang, Guangming Shi, and Xiaolin Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *Image Processing, IEEE Transactions on*, **20**[7]:533–549, July 2011. 6, 113

[25] Yiqiu Dong, Raymond H. Chan, and Shufang Xu. A detection statistic for random-valued impulse noise. *Image Processing, IEEE Transactions on*, **16**[4]:1112–1120, April 2007. 113

[26] David L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, **52**[4]:1289–1306, April 2004. 46, 48

[27] M. Elad. On the bilateral filter and ways to improve it. *Image Processing, IEEE Transactions on*, **11**[5]:1141–1151, October 2002. 45

[28] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on*, **15**[12]:1013–1027, May 2006. 6, 94

[29] M. Elad and A. Feuer. Restoration of a single super-resolution image from several blurred, noisy and under-sampled measured images. *Image Processing, IEEE Transactions on*, **6**[12]:1646–1658, December 1997. 3, 5, 11, 21, 45, 46, 48

[30] Michael Elad and Dmitry Datsenko. Example-based regularization deployed to super-resolution reconstruction of a single image. *The Computer Journal*, **50**[4]:1–16, April 2007. 6, 88

[31] Michael Elad and Dmitry Datsenko. Example-based regularization deployed to super-resolution reconstruction of a single image. *The Computer Journal*, **52**[2]:15–30, April 2009. 3

[32] Farzam Farbiz and Mohammad Bagher Menhaj. A fuzzy logic control based approach for image filtering. *Studies in Fuzzyness and Soft Computing*, **52**:194–221, 2000. 113

[33] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super-resolution. *Image Processing, IEEE Transactions on*, **13**[10]:1327–1344, October 2004. xiii, xiv, xvi, xvii, xviii, 5, 7, 9, 10, 11, 13, 28, 30, 32, 33, 34, 39, 41, 45, 48, 59, 62, 64, 65, 66, 67, 68, 71, 84, 85, 86

[34] Raanan Fattal. Image upsampling via imposed edge statistics. *ACM Transactions on Graphics*, **26**[3]:56–65, July 2007. xix, xx, 6, 106, 109, 111, 129, 131, 133, 136, 137

[35] Gilad Freedman and Raanan Fattal. Image and video upscaling from local self-examples. *ACM Transactions on Graphics*, **28**[3]:1–10, 2010. 97

[36] Gilad Freedman and Raanan Fattal. Image and video upscaling from local self-examples. *ACM Transactions on Graphics*, **30**[2]:12:1–12:11, April 2011. xix, xx, xxi, 108, 109, 110, 111, 130, 135, 137, 138, 139

[37] W. T. Freeman and T.R. Jones. Example-based super resolution. *IEEE Computer Graphics and Applications*, **22**[2]:56–65, March 2002. 5, 6, 87, 88, 90

[38] Daniel Glasner, Shai Bagon, and Michal Irani. Super-resolution from a single image. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 349–356, October 2009. xix, xx, xxi, 6, 16, 88, 106, 108, 109, 110, 111, 129, 131, 133, 135, 136, 137, 138, 139

[39] Tom Goldstein and Stanley J. Osher. The split Bregman method for l_1-regularized problems. *SIAM Journal on Imaging Sciences*, **2**[2]:323–343, April 2009. 48, 50, 51, 54, 61

[40] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography. *Journal of Theoretical Biology*, **29**[3]:471–482, December 1970. 21, 22

[41] D.E. Gustafson and W.C. Kessel. Fuzzy clustering with a fuzzy covariance matrix. In *IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, **17**, pages 761–766, 1978. 120

[42] Elaine T. Hale, Wotao Yin, and Yin Zhang. Fixed-point continuation for l_1-minimization: Methodology and convergence. *SIAM Journal on Optimization*, **19**[3]:1107–1130, October 2008. 46, 48

[43] He He and Wan-Chi Siu. Single image super-resolution using gaussian process regression. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 449–456, june 2011. 6

[44] Thomas Hofmann. Probabilistic latent semantic indexing. In *Proceedings of ACM SIGIR*, pages 50–57, June 1999. 16, 92, 93

[45] CÃline Hudelot, Jamal Atif, and Isabelle Bloch. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, **159**[15]:1929–1951, 2008. 113, 114

[46] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, **53**[3]:231–239, May 1991. 3, 5

[47] Pekka J. and Toivanen. New geodosic distance transforms for gray-scale images. *Pattern Recognition Letters*, **17**[5]:437–450, May 1996. 80

[48] S. Kaczmarz. Angenäherte auflösung von systemen linearer gleichungen. *Bulletin International de l'Académie Polonaise des Sciences et des Lettres. Classe des Sciences Mathématiques et Naturelles*, **35**[A]:355–357, 1937. 21

[49] A. K. Katsaggelos. *Introduction - Digital Image Restoration*. Springer Series in Information Sciences, Springer, Berlin, 1991. 5, 44, 45

[50] Kwang In Kim and Younghee Kwon. Single-image super-resolution using sparse regression and natural image prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **32**[6]:1127–1133, June 2010. 6, 113

[51] S. P. Kim, N. K. Bose, and H. M. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, **38**[6]:1013–1027, June 1990. 3, 9

[52] George J Klir and Bo Yuan. *Fuzzy sets and fuzzy logic*. Prentice Hall New Jersey, 1995. 118

[53] Sung-Jea Ko and Yong Hoon Lee. Center weighted median filters and their applications to image enhancement. *Circuits and Systems, IEEE Transactions on*, **38**[9]:984–993, September 1991. 64

[54] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Pro-*

ceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pages 2169–2178, 2006. 92

[55] Shie-Jue Lee, Chen-Sen Ouyang, and Shih-Huai Du. A neuro-fuzzy approach for segmentation of human objects in image sequences. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, **33**[3]:420–437, june 2003. 113

[56] Xiaobing Lee, Ya-Qin Zhang, and A. Leon-Garcia. Information loss recovery for block-based image coding techniques-a fuzzy logic approach. *Image Processing, IEEE Transactions on*, **4**[3]:259–273, March 1995. 113

[57] Romain Lerallut, Ãtienne DecenciÃre, and Fernand Meyer. Image filtering using morphological amoebas. *Image and Vision Computing*, **25**[4]:395–404, 2007. 76

[58] S. Lertrattanapanich and N. K. Bose. High resolution image formation from low resolution frames using Delaunay triangulation. *Image Processing, IEEE Transactions on*, **11**[12]:1427–1441, December 2002. 4

[59] Xuelong Li, Yanting Hu, Xinbo Gao, Dacheng Tao, and Beijia Ning. A multi-frame image super-resolution method. *Signal Processing, Elsevier*, **90**[2]:405–414, February 2010. xiii, xiv, xv, xvi, 11, 13, 22, 23, 28, 30, 31, 32, 33, 34, 35, 36, 39, 44, 46, 48, 59, 62, 64, 66, 67, 71

[60] Chin-Teng Lin, Yin-Cheung Lee, and Her-Chang Pu. Satellite sensor image classification using cascaded architecture of neural fuzzy network. *Geoscience and Remote Sensing, IEEE Transactions on*, **38**[2]:1033–1043, March 2000. 113

[61] G. Sapiro J. Mairal and M. Elad. Learning multi-scale sparse representations for image and video restoration. *SIAM Multiscale Modeling and Simulation*, 2008. 6, 94

[62] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online dictionary learning for sparse coding. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2009. 6

[63] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research (JMLR)*, **11**:19–60, 2010. 6

[64] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *Image Processing, IEEE Transactions on*, **17**[1]:53–69, 2008. 6

[65] S. Mallat and Guoshen Yu. Super-resolution with sparse mixing estimators. *Image Processing, IEEE Transactions on*, **19**[11]:2889–2900, 2010. 113

[66] Antonio Marquina and Stanley J. Osher. Image super-resolution by TV-regularization and Bregman iteration. *Journal of Sciences Computational*, **37**[3]:367–382, December 2008. xvi, 39, 40, 46, 49, 50, 59, 62, 65, 66, 67, 68, 71

[67] Geoffrey McLachlan and David Peel. *Finite mixture models*. Wiley-Interscience, 2004. 121

[68] Peyman Milanfar. Super Resolution Dataset. http://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html, 2007. [Online; accessed 2007]. 31, 69

[69] Peyman Milanfar. *Super-resolution imaging*, **1**. CRC Press, 2010. 2, 7

[70] J. J. Moreau. Functions convexes duales et points proximaux dans un espace hilbertien. *C. R. Acad. Sciences Paris Ser. A. Math.*, **255**:2897–2899, 1962. 52

[71] Susanta Mukhopadhyay and Bhabatosh Chanda. An edge preserving noise smoothing technique using multiscale morphology. *Signal Processing*, **82**[4]:527–544, April 2002. 40, 47

[72] Susanta Mukhopadhyay and Bhabatosh Chanda. Multiscale morphological segmentation of gray-scale image. *Image Processing, IEEE Transactions on*, **12**[5]:533–549, May 2003. 40, 47

[73] Makoto Nakashizuka, Yu Ashihara, and Youji Iiguni. Morphological image regularization with a smoothness criterion of structuring elements. In *Proceedings of the International Symposium on Communications and Information Technologies (ISCIT)*, pages 137–142, October 2010. 40

[74] Igor Nedeljkovic. Image classification based on fuzzy logic. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **34**[Part XXX], 2004. 113

[75] N. Nguyen and P. Milanfar. An efficient wavelet-based algorithm for image super resolution. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, **2**, pages 351–354, Vancouver, BC, Canada, September 2000. 4

[76] N. Nguyen, P. Milanfar, and G. Golub. Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. *Image Processing, IEEE Transactions on*, **10**[9]:1299–1308, September 2001. 4, 11, 44, 46

[77] Nhat Nguyen, P. Milanfar, and G. H. Golub. A computationally efficient image super resolution algorithm. *Image Processing, IEEE Transactions on*, **10**[4]:573–583, April 2001. 11, 44, 46

[78] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin. An iterative regularization method for total variation based image restoration. *SIAM Journal on Multiscale Modeling and Simulation*, **4**[2]:460–489, April 2005. 46, 48, 49, 50, 52, 61

[79] Stanley Osher, Yu Mao, Bin Dong, and Wotao Yin. Fast linearized Bregman iteration for compressive sensing and sparse denoising. *Communications in Mathematical Sciences*, **8**[1]:93–111, January 2010. 48

[80] N.R. Pal, V.K. Eluri, and G.K. Mandal. Fuzzy logic approaches to structure preserving dimensionality reduction. *Fuzzy Systems, IEEE Transactions on*, **10**[3]:277–286, June 2002. 114

[81] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, **3**[20]:21–36, March 2003. 3

[82] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, **20**[3]:21–36, May 2003. 21

[83] A. J. Patti and Y. Altunbasak. Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *Image Processing, IEEE Transactions on*, **10**[1]:179–186, January 2001. 4

[84] E. Le Pennec and S. Mallat. Sparse geometric image representations with bandelets. *Image Processing, IEEE Transactions on*, **14**[4]:423–438, April 2005. 6

[85] D.L. Pham and J.L. Prince. Adaptive fuzzy segmentation of magnetic resonance images. *Medical Imaging, IEEE Transactions on*, **18**[9]:737–752, September 1999. 113

[86] Gemma Piella. A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion, Elsevier*, **4**[4]:259–280, April 2003. 40

[87] Matan Protter, Michael Elad, Hiroyuki Takeda, and Peyman Milanfar. Generalizing the nonlocal-means to super-resolution reconstruction. *Image Processing, IEEE Transactions on*, **18**[1]:36–51, January 2009. 14

[88] Nathan D. Ratli, J. Andrew Bagnell, and Martin A. Zinkevich. Subgradient methods for maximum margin structured learning. In

Workshop on Learning in Structured Output Spaces. ICML 2006, June 2006. 55

[89] S. H. Rhee and M. K. Kang. Discrete cosine transform based regularized high-resolution image reconstruction algorithm. *Optical Engineering, SPIE*, **8**[38]:1348–1356, August 1999. 3, 4

[90] P. Rodriguez and B. Wohlberg. Efficient minimization method for a generalized total variation functional. *Image Processing, IEEE Transactions on*, **18**[2]:322–332, February 2009. 46

[91] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, **60**[1-4]:259–268, November 1992. xvii, xviii, 5, 11, 39, 59, 62, 66, 71, 83, 84, 85, 86

[92] Fabrizio Russo. Fire operators for image processing. *Fuzzy Sets and Systems*, **103**[2]:265–275, 1999. Soft Computing for Pattern Recognition. 113

[93] S. Schulte, M. Nachtegael, V. De Witte, D. Van der Weken, and E.E. Kerre. A fuzzy impulse noise detection and reduction method. *Image Processing, IEEE Transactions on*, **15**[5]:1153–1162, May 2006. 113

[94] Stefan Schulte, Valarie De Witte, Mike Nachtegael, Dietrich Van der Weken, and Etienne E. Kerre. Fuzzy random impulse noise reduction method. *Fuzzy Sets and Systems*, **158**[3]:270–283, 2007. 113

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

[96] J. A. Sethian. Fast marching methods. *SIAM Review*, **41**[2]:199–235, January 1999. 80

[97] Simon Setzer. Operator splittings, Bregman methods and frame shrinkage in image processing. *International Journal of Computer Vision*, **92**[3]:265–280, July 2011. 55

[98] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **27**[4]:531–545, April 2005. 5

[99] H. Shen, L. Zhang, B. Huang, and P. Li. A MAP approach for joint motion estimation segmentation and super resolution. *Image Processing, IEEE Transactions on*, **16**[2]:479–490, February 2007. 5, 45, 48

[100] Josef Sivic, Bryan C. Russell, Alexei A. Efros, Andrew Zisserman, and William T. Freeman. Discovering objects and their location in images. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 3070–3077, 2005. 92

[101] Pierre Soille. Generalized geodesy via geodesic time. *Pattern Recognition Letters*, **15**[12]:1235–1240, 1994. 76

[102] Pierre Soille. *Morphological image analysis - principles and applications*. Engineering Online Library, Springer, Berlin, 2003. 25, 40

[103] H. Stark and P. Oskoui. High resolution image recovery from image plane arrays, using convex projections. *Journal of the Optical Society of America A*, **6**[11]:1715–1726, November 1989. 4, 9

[104] D. Stathakis and A. Vasilakos. Comparison of computational intelligence based classification techniques for remotely sensed optical image classification. *Geoscience and Remote Sensing, IEEE Transactions on*, **44**[8]:2305–2318, August 2006. 113

[105] Jian Sun, Zongben Xu, and Heung-Yeung Shum. Image super-resolution using gradient profile prior. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, June 2008. xix, xx, 5, 6, 97, 106, 109, 129, 136, 137

[106] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*, **15**[1]:116–132, January-February1985. 117, 118, 121

[107] Hiroyuki Takeda, Peyman Milanfar, Matan Protter, and Michael Elad. Super-resolution without explicit subpixel motion estimation. *Image Processing, IEEE Transactions on*, **18**[09]:1958–1975, September 2009. 14

[108] A. M. Tekalp, M. K. Ozkan, and M. I. Sezan. High resolution image re-construction from lower-resolution image sequences and space-varying image restoration. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 169–172, San Francisco, CA, March 1992. Speech and Signal Processing. 3, 4

[109] J. Tian, L. Chen, L. Ma, and W. Yu. Phase-driven spatially-variant regularization for image resolution enhancement. *International Journal of Electronics and Communications*, **21**[11]:975–977, November 2011. xiii, xiv, xv, 28, 30, 31, 32, 33, 34, 35, 36

[110] B. C. Tom and A. K. Katsaggelos. Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, **2**, pages 539–542, October 1995. 4, 9, 41

[111] B. C. Tom, A. K. Katsaggelos, and N. P. Galatsanos. Reconstruction of a high resolution image from registration and restoration of low resolution images. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, **3**, pages 553–557, November 1994. 4, 9, 41

[112] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 836–846, New Delhi, India, January 1998. 76

[113] Roger Y. Tsai and Thomas S. Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **6**[1]:13–27, January 1984. 3

[114] Patrick Vandewalle, Sabine Susstrunk, and Martin Vetterli. A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP Journal on Applied Signal Processing (special issue on Super-resolution)*, **2006**[11]:233–233, January 2006. 36, 70

[115] D. Van De Ville, M. Nachtegael, D. Van der Weken, E.E. Kerre, W. Philips, and I. Lemahieu. Noise reduction by fuzzy image filtering. *Fuzzy Systems, IEEE Transactions on*, **11**[4]:429–436, August 2003. 113

[116] Luc Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *Image Processing, IEEE Transactions on*, **2**[2]:176–201, April 1993. 40

[117] N. A. Woods, N. P. Galatsanos, and A. K Katsaggelos. Em-based simultaneous registration and restoration, and interpolation of super-resolved images. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pages II–303–306, September 2006. 21

[118] Liang Xiao and Zhihui Wei. A super-resolution reconstruction via local and contextual information driven partial differential equations. In *Proceedings of Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, pages 726–730, Haikou, China, August 2007. 5

[119] Haixiang Xu, Guangxi Zhu, Fuyuan Peng, and Desheng Wang. Adaptive fuzzy switching filter for images corrupted by impulse noise. In *International Conference on Communications, Circuits and Systems (ICCCAS)*, **2**, pages 792–795, 2004. 113

[120] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang. Coupled dictionary training for image super resolution. *Image Processing, IEEE Transactions on*, **21**[8]:233–233, January 2012. 6

[121] Jianchao Yang, John Wright, Thomas S. Huang, and Yi Ma. Image super-resolution via sparse representation. *Image Processing, IEEE Transactions on*, **19**[11]:2861–2873, November 2010. xviii, xix, 6, 16, 88, 90, 94, 97, 105, 106, 107, 126, 127, 130

[122] Liron Yatziv, Alberto Bartesaghi, and Guillermo Sapiro. O(n) implementation of the fast marching algorithm. *Journal of Computational Physics*, **212**[2]:393–399, February 2006. 80

[123] Wotao Yin, Stanley Osher, Donald Goldfarb, and Jerom Darbon. Bregman iterative algorithms for l_1-minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, **1**[1]:143–168, March 2008. 48, 51

[124] M Emin Yüksel and Alper Başturk. Efficient removal of impulse noise from highly corrupted digital images by a simple neuro-fuzzy operator. *AEU - International Journal of Electronics and Communications*, **57**[3]:214–219, 2003. 114

[125] Deng Ze-Feng, Yin Zhou-Ping, and Xiong You-Lun. High probability impulse noise-removing algorithm based on mathematical morphology. *Signal Processing, IEEE Transactions on*, **14**[1]:31–34, January 2007. 47

[126] Xiaoqun Zhang, Martin Burger, Xavier Bresson, and Stanley Osher. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM Journal on Imaging Sciences*, **3**[3]:226–252, July 2010. 53

[127] Xin Zhang, Edmund Y. Lam, Ed X. Wu, and Kenneth K. Wong. Application of Tikhonov regularization to super-resolution reconstruction of brain MRI image. *Medical Imaging Informatics*, **49**[87]:51–56, 2008. 11, 44, 45, 46

[128] B. Zitov and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, **21**[11]:977–1000, 2003. 20

Publications related to this Thesis

Journal Publications:

J1. Purkait, P., Chanda, B., **Super resolution image reconstruction through bregman iteration using morphologic regularization**, *Image Processing, IEEE Transactions on (TIP)*, 21,9,4029-4039,2012, IEEE

J2. Purkait, P., Chanda, B., **Morphologic gain-controlled regularization for edge-preserving super-resolution image reconstruction**, *Signal, Image and Video Processing (SIVP)*, 1-14, 2011, Springer

J3. Purkait, P., Pal, N.R., Chanda, B., **A Fuzzy-Rule Based Approach for Single Frame Super Resolution.**, *Image Processing, IEEE Transactions on (TIP), (Under Review).*

Conference Publications:

C1. Purkait, P., Chanda, B., **Fuzzy-Rule Based Approach for Single Frame Super Resolution.**, *In Proceedings of the IEEE International Conference on Fuzzy Systems (Fuzz-IEEE '13)*, Hyderabad, India, July 2013. (To appear)

C2. Purkait, P., Chanda, B., **Adaptive Morphologic Regularizations for inverse problems**, *In Proceedings of the* 11^{th} *International Symposium on Mathematical Morphology (ISMM '13)*, Uppsala, Sweden, May 2013.

C3. Purkait, P., Chanda, B., **Image upscaling using multiple dictionaries of natural image patches**, *In Proceedings of the* 11^{th} *Asian Conference on Computer Vision (ACCV '12)*, Daejeon, South Korea, November 2012.

Other Publications:

C1. Purkait, P., Chanda, B., **Digital restoration of damaged mural images**, *In Proceedings of the* 8^{th} *Indian Conference on Computer Vision, Graphics*

and Image Processing (ICVGIP '12), Mumbai, India, December 2012.

C2. Purkait, P., Chanda, B., **A Fast and Robust Deblurring Technique on High noise Environment.**, *In Proceedings of the IEEE international conference on image processing (ICIP '13)* Melbourne Australia, September 2013. (To appear)

C3. Samanta, S., Purkait, P., Chanda, B., **Indian Classical Dance Classification by Learning Dance Pose Bases**, *In Proceedings of IEEE Workshop on Applications of Computer Vision (WACV '12)*, Breckenridge, CO, USA, January 2012.

C4. Atal, K., Arora, A., Purkait, P., Chanda, B., **Face Image Retrieval Based on Probe Sketch Using SIFT Feature Descriptors**, *In Proceedings of Perception and Machine Intelligence (PerMin '12)*, Kolkata, India, January 2012.

C5. Purkait, P., Chanda, B., Kulkarni, S., **A Novel Technique for Sketch to Photo Synthesis**, *In Proceedings of the* 7^{th} *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '10)*, Chennai, India, December 2010.

C6. Purkait, P., Chanda, B., **Off-line Recognition of Hand-written Bengali Numerals Using Morphological Features**, *In Proceedings of the* 12^{th} *International Conference on Frontiers in Handwriting Recognition (ICFHR '10)*, Kolkata, India, November 2010.

C7. Purkait, P., Kumar, R., Chanda, B., **Writer Identification for Telegu Documents Using Directional Morphological Features**, *In Proceedings of the* 12^{th} *International Conference on Frontiers in Handwriting Recognition (ICFHR '10)*, Kolkata, India, November 2010.

C8. Purkait, P., Ghorai, M., Chanda, B., **A fast Box-based Video Inpainting Techniques.**, *In Proceedings of the* 5^{th} *International Conference on Pattern Recognition and Machine Intelligence, (PREMI'13)* Kolkata, September 2013. (Submitted)