Ph.D Thesis

DEVELOPMENT OF DIRECTIONALLY ADAPTIVE TECHNIQUES FOR SINGLE IMAGE SUPER-RESOLUTION

Submitted to the
Cochin University of Science and Technology
in partial fulfillment of the requirements
for the award of the Degree of
Doctor of Philosophy
Under the Faculty of Technology

Reji A.P

under the supervision of

Prof. Tessamma Thomas

Department of Electronics
Cochin University of Science and Technology
Cochin 22
April 2014

Development of Directionally Adaptive Techniques for Single Image Super-Resolution

Ph.D. Thesis in the field of Image Processing

Author

Reji A.P
Research Scholar
Department of Electronics
Cochin University of Science and Technology
Cochin -682 022
Kerala, India.
e-mail:rejimoodethu@yahoo.com

Research Advisor

Dr.Tessamma Thomas
Professor
Department of Electronics
Cochin University of Science and Technology
Cochin -682 022
Kerala, India.
e-mail:tess@cusat.ac.in

CERTIFICATE

Certified that this thesis entitled *Development of Directionally Adaptive Techniques for Single Image Super-Resolution* is a bonafide record of the research work carried by Ms.Reji A.P under my supervision in Department of Electronics, Cochin University of Science and Technology. The results presented in this thesis or part of it has not been presented for the award of any other degree.

Cochin-22
23-04-14

Prof.Tessamma Thomas
Department of Electronics
Cochin University of science and Technology

DECLARATION

I hereby declare that the work presented in the thesis entitled *Development of Directionally Adaptive Techniques for Single Image Super-Resolution* is a bonafide record of the research work done by me under the supervision of Prof.Tessamma Thomas, Department of Electronics, Cochin University of Science and Technology, India and that no part thereof has been presented for the award of any other degree.

Reji A.P
Research Scholar
Department of Electronics
Cochin University of Science and Technology

ACKNOWLEDGEMENT

I bow before THE ALMIGHTY for all the blessings. Only with his blessings, I could bring all my efforts to a successful completion.

I deem it a rare privilege to extend my deep sense of gratitude and thankfulness to Prof.Tessamma Thomas for her inspiring guidance, encouragement and support throughout the course of this work. I consider it as a great boon to have Prof.Tessamma Thomas, an eminent academician of international reputation as my research supervisor. I also extend my sincere thanks to Prof. C.K Aanandan, Head of the Department of Electronics, Cochin University of Science and Technology, Prof. K Vasudevan, Dean of the Department, Prof.P. R.S Pillai for extending the facilities in the Department of Electronics for my research work. I thank from the bottom of my heart Prof.P.Mohanan, Dr.James Kurian, Dr. Supriya M.H for their support and encouragement through all my years in the department.

Let me place on record my sincere gratitude to Prof. Deepu Rajan, School of Computer Engineering, Nanyang Technological University, Singapore for his timely guidance and constant interest in the progress of this work. I wish to make a special mention for the help and encouragement from Prof. Narayanan Nampoothiri, Department of Mathematics, Cochin University of Science and Technology.

I thank all the non-teaching staff and technical staff in the Department of Electronics and the staff at the Administrative Section of CUSAT for their cooperation. I take this opportunity to thank, Mr.Anu Shabarish, Mr.Praveen.N, Mr.Nobert Thomas, Ms.Deepa.J, Ms.Ananada Resmi, Nishamol M.S, Dinesh R, Prajas John, Ullas G.K and Lindo A.O in the Department of Electronics for all their support.

I wish to record my sincere thanks to Dr. Ajitha R.S, Assistant professor , Department of Computer Applications, N.S.S College Rajakumari, for providing technical support with LaTeX, and to Ms Veena M.S, Assistant professor, Department of English, N.S.S College Rajakumari, for correcting the proof. I express my sincere thanks to Mr. Arun S Nair, Lab Technician, N.S.S college Rajakumari for being at the beck and call to sort out all my hardware and software problems. I sincerely express my gratitude to the Secretary N. S. S Colleges' Central Committee, Changanachery for granting me study leave to complete this doctoral work. I am thankful to faculty members and staff, N.S.S College, Rajakumary, for the timely support and encouragement given to me.

At a very personal level, I am most grateful to my dear family members for their patience and encouragement and friends for their support during the period of the study.

Reji A.P

ABSTRACT

Super Resolution problem is an inverse problem and refers to the process of producing a High resolution (HR) image, making use of one or more Low Resolution (LR) observations. It includes up sampling the image, thereby, increasing the maximum spatial frequency and removing degradations that arise during the image capture namely aliasing and blurring.

The work presented in this thesis is based on learning based single image super-resolution. In learning based super-resolution algorithms, a training set or database of available HR images are used to construct the HR image of an image captured using a LR camera. In the training set, images are stored as patches or coefficients of feature representations like wavelet transform, DCT, etc. Single frame image super-resolution can be used in applications where database of HR images are available. The advantage of this method is that by skilfully creating a database of suitable training images, one can improve the quality of the super-resolved image.

A new super resolution method based on wavelet transform is developed and it is better than conventional wavelet transform based methods and standard interpolation methods. Super-resolution techniques based on skewed anisotropic transform called directionlet transform are developed to convert a low resolution image which is of small size into a high resolution image of large size. Super-resolution algorithm not only increases the size, but also reduces the degradations occurred during the process of capturing image. This method outperforms the standard interpolation methods and the wavelet methods, both visually and in terms of SNR values. Artifacts like aliasing and ringing effects are also eliminated in this method. The super-resolution methods are implemented using, both critically sampled and over sampled directionlets. The conventional directionlet transform is computationally complex. Hence lifting scheme is used for implementation of directionlets. The new single image super-resolution method based on lifting scheme reduces computational complexity and thereby reduces computation time. The quality of the super resolved image depends on the type of wavelet basis used. A study is conducted to find the effect of different wavelets on the single image super-resolution method. Finally this new method implemented on grey images is extended to colour images and noisy images.

Contents

List of Tables

List of Figures

List of Abbreviation

CCD = Charge Coupled Device.
DVM = Directional Vanishing Moment.
DWT = Discrete Wavelet Transdorm.
HR = High Resolution.
IDWT = Inverse Discrete Wavelet Transdorm.
LR = Low Resolution.
MAD = Minimum Absolute Difference.
PDF = Probability Distribution Function.
S-AWT = Skewed Anisotropic Wavelet Transform.
SNR = Signal to Noise Ratio.
SR = Super-Resolution.

Chapter 1

Introduction

This chapter gives a brief introduction to the topic of research work undertaken. The chapter presents basic concepts of images, resolution, super resolution etc. It also explains why super resolution methods are needed in different image processing applications. Finally, objective of the presented work and a brief layout of the thesis are given.

1.1 Introduction

From time immemorial, human beings have always been interested in recording events. The history of transformation of these records ranges from primitive cave paintings to latest digital images and videos. Ibn al-Haytham had developed the earliest imaging system as a crude pinhole projection, over thousand of years ago[67]. The next eight hundred years witnessed development of imaging systems into increasingly sophisticated instruments. During the period 1826-27 the first permanent photograph was developed by Joseph Nicephore and the first color photograph was made by James Clerk Maxwell in 1861 [4],[101]. Highly sophisticated imaging systems such as thermal imaging and magnetic resonance imaging have been developed to capture and process visual information in entirely new domains later.

Images are of two types : analog and digital. Analog images are the type of images that human beings look at. They include things such as photographs, paintings, TV images, and all medical images recorded on film. An analog image contains various levels of brightness and colors. They are continuous in values.

When considering the case of technique and style, there is no difference between digital photography and film/analog photography. The only difference is that traditional film is replaced by a charged coupled device (CCD), which contains tiny grids containing millions of photosensitive elements. One can think of an image as a function, f(x,y). The function, f(x,y) gives the intensity at position (x,y) (spatial values) and is proportional to the brightness or gray value of the image at that point.

In digital technology, (a) Spatial and gray scale values have been made discrete. (b) Intensities are measured across a regularly spaced grid in x and y directions and each intensity data is represented by a finite number of bits. For example, the intensity values can be sampled to 8 bits (256 values) per point for black and white and 3x8 bits per point for color images. Examples of digital images include digital

photos, image sequences used for video broadcasting and playback, multi-sensor data like satellite images in the visible, infrared and microwave bands, medical images like ultra-sound, Gamma-ray images, X-ray images and radio-band images like MRI etc, astronomical images, electron microscope images used to study material structure etc.

The main difference between an analog and a digital image is the nature of the signal actually used to produce the image. The visible spectrum of human eye is limited to wavelengths between 400 to 700 nanometers, whereas digital images can extend this limit outside the visible spectrum, which includes microwave, infrared, ultraviolet, or even X-ray regions of the spectrum. In this way, digital images extend the capabilities of human vision. For example, ultrasound imaging produces images based on response (echoes) to sound frequencies, and a MRI scan creates a digital image based on the responses of molecules to magnetic fields.

The main advantage or applications of digital image over the analog counterpart is that it can be stored in computer memory and subsequently used for a variety of applications and can be modified by computer programs. Human vision is analog. It blends image elements smoothly. For example, during sunset, human eye captures images of the red horizon which merges with the deep blue of the darkening sky. One can see objects that are large as well as details that are quite small and differentiate subtle variations in light value, hue and saturation. When a real world is captured and stored in a computer, this continuous range of detail, intensities, and colors are converted to discrete signal values which are of limited range. The human eye, can distinguish a very large number of subtly different shades of red. But in a digital image there might be only a few hundred, or at most a few thousand, distinguishable shades of red.

An image consists of local image features that can be used to model a complete image. There are three main image features: Smooth regions, Edges, and Textures. Smooth regions comprise the largest portion of most images. The simplest model

for this region is to assign a random variable with low variances to model the intensity level locally. Edges represent the abrupt transitions between smooth regions and constitute the smallest region of most images. While they consist of smallest area in an image, they have the most information in an image. A simple model for edges is to assign a random variable with high variance to the gray level value. Textures have a noise-like appearance. But they are distinct from noise as there is a pattern within them and they have self-similarity [49].

1.2 Image Processing

Digital image processing is the processing of information for which the input signal is an image; using standard tools of signal processing techniques. Here the output is not necessarily an image. In image processing, the input image is treated as a 2-dimensional signal. The basic elements of image processing system are (1) image acquisition module such as camera, scanner (2) processing and storage module (3) display module: monitor, printer. Image acquisition model is shown in Figure1.1.

The digital imaging system is not perfect due to hardware limitations, and the acquiring of images is affected with various kinds of degradations. One such degradation is motion blur, which is very common in videos, is due to the finite aperture time. The finite sensor size can cause sensor blur; the image pixel is generated by integration over the sensor area instead of impulse sampling. The limited sensor density leads to aliasing effects, limiting the spatial resolution of the captured image. These signals may be contaminated by atmospheric turbulence before reaching the imaging system. Thus captured images are further affected by the sensor noise and color filtering noise. Finally, the frame captured by imaging system are blurred, decimated, and noisy versions of the underlying true scene. The captured images are called Low Resolution (LR) images. These degradations need to be modeled fully or partially when low resolution image is analyzed.

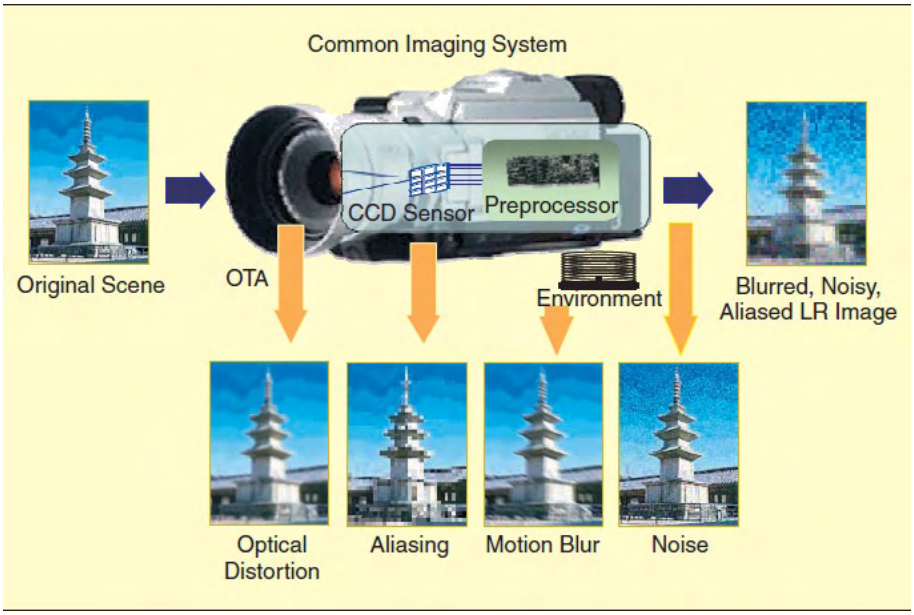


Figure 1.1: Image acquisition with a digital camera

Digital image processing has several applications in surveillance, satellite imaging, plate reading, criminal identification, target identification, diagnostics, etc. It is more advantageous when images used in these applications contain more detailed information, ie, these applications need High Resolution (HR) images. The HR images not only give the viewer a pleasing appearance but also offer additional information that is important for the analysis in many applications. Many image recognition and segmentation algorithms, do not work well with blurred or noisy images. Acquisition environment condition, resolution of image sensors employed, etc are some of the factors that affect the quality of digital image.

1.3 Image Resolution

The quality of conversion from the analog to the digital, determines image resolution. This conversion consists of two steps: sampling and quantization. The process of digitizing the spatial co-ordinates (x,y) is called sampling. Quantization

is the process of digitizing the amplitude values. The resultant image will get better as sampling frequency and quantization level increases. Image resolution is defined as the smallest measurable detail of visual representation. In optics the resolution of a device is determined by measuring the modulation transfer function (MTF) which measures the response of the system to different spatial frequencies [14]. There are two types of digital image resolution - spatial resolution and color or intensity (depth) resolution. Spatial resolution depends on sampling frequency and depth resolution on quantization levels. Analog images are continuous in detail (spatial resolution) and in color (depth resolution), while digital images inherently have limited spatial and color information.

1.3.1 Spatial resolution

Spatial resolution is the rate at which an image is sampled during scanning. Sampling is the principal factor determining the spatial resolution of an image. Basically, spatial resolution is the smallest discernible detail in an image [76]. An image is composed of various number of picture elements called pixels. Pixels can be considered as of samples of a continuous image which are arranged in a 2D rectilinear array. Spatial resolution refers to pixel spacing in an image. Higher spatial resolution allows more sharp details and subtle color transitions in an image [16]. If an image having high levels of details is not represented by a spatially dense set of pixels, the image is said to suffer from aliasing artifacts giving rise to blocky effects [16]. The more the number of pixels in an image, the higher is its resolution and the more pleasing the image appears. Figure 1.2 provides an example of the same image at different pixel resolutions. While going from left to right, the number of pixels increases from 1×1 to 100×100. Here it is clear that the right most image which contains more pixels have more image details. Spatial resolution is usually determined by the number of sensors used to digitize the image.

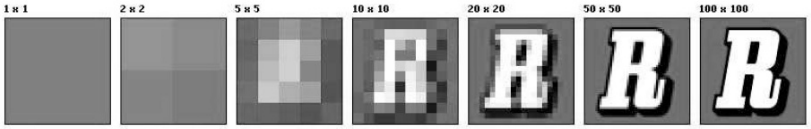


Figure 1.2: An illustration of how the same image appears at different pixel resolutions

1.3.2 Depth Resolution

The other resolution in digital imagery is color resolution or color depth resolution. This refers to the amount of information, the number of digital bits used to store each image pixel. Depth resolution is also defined as the gray level or color range of an individual pixel. It also refers to the smallest discernible change in gray or color level. Due to hardware considerations, the number of grey levels is usually an integer power of 2. One or more bits can define the color intensity of a pixel. The most common number is 8 bits. A single bit can represent a pixel value of black (0) or white (1). Two bits combined can represent a pixel value of black (00), dark gray (01), light gray (10), and white (11). Similarly, three bits can represent eight levels of intensity. Eight bits, or 1 byte, provide 256 levels of intensity. The number of bits used to hold each pixel value defines the depth resolution of a digital image.

The depth resolution required for various digital images depends on the content of the images and the application. The text on a printed page can be represented with a depth resolution of 1 bit, or with pixels of either black or white. Television images contain only about 7 bits of depth resolution, or 128 levels of intensity. Computer displays typically present up to 256 levels of intensity each for the red, green, and blue primary colors. Therefore, most displayed color images usually have a depth of 24 bits; 8 for the red component, 8 for the green, and 8 for the blue information. High end, full color images, such as those for film or detailed satellite imagery, often require 10 bits or even 12 bits for each color component.

This means that a total of up to 36 bits may be needed for each image pixel. One important point about spatial and depth resolutions is that they are not independent of one another. In other words, an image that has low spatial resolution and high depth resolution could, in its final output form, look similar in quality to the one that has high spatial resolution but low depth resolution.

1.3.3 Physical Limits of Resolution

The imaging sensors or the image acquisition device limits the image spatial resolution. Spatial resolution is important for the diagnostic quality for a cardiac CT examination because it directly translates to accurate, clear images—giving the physician the data to help correctly establish the extent of coronary artery disease (CAD). 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. With higher density of the sensors, higher spatial resolution possible for the imaging system. An imaging system with inadequate detectors will generate low-resolution images with blocky effects, due to the aliasing from low spatial sampling frequency. Aliasing is the jagged edges that appear in digital images. Aliasing occurs when the frequency at which the image sampled is lower than the maximum sampling frequency. If the image is sampled below Nyquist rate, frequencies fold over on one another causing these undesirable effects. The intuitive solution is to change the sampling rate of the hardware, but in many cases this sampling rate is fixed [70].

As the sensor size decreases, the amount of light incident on each sensor also decreases, causing the so-called shot noise. Therefore, there exists a limit for pixel size reduction. The optimally limited pixel size is estimated at about $40\mu m^2$ for a $0.35\mu m$ CMOS process. The present-day imaging technology has attained this level. To enhance the spatial resolution, if the chip size is increased, it leads to an increase in capacitance. This approach is not considered effective, since large capacitance makes it difficult to speed up a charge transfer rate. Also, the hardware

cost of a sensor increases with the increase of sensor density or corresponding image pixel density. The spatial resolution of an image that can be captured is restricted by the hardware limitation of sensor technology. While the image sensors limit the spatial resolution of the image, 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, aperture diffractions, and optical blurring due to motion. Constructing imaging chips and optical components to capture very high-resolution images is prohibitively expensive and is not practical in most real applications. Besides the cost, the resolution of a surveillance camera is also limited in the camera speed and hardware storage. In some other scenarios such as satellite imaging, it is difficult to use high resolution sensors due to physical constraints. One solution to this problem is to accept the image degradations, and use signal processing techniques to post-process the captured images.

Therefore, some image processing methods are needed to construct a high resolution (HR) image from one or more available LR images. LR images are typically under-sampled and blurred versions of the original scene. Super resolution problem is an inverse problem and refers to the process of producing a high resolution image , making use of one or more LR observations. It includes up sampling the image, thereby, increasing the maximum spatial frequency and removing degradations that arise during the image capture namely aliasing and blurring.

1.4 Super resolution

Super-Resolution (SR) is considered as a technique to increase the resolution of an image or a sequence of images beyond the resolving power of the imaging system. Pixel count (i.e; the number of pixels in an image) is not an appropriate measure of image resolution and simply increasing the pixel count does not improve the resolution. There has always been a demand for detail in imaging. Super resolved image reconstruction has proved to be effective in many areas including medical imaging, satellite imaging, video applications, image enlarging in web

pages and restoration of old historic photographs, tracking, license plate recognition system, etc. Techniques such as bilinear and bicubic interpolation only consider low resolution image information. So they only increase the pixel count without actually adding the details, and the resulting image from these techniques is often blurry and contains artifacts. These techniques perform well in smoother regions of the images and tend to blur edges and other sharp details in the images.

Super-Resolution (SR) techniques are used to construct high resolution (HR) images from several observed low resolution (LR) images or from a single low resolution image, by increasing the high-frequency components and removing the degradations caused by the imaging process of the low-resolution camera. In other words, image super resolution is the process by which additional higher frequency information is incorporated to enhance a low resolution image thereby producing a high resolution image. Along with the original information inherent within the low-resolution image, this additional information may come in several forms: a group of several shifted versions of the low resolution image, a collection of optimally estimated filters selected for specific image content, or as in the case presented in this thesis, a relationship using a training set that contains low and high resolution image pairs.

Super resolution techniques not only increase the size of the input low resolution image but they also reduce the degradations present in the LR image to obtain its high resolution image. Work presented in this thesis super resolves input low resolution images to two and four times their size (with magnification factors 2 and 4).

In general, there are two types of super resolution techniques – reconstruction based and learning based.

1.4.1 Reconstruction based Super resolution

The basic idea behind reconstruction based Super Resolution (SR) is to combine the non redundant information contained in multiple low-resolution frames to

generate a high-resolution image. The camera captures several low resolution images, which are actually down sampled form of the original scene with sub pixel shifts between each other. The sub pixel shifts are formed from the motion between the camera and scene, for example movements of objects, or due to controlled motion like that of the satellite imaging system that orbits round the earth with predefined speed and path. Super resolution reconstruction reverses this process by aligning the low resolution observations to sub pixel accuracy and combining them into a HR image grid (interpolation), thereby overcoming the imaging limitation of the camera. The non redundant information contained in the these low resolution images is due to the sub pixel shifts between them. One disadvantage of reconstruction based super resolution is that effective reconstruction is possible only if there exists sub pixel motion between low resolution images.

The problem of merging or fusing information from a number of low resolution images into a single high-resolution image is referred to as multi-frame super-resolution. Figure 1.3 shows the basic idea of multi- frame reconstruction based super resolution.

Figure 1.3: Multi-frame super-resolution

A typical multi frame super-resolution involves three sub-tasks: registration, fusion and deblurring. First, the LR images are registered against a common reference to a sub pixel accuracy. An intermediate image at a higher resolution is then constructed from the scattered input samples. The HR fusion image is finally deblurred to amplify the frequency spectrum beyond the Nyquist frequency of the imaging sensor. These process are shown in Figure 1.4.

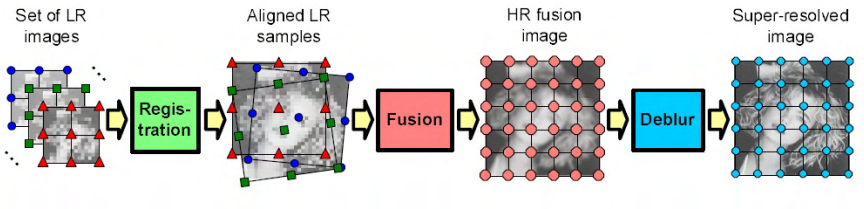


Figure 1.4: Image registration, image fusion, image deblurring

1.4.2 Learning based Super resolution

Recent super-resolution algorithms, based on machine learning take advantage of learning and training images to preserve spatial-frequency and recover information of details such as edges and textures. In learning based super-resolution algorithms, a training set or database of available HR images are used to construct the HR image of an image captured using a LR camera. In the training set, images are stored as patches or coefficients of feature representations like wavelet transform, DCT, etc. Unlike the reconstruction-based method which requires multiple LR input images, here only one input image (single frame image super-resolution) is required. Single frame image super-resolution can be used in applications where database of HR images are available. These methods are classified under the motion free super-resolution scheme. Figure 1.5 shows schematic diagram of learning based super-resolution.

From the viewpoint of learning based super-resolution, there is loss of information in low resolution images and the task of super-resolution is to retrieve the missing information correctly from training samples and patterns. Some of the research work based on machine learning addresses the single image super resolution problem. The seminal work in this area is done by William T. Freeman's group in the Artificial Intelligence Lab at Massachusetts Institute Technology. Generally speaking, the main concept of single image super-resolution relies on a carefully selected training set of high-resolution images. The normalized training images are down sampled and divided into low resolution patches, each one forming

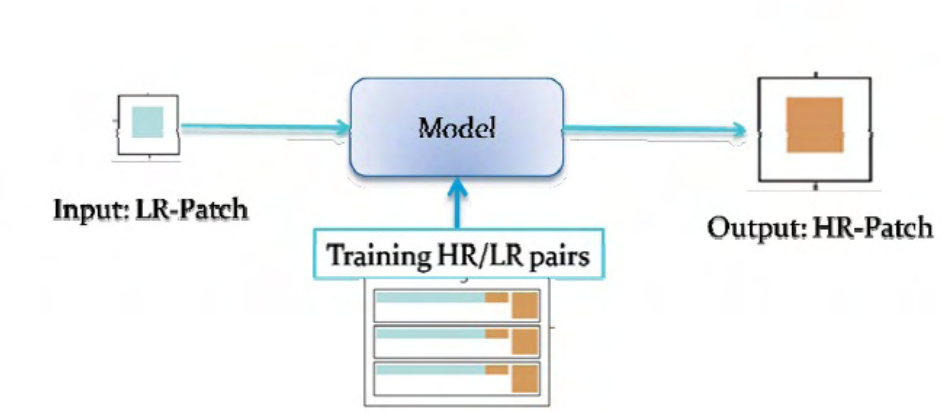


Figure 1.5: Learning based super-resolution

a training pair with the corresponding high resolution patch. Such super-resolution algorithms try to learn the connection between the patches from LR input images and the ones from training set, and try to reconstruct the high-resolution counterpart using the high resolution training patches. Different learning based single image super-resolution approaches share similar concepts as mentioned above, but vary in their ways of choosing effective features, designing mechanisms to find appropriate matching patterns, and so on. Freeman et al presented the method which first generates the high resolution estimate with traditional interpolation, and then finds the missing high frequency information from training images and finally combine the estimate with high frequency information to obtain the target HR image.

It is clear that example based (learning based) super-resolution is an image interpolation algorithm which uses a database of training images to create plausible high frequency details in zoomed images. Though this method is fairly simple, its performance heavily depends on the images in the database. In particular, when the characteristics of a target image to be magnified are different from the training

images, the quality of the super resolved image degrades. The performance is improved by transforming selected images which are downloaded from Internet photo sharing sites, to match their characteristics with those of the target image before adding them to the database. The advantage of this method is that by skillfully creating a database of suitable training images, one can improve the quality of the super-resolved image. On the other hand, the shortcoming of this method is the possibility of failure unless images being similar to the target image are available on the internet. This problem should be solved as the number of uploaded images increases everyday.

Super resolution is a computationally intensive inverse problem typically involving tens of thousands unknowns. For example, super resolving a sequence of 32×32 pixel LR image by a factor of 4 in each spatial dimension involves 128×128 unknown pixel values in the HR image.

1.5 Objective of the present Work

Super-resolution is a restoration technique which can be used to increase the resolution of a low resolution image. Super-resolution method not only increases the resolution of an image, but also removes degradations and artifacts present in the low resolution image. However, most methods used to produce super-resolution are found to have some defects. The major defect noticed in these methods is that they fail to remove artifacts like aliasing, noise etc. They also require high computation time and computation complexity. The aims of the work presented in this thesis are:

1. To develop a learning based super-resolution method which can produce high quality images with high spatial frequency and less artifacts like blurring, noise and aliasing.

2. To develop a method that requires less computation time.

3. To reduce the computational complexity of the super-resolution method.

In brief the aim of this work is to optimise a super resolution method in terms of quality, time and computational complexity.

1.6 Organisation of the thesis

The main problem addressed in this thesis is learning based single image super-resolution method.

The second chapter is an overview of the previous literature work in this field .

The main body of the Ph.D thesis is presented in five chapters from Chapter 3 to Chapter 7. In the third chapter, three single image super-resolution methods are discussed. First method is a study of the work done by William T.Freemann [108]. Second and third methods are based on wavelets. Here the third method is obtained by modifying the second method[15] by patch based approach.

The fourth chapter explains a new single image super-resolution method based on skewed anisotropic transform called directionlets. This method outperforms the standard interpolation methods and the wavelet methods, both visually and in terms of SNR values. The super-resolution methods are implemented using, both critically sampled and over sampled directionlets.

The conventional directionlet transform is computationally complex. Hence lifting algorithms are used for implementation of directionlets. The fifth chapter describes single image super-resolution method based on lifting based directionlets. This chapter also describes a study of different wavelets on this single image super-resolution method.

The sixth chapter extends the new single image super-resolution method to color images. Two methods are presented here. In the first method images in RGB format is converted to YC_bC_r format and the super-resolution method is applied to the luminance (Y) component. The other two color components C_b, C_r are

interpolated using standard methods. In the second method three color components r,g,b are super resolved separately to obtain the high resolution color image.

All the above chapters describe single image super-resolution method on noise free images. Seventh chapter explains the effect of the single image super-resolution method on noisy images.

The chapter 8 presents conclusions of the thesis and future work directions.

Chapter 2

Literature Review

The work in super resolution area is progressing very rapidly with the introduction of many signal processing techniques such as Wavelet transform, different learning methods etc. Like all fields and particularly emerging ones, it has a plethora of different motivations. This chapter gives a detailed account of the previous work done in reconstruction and learning based super resolution area.

2.1 Super-resolution : An overview

As explained in the previous chapter, there are two types of super resolution techniques- reconstruction based and learning based. In reconstruction based techniques, high resolution image is recovered from several low resolution observations of the input.Impressive amount of work has been reported in this field. Frequency domain approach proposed by Tsai in "Multiframe image restoration and registration" [100] is the first work in super resolution. They considered the super resolution problem described above subject to the assumption that the low resolution frames have neither been corrupted by noise nor degraded by a blurring phenomenon. In the paper " High-resolution image recovery from image-plane arrays, using convex projections" [96], Stark et al considers the problem of reconstructing remotely obtained images from image-plane detector arrays. Although the individual detectors may be larger than the blur spot of the imaging optics, high-resolution reconstructions can be obtained by scanning or rotating the image with respect to the detector. As an alternative to matrix inversion or least-squares estimation, the method of convex projections is proposed. It is also shown that readily obtained prior knowledge can be used to obtain good-quality imagery with reduced data.

Kim et al [45] in their paper, propose a recursive algorithm for restoration of super resolution images from noisy and blurred observations. They use the aliasing relationship between the under sampled frames and the reference image, to develop a weighted recursive least-squares theory based algorithm in the wave number domain. This algorithm is efficient because interpolation and noise removal are performed recursively and in addition, it is highly suitable for implementation through the massively parallel computational architectures currently available. Accurate knowledge of the relative scene locations sensed by each pixel in the observed images is necessary for super resolution. This information is available in image regions where local deformation can be described by some parametric function. Such functions can describe, for example, perspective transformation. Authors assumed that local motion can be described by translations

and rotations only, but the approach is applicable also for other image motion models. In 1990 Irani et al , propose a method in the paper [62] which is similar to back-projection used in tomography. In tomography, images are reconstructed from their projections in many directions. In the super resolution case, each low resolution pixel is a "projection" of a region in the scene whose size is determined by the imaging blur. Accurate knowledge of the relative scene locations sensed by each pixel in the observed images is necessary for super resolution. This information is available in image regions where local deformation can be described by some parametric function. Such functions can describe, for example, perspective transformation. Like[45], here also it is assumed that local motion can be described by translations and rotations only, but the approach is applicable also for other image motion models. Here it is shown that when the algorithm is applied to a single image without increasing the sampling rate, super resolution reduces to deblurring. In 1993, Kim et al present an approach in the article [44] to obtain high resolution image reconstruction from low-resolution, blurred, and noisy multiple-input frames. Here a recursive least squares approach with iterative regularization is developed in the discrete Fourier transform (DFT) domain. When the input frames are processed recursively, the reconstruction does not converge in general due to the measurement noise and ill-conditioned nature of the deblurring. Through the iterative update of the regularization function and the proper choice of the regularization parameter, good high-resolution reconstructions of low-resolution, blurred, and noisy input frames are obtained. The reconstruction is done independently for each DFT element. It is shown by Bose et al, in their paper[68] how the total least squares recursive algorithm for the real data FIR (finite impulse response) adaptive filtering problem can be applied to reconstruct a high-resolution filtered image from undersampled, noisy multiframes, when the interframe displacements are not accurately known. This is done in the wave number domain after transforming the complex data problem to an equivalent real data problem. The procedure developed also applies when the multiframes are degraded by linear shift-invariant blurs.

Richard R. Schultz introduces a method in the article titled "A bayesian approach to image expansion for improved definition" [79] for nonlinear image expansion which preserves the discontinuities of the original image, producing an expanded image with improved definition. The Maximum A Posteriori (MAP) estimation techniques that are proposed for noise-free and noisy images result in the optimization of convex functionals. Another approach toward the super-resolution restoration problem is presented by the same author in the paper "Extraction of high-resolution frames from video sequences" [78]. Their approach uses MAP estimator, with the Huber-Markov Random Field (HMRF) prior. The blur of the measured images is assumed to be simple averaging, and the measurements additive noise is assumed to be independent and identically distributed Gaussian vector. This choice of prior causes the entire problem to be non quadratic, thus complicating the resulting minimization problem. The maximum a posteriori (MAP) estimation techniques that are proposed for noise-free and noisy images result in the optimization of convex functionals.This paper addressed how to utilize both the spatial and temporal information present in a short image sequence to create a single high-resolution video frame. A novel observation model based on motion compensated subsampling was proposed for a video sequence. Since the reconstruction problem is ill-posed, Bayesian restoration with a discontinuity-preserving prior image model is used to extract a high-resolution video still given a short low-resolution sequence. Tekalp et al [5] propose i) a complete model of video acquisition with an arbitrary input sampling lattice and a nonzero aperture time, and ii) an algorithm based on this model using the theory of projections onto convex sets to reconstruct SR still images or video from an LR time sequence of images. Michael Elad et al in 1997, propose a unified methodology in the paper[20] toward the more complicated problem of super resolution restoration using the tools maximum likelihood (ML) estimator, the maximum a posteriori probability (MAP) estimator, and the set theoretic approach using projection onto convex sets (POCS). In this super resolution restoration problem, an improved resolution image is restored from several geometrically warped, blurred, noisy and down sampled measured images. For this the super resolution restoration problem

is modeled and analyzed from the ML, the MAP, and POCS points of view, yielding a generalization of the known super resolution restoration methods. It assumes explicit knowledge of the linear space and time variant blur, the (additive Gaussian) noise, the different measured resolutions, and the (smooth) motion characteristics. It is also shown that super resolution can be achieved even without motion. The same authors presents a new method in the paper "Super resolution restoration of an image sequence: Adaptive filtering approach" [58] based on adaptive filtering theory for super resolution restoration of continuous image sequences. They modify the previous work as approximations of the Kalman filter and then carry out a thorough analysis of their performance in paper[57]. For each algorithm, they calculated a bound on its deviation from the Kalman filter performance. They also show that the propagated information matrix within the R-SD algorithm remains sparse in time, thus ensuring the applicability of this algorithm. This analysis includes two parts. The first part corresponds to the convergence properties of the proposed algorithms and the second part relates to the computational complexity of the R-SD algorithm. In the year itself the same authors propose a methodology which suggests least squares (LS) estimators which adapt in time, based on adaptive filters, least mean squares (LMS) or recursive least squares (RLS). The adaptation enables the treatment of linear space and time variant blurring and arbitrary motion, both of them assumed known. The above mentioned algorithms are adequate for tasks such as satellite fusion of several images and in combining several sources to improve targets detection and identification, or generating an improved still picture in VCR's. However, when applied to a typical video scene, the production of one super resolution output image cannot be regarded as a restoration of the scene. The required output should be a sequence of super resolution images with the same length as the source sequence and the same geometric behavior.

The paper "A regularized multichannel restoration approach for globally optimal high resolution video sequence" [63] introduces an iterative regularized approach to obtain a high resolution video sequence. A multiple input smoothing convex functional is defined and used to obtain a globally optimal high resolution video sequence. A mathematical model of multiple inputs is described by using the

point spread function between the original and bilinearly interpolated images in the spatial domain, and motion estimation between frames in the temporal domain. Properties of the proposed smoothing convex functional are analyzed. An iterative algorithm is utilized for obtaining a solution. The regularization parameter is updated at each iteration step from the partially restored video sequence

In paper "An iterative weighted regularized algorithm for improving the resolution of video sequences" [55] authors introduce an iterative regularized approach to increase the resolution of a video sequence. A multiple input smoothing convex functional is defined and used to obtain a globally optimal high resolution video sequence. A mathematical model of multiple inputs is described by using the point spread function between the original and bilinearly interpolated images in the spatial domain, and motion estimation between frames in the temporal domain. An iterative algorithm is utilized for obtaining the solution. The regularization parameter is updated at each iteration step from the partially restored video sequence.

In 1997, Hardie et al [84] propose a maximum a posteriori (MAP) framework for jointly estimating image registration parameters and the high-resolution image. Several previous approaches had relied on knowing the registration parameters a priori or have utilized registration techniques not specifically designed to treat severely aliased images. In the proposed method, the registration parameters are iteratively updated along with the high-resolution image in a cyclic coordinate-descent optimization procedure. Therefore,in the above paper authors seek to minimize a MAP cost function with respect to the high-resolution image and the registration parameters simultaneously using a cyclic coordinate-descent optimization procedure. In this iterative technique, the registration parameter estimates are updated using the current best estimate of the high-resolution image. Ming-Chao et al [53] introduces two algorithms for enhancing image resolution from an image sequence. The “image-based” approach presumes that the images were taken under the same illumination conditions and uses the intensity information provided by the image sequence to construct the high-resolution image. This ideal, however, is almost always not true when the illumination varies.

The “edge-based” approach, based on edge models and a local blur estimate, circumvents these difficulties.

A discrete cosine transform (DCT)-based method is proposed by Kang et al in the paper[87]. 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.

The approach, the authors propose in the paper titled "Efficient super-resolution via image warping" [54] uses the integrating resampler for warping. The method is a direct computation, which is fundamentally different from the iterative back-projection approaches proposed in previous work. This paper shows that image-warping techniques may have a strong impact on the quality of image resolution enhancement. By coupling the degradation model of the imaging system directly into the integrating resampler, the warping characteristics of real sensors can be approximated effectively, which also significantly improve the quality of super-resolution images.

Nhat Nguyen et al [71] propose efficient block circulant pre conditioners for solving the regularized super resolution problem. In the paper "An efficient wavelet based algorithm for image superresolution" [65], the same author presents a new and efficient wavelet based algorithm for image superresolution. The algorithm is a combination of interpolation and restoration processes. This method exploits the interlaced sampling structure in the low resolution data. In the paper, [8] Peleg et al introduces two novelties. First, a framework for super resolution algorithms is presented, which enables the development of very efficient algorithms. Second, a method for applying super resolution to panoramic mosaics is presented. Another robust approach for super resolution is presented in the paper "Robust super-resolution" [7], which is essentially valuable in the presence of outliers. Such outliers may be due to motion errors, inaccurate blur

models, noise, moving objects, motion blur etc. This robustness is needed since super-resolution methods are very sensitive to such errors. A robust median estimator is combined in an iterative process to achieve a super resolution algorithm. This process can increase resolution even in regions with outliers, where other super resolution methods actually degrade the image. Nguyen et al [66] propose an efficient block circulant preconditioners for solving the Tikhonov regularized superresolution problem by the conjugate gradient method. They also extend their work to underdetermined systems with the derivation of the generalized cross-validation method for automatic calculation of regularization parameters. Here, they present efficient circulant block preconditioners that take advantage of the inherent structures in the superresolution system matrix to accelerate CG. They adopt the generalized cross-validation (GCV) method, which is often used to calculate regularization parameters for Tikhonov-regularized overdetermined least squares problems without accurate knowledge of the variance of noise, to our underdetermined problem.

Elad et al [60] address the problem of recovering a super-resolved image from a set of warped, blurred and decimated versions there of. The main contribution of this paper corresponds to the fusion stage, where the measurements are merged into a higher resolution image. It is shown that through a very simple non iterative algorithm, this fusion is achieved, while preserving the optimality in the Maximum-Likelihood sense. In their next paper [91], they prove that additive Gaussian distribution is not a proper model for super-resolution noise. Specifically, they show that Lp norm minimization results in a pixel wise weighted mean algorithm which requires the least possible amount of computation time and memory and produces a maximum likelihood solution. They also justify the use of a robust prior information term based on bilateral filter idea.With the objective of improving the performance of the signal processing algorithms in the presence of the ubiquitous perturbation errors of displacements around the ideal subpixel locations (because of imperfections in fabrication), in addition to noisy observation, the errors-invariables or the total least-squares method is used in paper "Constrained total least squares computations for high- resolution image reconstruction with

multisensors" [61]. A regularized constrained total least-squares (RCTLS) solution to the problem is given, which requires the minimization of a nonconvex and nonlinear cost functional. Simulations indicate that the choice of the regularization parameter influences significantly the quality of the solution. The L-curve method is used to select the theoretically optimum value of the regularization parameter instead of the unsound but expedient trial-and-error approach. Keren et al, in their paper [113] suggest a new approach which exploits the correlation between neighboring information in the “steerable wavelet” representation. The advantage of working in the wavelet domain is that the smoothness assumption is applied in the appropriate scale and in the appropriate orientation. The demosaic results using this approach provides a faithful interpolation of missing samples while preserving edges and textures in the resulting image. In the paper[118], Peleg et al present a new way to combine the information from different non-registered sensors. Given a set of images of possibly different sensors viewing the same scene, the resolution of one image is improved by using the other images. In color RGB sensors, for example, the red channel is enhanced using the green and blue channels. Similarly the green channels are enhanced using the red and blue channels. The result of combining the enhanced resolution channels is a higher resolution color image.

In 2003, Chan et al analyze the super resolution problem from the wavelet point of view [77]. By expressing the true image as a function in L(R2), the authors derive iterative algorithms which recover the function completely in the L sense from the given low-resolution functions. These algorithms decompose the function obtained from the previous iteration into different frequency components in the wavelet transform domain and add them into the new iterate to improve the approximation. Wavelet (packet) thresholding methods are applied to denoise the function obtained in the previous step before adding it into the new iterate. In paper titled "Variational approaches to super-resolution with contrast enhancement and anisotropic diffusion " [30], Kim et al present super-resolution methods that enhance image contrast and perform anisotropic diffusion simultaneously. Since the super-resolution problem is solved by encouraging Mach-band profiles while incorporating anisotropic diffusion, this technique not only reconstructs a high-resolution image from several

overlapping noisy low-resolution images, but also enhances edges and image contrast while suppressing image noise during the restoration process.

Elad et al [92] propose a fast and robust hybrid method of super-resolution and demosaicing, based on a maximum a posteriori (MAP) estimation technique by minimizing a multi-term cost function. The L1 norm is used for measuring the difference between the projected estimate of the high-resolution image and each low-resolution image, removing outliers in the data and errors due to possibly inaccurate motion estimation. Bilateral regularization is used for regularizing the luminance component, resulting in sharp edges and forcing interpolation along the edges and not across them. Simultaneously, Tikhonov regularization is used to smooth the chrominance component. Finally, an additional regularization term is used to force similar edge orientation in different color channels.

In the paper "Fast and robust multiframe super resolution" [22], the authors review some of these methods and addresses their shortcomings. They propose an alternate approach using L1 norm minimization and robust regularization based on a bilateral prior to deal with different data and noise models.This paper reviews some of these methods and addresses their shortcomings. The authors propose an alternate approach using L1 norm minimization and robust regularization based on a bilateral prior to deal with different data and noise models.

The work presented in the paper titled "Direct super-resolution and registration using raw cfa images" [27] aims at producing a high-resolution color image directly from raw color mosaic images obtained by a single CCD equipped with a color filter array. This method is based on a generalized formulation of superresolution that simultaneously performs both resolution enhancement and demosaicing.

In the paper, [52] Bose et al investigate the effect of the threshold level on reconstructed image quality in second-generation wavelet super-resolution. The choice of optimal threshold involves a tradeoff between noise filtering and blurring introduced by thresholding. This can then be formulated as an optimization problem whose solution is obtained by minimizing a cost function. The approaches adopted and their analysis to solve the formulated optimization problem are crucial,

The image acquisition scheme is important in the modeling of the degradation process. The need for model accuracy is undeniable in the attainment of SR along with the design of the algorithm whose robust implementation will produce the desired quality in the presence of model parameter uncertainty.

Using a stochastic framework, Woods et al propose two algorithms for the problem of obtaining a single high-resolution image from multiple noisy, blurred, and undersampled images in the paper[64]. The first is based on a Bayesian formulation that is implemented via the expectation maximization algorithm. The second is based on a maximum a posteriori formulation. In their formulations, the registration, noise, and image statistics are treated as unknown parameters. These unknown parameters and the high-resolution image are estimated jointly based on the available observations.

In paper, [25] Nathan et al propose a maximum a posteriori (MAP) framework for the super resolution problem, i.e. reconstructing high-resolution images from shifted, rotated, low-resolution degraded observations. The main contributions of this work are two; first, the use of a new locally adaptive edge preserving prior for the super resolution problem. Second, an efficient two-step reconstruction methodology that includes first an initial registration using only the low resolution degraded observations. This is followed by a fast iterative algorithm implemented in the discrete fourier transform domain in which the restoration, interpolation and the registration subtasks of this problem are preformed simultaneously.

In the paper "Super resolution blind reconstruction of low resolution images using wavelets based fusion", [50] authors propose a super resolution blind reconstruction technique for linearly degraded images. In this technique the algorithm is divided into three parts - an image registration, wavelets based fusion and an image restoration. For this, three low resolution images are considered which may sub pixels shifted, rotated, blurred or noisy. The sub pixel shifted images are registered using affine transformation model. A wavelet based fusion is performed and the noise is removed using soft thresholding.

Multiframe super resolution (SR) reconstruction aims to produce a high-resolution (HR) image using a set of low-resolution (LR) images. In the process of reconstruction, fuzzy registration usually plays a critical role. It mainly focuses on the correlation between pixels of the candidate and the reference images to reconstruct each pixel by averaging all its neighboring pixels. Therefore, the fuzzy-registration-based SR performs well and has been widely applied in practice. However, if some objects appear or disappear among LR images or different angle rotations exist among them, the correlation between corresponding pixels becomes weak. Thus, it will be difficult to use LR images effectively in the process of SR reconstruction. Moreover, if the LR images are noised, the reconstruction quality will be affected seriously. To address or at least reduce these problems, Xinbo Gao et al presents a novel SR method based on the Zernike moment, to make the most of possible details in each LR image for high-quality SR reconstruction in paper [112].

In paper[24], Gao et al propose a sparse neighbor selection (SpNS) scheme for SR reconstruction. They first predetermine a larger number of neighbors as potential candidates and develop an extended Robust- SL0 algorithm to simultaneously find the neighbors and solve the reconstruction weights. Recognizing that the k-NN for reconstruction should have similar local geometric structures based on clustering, they employ a local statistical feature, namely Histograms of oriented gradients (Hog) of low-resolution (LR) image patches, to perform such clustering. By conveying local structural information of Hog in the synthesis stage, the k-NN of each LR input patch are adaptively chosen from their associated subset, which significantly improves the speed of synthesizing the HR image while preserving the quality of reconstruction.

In the paper [43], Zhang et al propose a simple preconditioning method for accelerating the solution of edge-preserving image super-resolution (SR) problems in which a linear shift-invariant point spread function is employed. Their technique involves reordering the high-resolution (HR) pixels in a similar manner to what is done in preconditioning methods for quadratic SR formulations. However, due to the edge preserving requirements, the Hessian matrix of the cost function varies

during the minimization process. They develop an efficient update scheme for the preconditioner in order to cope with this situation. Unlike some other acceleration strategies that round the displacement values between the low-resolution (LR) images on the HR grid, the proposed method does not sacrifice the optimality of the observation model. In addition, they describe a technique for preconditioning SR problems involving rational magnification factors.

Accurate image registration plays a preponderant role in image super-resolution methods and in the related literature landmark based registration methods have gained increasing acceptance in this framework. However, their solution relies on point correspondences and on least squares estimation of the registration parameters necessitating further improvement. In paper[104], Vrigkas et al propose a maximum a posteriori scheme for image super resolution is presented where the image registration part is accomplished in two steps. At first, the low resolution images are registered by establishing correspondences between robust SIFT features. In the second step, the estimation of the registration parameters is fine-tuned along with the estimation of the high resolution image, in an iterative procedure, using the maximization of the mutual information criterion. In 2012, Pelletier et al propose a simple preconditioning method for accelerating the solution of edge-preserving image super-resolution (SR) problems in which a linear shift-invariant point spread function is employed [69]. Their technique involves reordering the high-resolution (HR) pixels in a similar manner to what is done in preconditioning methods for quadratic SR formulations. However, due to the edge preserving requirements, the Hessian matrix of the cost function varies during the minimization process. They develop an efficient update scheme for the preconditioner in order to cope with this situation. Unlike some other acceleration strategies that round the displacement values between the low-resolution (LR) images on the HR grid, the proposed method does not sacrifice the optimality of the observation model. In addition, they describe a technique for preconditioning SR problems involving rational magnification factors.

Reconstruction based super resolution can be used when more than one low resolution images are available. In some situations, only one low resolution image is available. In that case, a database with available high resolution images are used to super resolve the input low resolution image. But in learning based super resolution algorithms, a training set of available high resolution images are used to obtain the high resolution of an image captured using a low resolution camera. In the training set, images are stored as patches or coefficients of feature representations like wavelet transform, DCT ,etc. Unlike the reconstruction based method which requires multiple low resolution input images, here only one input image (single frame image super resolution) is required. Single frame image super resolution can be used in applications where database of high resolution images are available. These methods are classified under the motion free super resolution scheme. Because the richness of real-world images is difficult to capture analytically, for the past several years. They have been exploring a learning-based approach for enlarging images. In a training set, the algorithm learns the fine details that correspond to different image regions seen at a low-resolution and then uses those learned relationships to predict fine details in other images. In the first part of the paper [11], the authors derive a sequence of analytical results which show that the reconstruction constraints provide less and less useful information as the magnification factor increases. They also validate these results empirically and show that for large enough magnification factors any smoothness prior leads to overly smooth results with very little high-frequency content (however many low resolution input images are used.) In the second part of this paper, they propose a super-resolution algorithm that uses a different kind of constraint, in addition to the reconstruction constraints. The algorithm attempts to recognize local features in the low resolution images and then enhances their resolution in an appropriate manner.

The paper [2] describes a new framework for processing images by example, called “image analogies.” The framework involves two stages: a design phase, in which a pair of images, with one image purported to be a “filtered” version of the other, is presented as “training data”; and an application phase, in which the

learned filter is applied to some new target image in order to create an “analogous” filtered result. Image analogies are based on a simple multi-scale auto regression, inspired primarily by recent results in texture synthesis. By choosing different types of source image pairs as input, the framework supports a wide variety of “image filter” effects, including traditional image filters, such as blurring or embossing; improved texture synthesis, in which some textures are synthesized with higher quality than by previous approaches; super-resolution, in which a higher-resolution image is inferred from a low-resolution source; texture transfer, in which images are “texturized” with some arbitrary source texture; artistic filters, in which various drawing and painting styles are synthesized based on scanned real-world examples; and texture-by-numbers, in which realistic scenes, composed of a variety of textures, are created using a simple painting interface.

In paper [108], Freeman et al propose an example based super resolution method in which he had developed a Bayesian propagation algorithm using Markov Network.

The paper titled "Is super-resolution with optical flow feasible" [107] is an attempt at understanding the influence of image alignment and warping errors on super-resolution. Requirements on the consistency of optical flow across multiple images are studied and it is shown that errors resulting from traditional flow algorithms may render super-resolution infeasible.

Pickup et al present a domain-specific image prior based upon sampled images in the paper[72]. The paper "Learning based super-resolution imaging: Use of zoom as a cue" [42] propose a technique for super-resolution imaging of a scene from observations at different camera zooms. Given a sequence of images with different zoom factors of a static scene, they obtain a picture of the entire scene at a resolution corresponding to the most zoomed observation. The high-resolution image is modeled through appropriate parameterization, and the parameters are learned from the most zoomed observation. Assuming a homogeneity of the high-resolution field, the learned model is used as a prior while super-resolving the scene. They suggest the use of either a Markov random field (MRF) or an simultaneous autoregressive (SAR) model to parameterize the

field based on the computation one can afford. In this paper[15], Jiji et al propose a single frame, learning-based super-resolution restoration technique by using the wavelet domain to define a constraint on the solution. Wavelet coefficients at finer scales of the unknown high-resolution image are learned from a set of high resolution training images and the learned image in the wavelet domain is used for further regularization while super resolving the picture. They use an appropriate smoothness prior with discontinuity preservation in addition to the wavelet based constraint to estimate the super-resolved image. The smoothness term ensures the spatial correlation among the pixels, whereas the learning term chooses the best edges from the training set. Because this amounts to extrapolating the high frequency components, the proposed method does not suffer from over smoothing effects

In the paper[17], the same authors propose an eigen image based super resolution reconstruction technique. In paper titled "Patch based blind image super resolution" [74], Qiang et al presented, a novel method for learning based image super resolution (SR) is presented. The basic idea is to bridge the gap between a set of low resolution (LR) images and the corresponding high resolution (HR) image using both the SR reconstruction constraint and a patch based image synthesis constraint in a general probabilistic framework. In this framework, the estimation of the LR image formation parameters is straightforward. The whole framework is implemented via an annealed Gibbs sampling method.

In the paper "Non-parametric image super-resolution using multiple images" [28], the authors present a novel learning based framework for performing super-resolution using multiple images. They model the image as an undirected graphical model over image patches in which the compatibility functions are represented as non-parametric kernel densities which are learnt from training data. The observed images are translation rectified and stitched together onto a high resolution grid and the inference problem reduces to estimating unknown pixels in the grid.

In 2006, Todd et al [99] propose to use even stronger prior information by extending MRF-based super-resolution to use adaptive observation and transition

functions, that is, to make these functions region-dependent. They adapt the modeling for each image patch so as to improve the resolution.

In paper [17], Jiji et al propose a learning-based, single-image super-resolution reconstruction technique using the contourlet transform, which is capable of capturing the smoothness along the contours, making use of directional decompositions. The contourlet coefficients at finer scales of the unknown high-resolution image are learned locally from a set of high-resolution training images, the inverse contourlet transform of which recovers the super-resolved image. In effect, they learn the high-resolution representation of an oriented edge primitive from the training data.

The paper titled "Learning-based nonparametric image super-resolution" present a novel learning-based framework for zooming and recognizing images of digits obtained from vehicle registration plates, which have been blurred using an unknown kernel[89]. They model the image as an undirected graphical model over image patches in which the compatibility functions are represented as nonparametric kernel densities. The crucial feature of this work is an iterative loop that alternates between super-resolution and restoration stages. A machine-learning-based framework has been used for restoration which also models spatial zooming. Image segmentation is done by a column-variance estimation-based “dissection” algorithm. Initially, the compatibility functions are learned by nonparametric kernel density estimation, using random samples from the training data.

In the paper[51], Pickup et al attempt to shed some light on this problem when the SR algorithms are designed for general natural images (GNIs). They first define an expected risk for the SR algorithms that is based on the root mean squared error between the super resolved images and the ground truth images. Then utilizing the statistics of GNIs, they derive a closed form estimate of the lower bound of the expected risk. The lower bound can be computed by sampling real images. By computing the curve of the lower bound with respect to the magnification factor, they can estimate the limits of learning-based SR algorithms, at which the lower bound of expected risk exceeds a relatively large threshold. The paper propose an

alternative super-resolution method based on image fusion (called super-fusion hereafter). Image fusion has been proven to be effective in many applications. Extending image fusion to super-resolve images, they show that super-fusion is a faster alternative that imposes less requirements and is more stable than traditional super-resolution methods

In the paper[33], Isabelle presents comparisons of two learning-based super-resolution algorithms as well as standard interpolation methods. To allow quality assessment of results, a comparison of a variety of image quality measures is also performed.

In the paper "Single frame image super-resolution: should we process locally or globally?" [17], the authors study the usefulness of different local and global, learning-based, single-frame image super-resolution reconstruction techniques in handling three specific tasks, namely, de-blurring, de-noising and alias removal. They start with the global, iterative Papoulis–Gerchberg method for super-resolving a scene. Next they describe a PCA-based global method which faithfully reproduces a super-resolved image from a blurred and noisy low resolution input. They also study several multi resolution processing schemes for super-resolution where the best edges are learned locally from an image database. They show that the PCA-based global method is efficient in handling blur and noise in the data. The local methods are adept in capturing the edges properly. However, both local and global approaches cannot properly handle the aliasing present in the low resolution observation. Hence they propose an alias removal technique by designing an alias-free upsampling scheme. In the paper[80] Freeman et al introduce a method to remove the effects of camera shake from seriously blurred images. The method assumes a uniform camera blur over the image and negligible in-plane camera rotation. In order to estimate the blur from the camera shake, the user must specify an image region without saturation effects.

In 2006,Ju Liu et al [38] propose a logarithmic-wavelet transform (Log-WT) based method to combine super-resolution and shadow removing into a single operation. First intrinsic, illumination invariant features of the image are extracted by exploiting logarithmic-wavelet transform. Then an initial estimation of high

resolution image is obtained based on the assumption that small patches in low resolution space and patches in high resolution space share the similar local manifold structure. Finally the target high resolution image is reconstructed by applying the reconstruction constraints in pixel domain. The paper[85] "Learning the kernel matrix for super resolution" proposes the application of learned kernels in support vector regression to super resolution in the discrete cosine transform (DCT) domain. Though previous works involve kernel learning, their problem formulation is examined to reformulate the semi-definite programming problem of finding the optimal kernel matrix. For the particular application to superresolution, downsampling properties derived in the DCT domain are exploited to add structure to the learning algorithm. The advantage of the proposed method over other learning-based super resolution algorithms include specificity with regard to image content, structured consideration of energy compaction, and the added degrees of freedom that regression has over classification-based algorithms. Convolutional networks have achieved a great deal of success in high-level vision problems such as object recognition. In the paper [34], Jain et al show that they can also be used as a general method for low-level image processing. As an example of their approach, convolutional networks are trained using gradient learning to solve the problem of restoring noisy or degraded images. For training data, they have used electron microscopic images of neural circuitry with ground truth restorations provided by human experts. On this data set, Markov random field (MRF), conditional random field (CRF), and anisotropic diffusion algorithms perform about the same as simple thresholding, but superior performance is obtained with a convolutional network containing over 34,000 adjustable parameters. When restored by this convolutional network, the images are clean enough to be used for segmentation, whereas the other approaches fail in this respect.

The paper [9], Ayan et al present a learning-based method to super-resolve face images using a kernel principal component analysis-based prior model. A prior probability is formulated based on the energy lying outside the span of principal components identified in a higher-dimensional feature space. This is used to regularize the reconstruction of the high-resolution image.

The paper [12] addresses the problem of super-resolving a single image and recovering the characteristics of the sensor using a learning-based approach. In particular, the point spread function (PSF) of the camera is sought by minimizing the mean Euclidean distance function between patches from the input frame and from degraded versions of high resolution training images. Once an estimate of the PSF is obtained, a supervised learning algorithm can then be used as is. The paper [59] reviews an emerging powerful family of regularization techniques that is drawing attention in recent years—the example-based approach. The authors describe how examples can and have been used effectively for regularization of inverse problems, reviewing the main contributions along these lines in the literature, and organizing this information into major trends and directions. A description of the state-of-the-art in this field, along with supporting simulation results on the image scale-up problem are given.

In the paper titled "Psf recovery from examples for blind super-resolution"[23], the authors propose a new learning based approach for super-resolving an image captured at low spatial resolution. Given the low spatial resolution test image and a training set consisting of low and high spatial resolution images, all captured using the same camera, they obtain super-resolution for the test image. They propose a new wavelet based learning technique that learns the high frequency details for the test image from the training set and thus obtain an initial high resolution estimate. Since super-resolution is an ill-posed problem they solve it using regularization framework. They model the low resolution image as the aliased and noisy version of the corresponding high resolution image and estimate the aliasing matrix using the test image and the initial high resolution (HR) estimate.

In the paper[106], Zhang et al address the problem of producing super- resolved image from a single low-resolution input. Unlike most previous work, the camera's point spread function (PSF) is not assumed to be known in advance and the single image super-resolution problem is formulated as a blind deconvolution problem under a MAP framework which can be optimized effectively in an iterative manner. In the paper, [117] Xiong et al propose a single-image super-resolution scheme for enlarging low quality thumbnail images widely distributed on the Web, which are

often generated by down sampling plus compression. To obtain visually pleasurable high-resolution versions for this kind of low-resolution images, they first adopt a PDE-based image regularization technique to alleviate the compression noise in the distorted thumbnails, and then use learning-based pair matching to further enhance the high-frequency details in the upsampled images. In the paper, [111] Xiaofeng et al develop a scale-invariant representation of images from the bottom up, using a piecewise linear approximation of contours and constrained Delaunay triangulation to complete gaps. They model curvilinear grouping on top of this graphical/ geometric structure using a conditional random field to capture the statistics of continuity and different junction types. The paper titled "Example-based learning for single- image super-resolution" [48] proposes a regression-based method for single-image super-resolution. Kernel ridge regression (KRR) is used to estimate the high-frequency details of the underlying high-resolution image. A sparse solution of KRR is found by combining the ideas of kernel matching pursuit and gradient descent, which allows time-complexity to be kept to a moderate level. The paper[36]titled "Image super resolution as sparse representation of raw image patches" addresses the problem of generating a super resolution (SR) image from a single low-resolution input image. They approach this problem from the perspective of compressed sensing. The low-resolution image is viewed as down sampled version of a high-resolution image, whose patches are assumed to have a sparse representation with respect to an over-complete dictionary of prototype signal atoms. The principle of compressed sensing ensures that under mild conditions, the sparse representation can be correctly recovered from the down sampled signal. They demonstrate the effectiveness of sparsity as a prior for regularizing the otherwise ill-posed super-resolution problem. In the paper "Image super-resolution using gradient profile prior" [39], authors propose an image super-resolution approach using a novel generic image prior gradient profile prior, which is a parametric prior describing the shape and the sharpness of the image gradients. Using the gradient profile prior learned from a large number of natural images, they can provide a constraint on image gradients when one estimate a hi-resolution image from a low-resolution image.

In paper [36] Jianchao et al addresses the problem of generating a super-resolution (SR) image from a single low-resolution input image. They approach this problem from the perspective of compressed sensing. The low-resolution image is viewed as downsampled version of a high-resolution image, whose patches are assumed to have a sparse representation with respect to an over-complete dictionary of prototype signal-atoms. The principle of compressed sensing ensures that under mild conditions, the sparse representation can be correctly recovered from the downsampled signal.

In paper [18] "Super-Resolution from a Single Image", Daniel Glasner et al propose a unified framework for combining classical multi image super resolution and example based super-resolution. They further show how this combined approach can be applied to obtain super resolution from as little as a single image. This approach is based on the observation that patches in a natural image tend to redundantly recur many times inside the image, both within the same scale, as well as across different scales.

In paper [47], Kim et al proposes a framework for single-image super-resolution. The underlying idea is to learn a map from input low-resolution images to target high-resolution images based on example pairs of input and output images. Kernel ridge regression (KRR) is adopted for this purpose. To reduce the time complexity of training and testing for KRR, a sparse solution is found by combining the ideas of kernel matching pursuit and gradient descent. As a regularized solution, KRR leads to a better generalization than simply storing the examples as it has been done in existing example-based algorithms and results in much less noisy images. However, this may introduce blurring and ringing artifacts around major edges as sharp changes are penalized severely. A prior model of a generic image class which takes into account the discontinuity property of images is adopted to resolve this problem

In 2010 [37] Jianchao et al present a new approach to single-image super resolution, based upon sparse signal representation. Research on image statistics

suggests that image patches can be well-represented as a sparse linear combination of elements from an appropriately chosen over-complete dictionary. Inspired by this observation, they seek a sparse representation for each patch of the low-resolution input, and then use the coefficients of this representation to generate the high resolution output. Theoretical results from compressed sensing suggest that under mild conditions, the sparse representation can be correctly recovered from the down sampled signals. By jointly training two dictionaries for the low and high resolution image patches,they can enforce the similarity of sparse representations between the low resolution and high resolution image patch pair with respect to their own dictionaries. Therefore, the sparse representation of a low-resolution image patch can be applied with the high-resolution image patch dictionary to generate a high-resolution image patch. The learned dictionary pair is a more compact representation of the patch pairs, compared to previous approaches, which simply sample a large amount of image patch pairs , reducing the computational cost substantially. The effectiveness of such a sparsity prior is demonstrated for both general image super-resolution (SR) and the special case of face hallucination. In both cases, this algorithm generates high-resolution images that are competitive or even superior in quality to images produced by other similar SR methods. In paper [114] Tang et al propose a method to improve the power of the nearest neighbor based algorithms in single-image based super-resolution, a local learning method is proposed in this paper. Similar to the nearest neighbor-based algorithms, a local training set is generated according to the similarity between the training samples and a given test sample. For super-resolving the given test sample, a local regression function is learned on the local training set. The generalization of nearest neighbor-based algorithms can be enhanced by the process of local regression. Based on such an idea, they propose a novel local-learning-based algorithm, where kernel ridge regression algorithm is used in local regression for its well generalization. Some experimental results verify the effectiveness and efficiency of the local learning algorithm in single-image based super-resolution.

In paper [116], Zheng et al present a new approach to generate a high-resolution (HR) remote sensing image from a single low-resolution (LR) input while

denoising simultaneously, based on sparse signal representation. Recent research on patch-based sparse representation suggests that the high resolution patch has the same sparse representation as the corresponding low resolution patch. Inspired by this observation, they jointly train two dictionaries for the low resolution and the high resolution image patches and enforce the similarity of sparse representations between them. Thus using Batch Orthogonal Matching Pursuit (Batch-OMP), they seek a sparse representation for each patch of the low-resolution input which can be applied with the high resolution dictionary to generate a high resolution patch.

The neighbor-embedding (NE) algorithm for single image super resolution (SR) reconstruction assumes that the feature spaces of low-resolution (LR) and high-resolution (HR) patches are locally isometric. However, this is not true for SR because of one-to-many mappings between LR and HR patches. To overcome or at least to reduce the problem for NE-based SR reconstruction, they apply a joint learning technique to train two projection matrices simultaneously and to map the original LR and HR feature spaces onto a unified feature subspace. Subsequently, the k nearest neighbor selection of the input LR image patches is conducted in the unified feature subspace to estimate the reconstruction weights. To handle a large number of samples, joint learning locally exploits a coupled constraint by linking the LR-HR counterparts together with the K-nearest grouping patch pairs. In order to refine further the initial SR estimate, they impose a global reconstruction constraint on the SR outcome based on the maximum a posteriori framework in the paper "Joint Learning for Single-Image Super-Resolution via a Coupled Constraint" [112] .

In paper [115], Yu Hu et al develop a new face hallucination framework termed from local pixel structure to global image super-resolution (LPS-GIS). Based on the assumption that two similar face images should have similar local pixel structures. The new framework first uses the input low-resolution (LR) face image to search a face database for similar example high-resolution (HR) faces in order to learn the local pixel structures for the target HR face. It then uses the input LR face and the learned pixel structures as priors to estimate the target HR face. They present a three-step implementation procedure for the framework. Step 1 searches the

database, for example the faces that are the most similar to the input, and then warps the example images to the input using optical flow. Step 2 uses the warped HR version of the example faces to learn the local pixel structures for the target HR face. An effective method for learning local pixel structures from an individual face, and an adaptive procedure for fusing the local pixel structures of different example faces to reduce the influence of warping errors, have been developed. Step 3 estimates the target HR face by solving a constrained optimisation problem by means of an iterative procedure. Neighbor embedding algorithm has been widely used in example-based super-resolution reconstruction from a single frame, which makes the assumption that neighbor patches embedded are contained in a single manifold. However, it is not always true for complicated texture structure. In paper[43], Xinbo Gao et al believe that textures may be contained in multiple manifolds, corresponding to classes. Under this assumption, they present a novel example-based image super-resolution reconstruction algorithm with clustering and supervised neighbor embedding (CSNE). First, a class predictor for low-resolution (LR) patches is learnt by an unsupervised Gaussian mixture model. Then by utilizing class label information of each patch, a supervised neighbor embedding is used to estimate high-resolution (HR) patches corresponding to LR patches. In paper [97], Subrahmanyam et al propose a method based on dictionary training, feature extraction from the trained data base images and regularization. They have used singular values as prior for regularizing the ill-posed nature of the single image superresolution problem. Method of Optimal Directions algorithm (MOD) has been used in the proposed algorithm for obtaining high resolution and low resolution dictionaries from training image patches. Using the two dictionaries the given low resolution input image is super-resolved.

Image super-resolution reconstruction has drawn a lot of attentions lately. But almost all existing SR algorithms do not consider the noisy image SR problem. In paper [21], Fang et al propose a novel super-resolution algorithm for noisy images based on sparse mixing estimators. Firstly, sparse mixing estimators are introduced to achieve a directional and sparse representation of noisy low resolution (LR)

image. Then, they employ the median filter to define thresholds using the local characters of the sparse representation. After the noise is removed by shrinkage thresholds, the adaptive interpolations are adopted to achieve high resolution (HR) image.

The search for efficient image denoising methods is still a valid challenge at the crossing of functional analysis and statistics. In spite of the sophistication of the recently proposed methods, most algorithms have not yet attained a desirable level of applicability. All show an outstanding performance when the image model corresponds to the algorithm assumptions but fail in general and create artifacts or remove image fine structures. The main focus of the paper "Self-similarity-based image denoising" [6] is, first, to define a general mathematical and experimental methodology to compare and classify classical image denoising algorithms and, second, to describe the nonlocal means (NL-means) algorithm introduced in 2005 and its more recent extensions. The mathematical analysis is based on the analysis of the "method noise," defined as the difference between a digital image and its denoised version. NL-means, which uses image self-similarities, is proven to be asymptotically optimal under a generic statistical image model.

Various methods for super resolving low resolution images have been presented here. Though they handled super position problem in different context, they are not so successful in reducing aliasing effect, ringing effect etc. Since most of the them use learning methods like Markove method, SVM, they are computationally more complex. So a method which handles defects like aliasing and also which is less complex is desirable. The directionlet based super resolution method presented in this thesis handles all these problems efficiently.

Chapter 3

Single image super-resolution: Freemann's Example based single image super resolution method and Wavelet methods

Levels of details within an image varies from location to location. Some locations contain significant details, where it requires finer resolution for analysis and there are other locations, where a coarser resolution representation suffices. A multi-resolution representation of an image gives complete idea about the extent of the details existing at different locations from which one can choose their requirements of desired details. Wavelet transforms is one of the popular, but not the only approach for multi resolution image analysis. Basic theory of wavelet transform is initially presented. A new block wavelet method to super resolve low resolution images to high resolution images is developed and its performance is compared with the available wavelet transform method which super resolves the image as a whole. The quality of super resolved image using the block method is better than the other method.

3.1 Introduction

As already stated there are two types of super-resolution methods : Reconstruction based and Learning based. Reconstruction based methods super resolve a low resolution image using a number of low resolution images. In satellite communication multiple views of a scene are available and in that situation, reconstruction based super resolution is applicable. But in some cases only a single view of a scene is available and hence other methods are needed for super-resolving the low resolution image. Different types of high resolution images are available today through various means like Internet. One can use these high resolution images to super resolve an entirely different low resolution image. This type of super- resolution methods are called learning based super resolution methods.

Learning based single image super-resolution consists of creating a sharpened version or high resolution image of bigger size from a low resolution image of lower size, using a database created beforehand, called the training set. The basic idea behind learning based super-resolution is to use some reference images (available high resolution images) with the aim of sharpening a low resolution image. These training set reference images are sharp images, containing low, mid and high frequency data components. Some signal processing techniques can be used to decompose these high resolution images into different frequency compositions. These frequency features are stored in the training set. Learning based super resolution means estimation of high resolution image using this training set, when its low resolution image is given. The assumption used is that low resolution image always contains low and mid frequency components and lacks high frequency contents. To obtain high resolution components of a low resolution image, different frequency bands must be known. In the learning based approach the low resolution image itself can be used to obtain the low and mid frequency band and the missing high frequency bands are learned from the training set. Learning based super resolution is considered as the estimation of high frequency bands when low frequency components are given.

In this chapter, three learning based methods are presented. The first one, which was developed by William T.Freeman [108], is considered as one of the initial works in the field of learning based super resolution. Here a study of this method is conducted to analyze the effect of training set on super resolved images. It uses a training set which contains high and mid frequencies of sharp images. These high and mid frequency details are obtained using Laplacian pyramid decomposition and cubic spline interpolation. The basic idea of this algorithm is that interpolation of a low resolution image using standard methods like cubic spline interpolation only increases its size and the resulting image still lacks high resolution frequency part. If the missing high resolution part is obtained somehow, adding it to the interpolated low resolution image gives an image having high resolution. Learning methods can be used to obtain the missing high frequency details. The training set which contains high and mid frequency information of different images is used to learn the correlation between the mid frequency information and those in the database. The high frequency corresponding to the best related mid frequency band is used as the missing high frequency band of input (given) low resolution image. In Freemann's algorithm, Laplacian image decomposition method is used to decompose an image into different frequency bands. A number of other signal processing techniques like DWT, DCT etc are also available for decomposition of an image.

Another method was proposed by Jiji et al [15], where he used discrete wavelet transform to decompose images into different frequency bands at different spatial resolutions. Here the available high resolution images are decomposed into different frequency bands using wavelet transform, and the wavelet coefficients of those bands are stored as the training set or database. Wavelet coefficients at finer scales of the given low resolution image are learnt from this training set. Here the low resolution image substitutes the coarser level. The inverse wavelet transform of the learned coefficients together with the low resolution image (coarser level coefficients) gives the high resolution version of low resolution image. The disadvantage of this method is that it needs different training sets for low resolution

images of different size. This is because wavelet transform of whole image is taken and parent child relationship is used to learn wavelet coefficient at finer scales.

This wavelet method is modified to develop a new method which divides high resolution training set images and their corresponding low resolution images into small patches or blocks and wavelet transform coefficients of these blocks are stored as training set. Hence new method needs only one training set to super resolve low resolution images of different sizes.

3.2 Low resolution model

Digital camera creates an image using a CCD. CCD consists a grid of many tiny light-sensitive cells, or sensors, arranged to divide the total picture area into rows and columns of a huge number of very tiny sub areas. For example, a 3 mega pixel camera CCD has a grid of 2048×1536 sensors (3 million of them). Each sensor samples the color of one of those tiny areas, creating an image of size 2048×1536 pixels.

As explained in chapter 1, the captured image suffers from different degradations. The degradations are optical distortion, blurring, noise, aliasing etc. The degradations occurred in capturing of digital image need to be modeled fully or partially in different SR techniques. The captured image is a degraded version of the original scene and is called low resolution image. Figure 3.1 shows a typical observation model relating the original scene (high resolution) with low resolution image. A low resolution image can be considered as a noisy, down sampled version of the high resolution image which has been blurred and shifted.

Single frame super-resolution algorithms attempt to estimate high resolution image from single low resolution observation. This is considered as an inverse problem. Solving such an inverse problem requires the devising of a forward model that represents the low resolution image formation process. The first step to comprehensively analyze the super resolution image problem is to formulate an

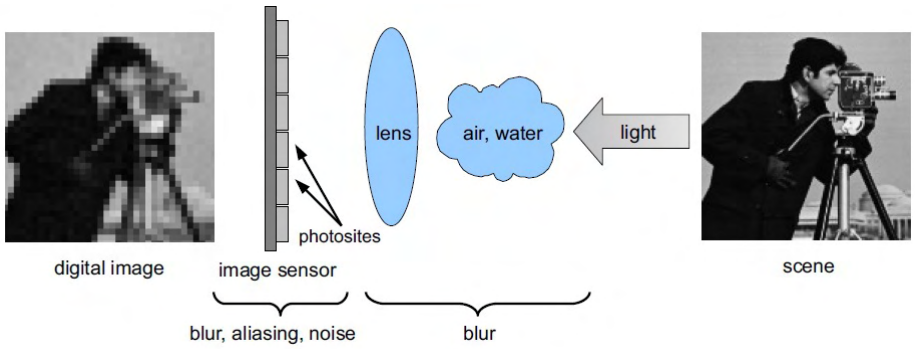


Figure 3.1: Image acquisition system

observation model that relates the original image (HR) to the observed LR image. Usually super resolution problem can be modeled using equation 3.1, given by;

$$Y = DBZ + n \tag{3.1}$$

The low resolution pixel intensity is the average of high resolution intensities over a neighborhood of q^2 pixels, (q is the magnification factor) [15]. Y represents the lexicographically ordered vector of size $M^2 \times 1$, which is formed from the observed low resolution image of size $M \times M$. Similarly Z is the lexicographically ordered vector of the high resolution image to be super resolved. For an integer decimation factor of q, the decimation matrix D consists of q^2 non-zero elements along each row at appropriate locations. Here n is the independent and identically distributed (i.i.d.) noise vector with zero mean and variance σ. It has same size as Y. Here the problem is to estimate Z, the HR image, given Y, the LR image. B is blurred matrix, here it is considered as an identity matrix. Generally, the decimation matrix D, used to obtain the aliased pixel intensities from the high

resolution image is given as;

$$D = \begin{bmatrix} 11..1 & & 0 \\ & 11..1 & \\ 0 & & 11..1 \end{bmatrix} \tag{3.2}$$

For example, with decimation factor $q = 2$ and with lexicographically ordered Z of size, say 16×1, the D matrix of size 4×16 can be written as

$$D = \begin{bmatrix} 1100110000000000 \\ 0011001100000000 \\ 0000000011001100 \\ 0000000000110011 \end{bmatrix} \tag{3.3}$$

Thus equation 3.1 indicates that the low resolution pixel intensity is obtained by averaging the intensities of q^2 pixels corresponding to the same scene in the high resolution image and adding noise.

Low resolution images are modeled using the equation 3.1 for different methods presented in this thesis.

3.3 Background

3.3.1 Multi Resolution Analysis

From a mathematical viewpoint, images are two dimensional arrays of intensity values with locally varying statistics that result from different combination of abrupt features like edges and contrasting homogeneous region. When an image is viewed, connected regions of similar texture and gray level that combine to form objects are visible. Objects with small size and low contrast are needed to be examined at high resolution but coarse view is essential when they are large in size and have high contrast. It will be advantageous to study images at several resolutions, if both small and large objects (low or high contrast objects) are present simultaneously [76]. This is the fundamental motivation for multi resolution processing. A multi

resolution decomposition enables to have a scale-invariant interpretation of the image. The scale of an image varies with the distance between the scene and the optical center of the camera. When the image scale is modified, the interpretation of the scene should not change. A multi resolution representation can be partially scale-invariant. A multi resolution representation provides a simple hierarchical framework for interpreting the image information. At different resolutions, the details of an image generally characterize different physical structures of the scene. At a coarse resolution, these details correspond to the larger structures which provide the image "context". It is therefore natural to first analyze the image details at a coarse resolution and then gradually increase the resolution. Such a coarse-to-fine strategy is useful for computer vision algorithms.

Laplacian pyramid decomposition

Burt and Crowley [13] have introduced pyramidal implementation for computing the signal details at different resolutions. Image pyramid is a powerful and simple structure for representing images at more than one resolution. It is a collection of decreasing resolution images arranged in a shape of a pyramid. In order to simplify the computations, Burt has chosen a resolution step q equal to 2. The details at each resolution 2 are calculated by filtering the original image with the difference of two low-pass filters and by sub sampling the resulting image by a factor 2. This operation is performed over a finite range of resolutions. In this implementation, the difference of low-pass filters gives an approximation of the Laplacian of the Gaussian. The details at different resolutions are regrouped into a pyramid structure called the Laplacian pyramid.

The Laplacian pyramid data structures, as studied by Burt and Crowley, suffer from the difficulty that data at separate levels are correlated. There is no clear model whether a similarity between the image details at different resolutions is due to a property of the image itself or to the intrinsic redundancy of the representation. Furthermore, the Laplacian multi resolution representation does not introduce any spatial orientation selectivity into the decomposition process. This spatial

homogeneity can be inconvenient for pattern recognition problems such as texture discrimination.

3.3.2 Background : Multi Resolution Analysis using wavelets

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. It was introduced by Morlet to overcome the short comings of Fourier transform, which is not a convenient tool for decomposition for image analysis. For many applications a function has to be analyzed in both time and frequency. Though Fourier transform has been considered as the backbone of transform based image processing, wavelet transform is more efficient for applications like image compression, image transmission etc. Fourier transform has sinusoids as basis function. Difference between fourier transform and wavelet transform is that fourier only gives frequency information, smearing time. Using fourier transform it is impossible to tell when a particular event takes place. Wavelet transform can capture both the frequency and time properties of a signal in a single representation. This allows more specific filtering to be done.

For WT, small waves called wavelets of varying frequency and limited duration is used as basis functions. Wavelet analysis is a version of windowing technique, but with a varying window size. It allows the use of longer windows when more precise low frequency information is required, and shorter windows where high frequency information is needed.

Continuous Wavelet Transform

Mathematically, the wavelet, is a function of zero average, having the energy concentrated in time[32].

$$\int_{-\infty}^{\infty} \Psi(t)d(t) = 0 \tag{3.4}$$

The WT of a signal represents the signal as a linear combination of scaled and shifted versions of wavelets and scaling functions. Continuous Wavelet Transform

(CWT) is defined as the sum over all time of the signal multiplied by scaled, shifted version of the wavelet function. Equation 3.5 gives CWT of signal x(t). In the CWT a function Ψ, which in practice looks like a little wave called mother wavelet, is used to create a family of wavelets $\Psi((t-b)/a)$ where a and b are real numbers. Here a is called scale factor and b is called translation factor. The term $1/\sqrt{a}$ is used as the energy normalisation factor across the different scales. The time shifted and time-scaled wavelet $\Psi((t-b)/a)$is sometimes called a baby wavelet. Figure3.2 shows wavelets at different scales. From figure it is clear that scaling a wavelet means stretching (compressing) it. When a>1 the signal gets dilated and a<1 it gets compressed. Figure 3.3 shows translation.

$$C(a,b) = 1/\sqrt{a} \int \Psi((t-b)/a) * x(t) * dt \tag{3.5}$$

This transformation in theory is infinitely redundant, but it can be useful in recognizing certain characteristics of a signal.

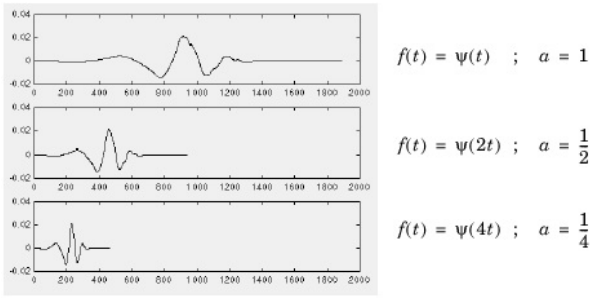


Figure 3.2: wavelets at different scales

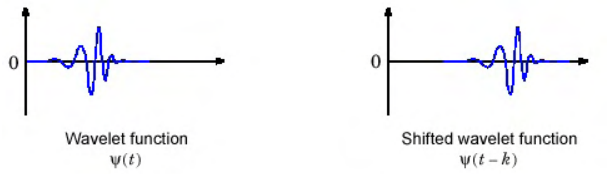


Figure 3.3: Translation

DWT

The CWT maps a signal of one independent variable 't' into a function of two independent variables a and b. The highly redundant nature of this transform makes it inefficient from a computational point of view. CWT provides a redundant representation of the signal, in the sense that the entire support of C(a,b) need not be used to recover x(t). One way to eliminate the problem of redundancy is to sample the CWT on a 2-D dyadic grid. That is, use wavelets only of the form $\Psi(2^j t - k)$ with k and j being whole numbers. The resulting WT is called Discrete Wavelet Transdorm (DWT). DWT of image signals produces a non-redundant image representation, which provides better spatial and spectral localization of image formation, compared with other multi scale representations such as Gaussian and Laplacian pyramid. The following relation gives non redundant wavelet representation, DWT.

$$C(1/2^j, k/2^j) = 2^{j/2} \int_{\infty}^{\infty} x(t)\Psi(2^j t - k)dt \tag{3.6}$$

This is called analysis formula. To simplify, the doubly indexed set of wavelet $2^{j/2}\Psi(2^j t - k)$ is represented as $\Psi_{j,k}$ and equation 3.6 becomes

$$DWT(analysis) : C_{j,k} = \int_{\infty}^{\infty} x(t)\Psi_{j,k}(t)dt \tag{3.7}$$

The recovery of the signal through the synthesis formula is called the Inverse Discrete Wavelet Transdorm (IDWT).

$$IDWT(synthesis) : x(t) = \sum_j \sum_k c_{j,k} \Psi_{j,k} \tag{3.8}$$

Multi resolution Analysis of Wavelets

Concept of multi resolution analysis was introduced by Mallat in their papers [94] and [95]. This concept is used to construct orthogonal bases of wavelets. Multiresolution view can be interpreted as a successive approximation procedure. A signal's approximation at resolution 2^{-j} is defined as an orthogonal projection on a space $V_j \subset L^2(\mathrm{R})$. The space V_j groups all possible approximations at the resolution 2^{-j}. The orthogonal projection of x on V_j is the function x_j that minimizes distance $\|x - x_j\|$. The details of a signal at resolution 2^{-j} are the difference between the approximations at the resolutions 2^{-j+1} and 2^j [32].

Multi resolution approximation

A multi resolution analysis consists of a sequence of successive approximation spaces V_j, $j \in Z$ and satisfy following properties

$$x(t) \in V_j \Longleftrightarrow x(t - 2^j k) \in V_j \tag{3.9}$$

$$V_{j+1} \subset V_j \tag{3.10}$$

The subspace spanned by the scaling functions at lower scales is contained within the subspace spanned by those at higher scales and is given by the following nested relationship,

$$...V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2}... \tag{3.11}$$

This subspace relationship is shown in Figure 3.4

For a given multi resolution approximation $V_{j,j \in Z}$, there exists a unique function $\Phi(t)$, called a scaling function, such that $\Phi_{j,k}(t)$ is an orthonormal basis of V_j. The orthogonal projection on V_j can be computed by decomposing the signal x(t) in the

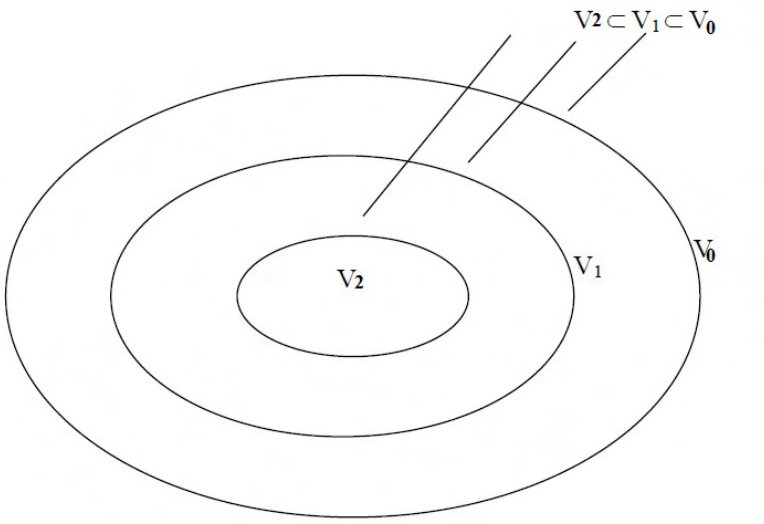


Figure 3.4: Subspace relationship of scaling functions

scaling orthonormal basis [32].

$$x(t) = \sum_{k} \alpha_k \Phi_k(t) \tag{3.12}$$

where k is an integer index of summation, $k \in Z$, the $\alpha(k)$s are the real valued expansion coefficients and $\Phi_k(t)$ represents expansion set.

The Detail Signal

The difference of information between the approximation of a signal x(t) at scales 2^{j-1} and 2^j is called the detail signal at scale 2^j. It is shown in the previous paragraph that the approximations of a signal at scales 2^{j-1} and 2^j are, respectively equal to its orthogonal projection on V_{j-1} and V_j. It can be easily proved that the detail signal at the scale 2^j is given by the orthogonal projection of the original signal on the orthogonal complement of V^j in V_{j-1}, denoted here by W_j. If W_j is

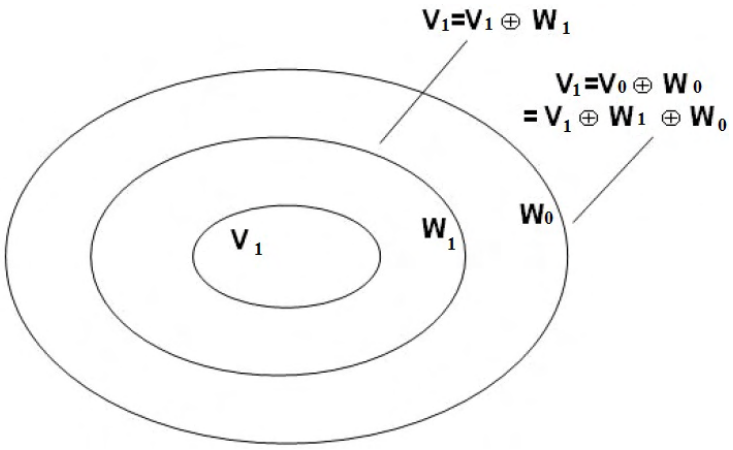


Figure 3.5: Relation between scaling and wavelet functions

the orthogonal complement, then W_j is orthogonal to V_j ; and

$$V_j = V_{j-1} \bigoplus W_{j-1} \tag{3.13}$$

By iterating equation 3.15

$$V_j = W_{j+1} \oplus W_{j+2} \oplus W_{j+3} + ... \tag{3.14}$$

There exists a function $\Psi(t)$, called an orthogonal wavelet, such that,

$$\Psi_{j,k}(t) = 1/\sqrt{2^j}\Psi\frac{t-2^j k}{2^j} \tag{3.15}$$

for any scale 2^j, $\Psi_{j,k}$ is an orthonormal basis of W_j and $\Psi_{j,k}$ is an orthonormal basis of $L^2(R)$, for all scales. The relation between scaling and wavelet function spaces is shown in Figure3.5

If a function x(t) belongs to the subspace V_0 but not V_1. In that case, the scaling functions V_1 of make an approximation of x(t) and the wavelet functions W_1 provide the details. In this sense, the scaling functions analyze x(t) into its low-pass filtered

form and the wavelet functions analyze x(t) into its high-pass filtered form.

Any function x(t) can be expressed as a series summation of scaling functions and wavelet functions as

$$x(t) = \sum a_{j-1,k}\Phi_{j-1,k}(t) + \Sigma_{j=j-1}^{\infty}\Sigma_k b_{j,k}\psi_{j,k}(t) \tag{3.16}$$

where, a and b are the corresponding expansion coefficients.

Filter banks and DWT

Multiresolution analysis allows to decompose a signal into approximations and details. Both approximation and detail coefficients can be obtained by filtering and sub-sampling of the original signal. Filter bank is the building block of discrete-time wavelet transform. When the WT is applied to a signal in the time domain, the result is a two dimensional, time scale domain analysis of the signal. For 1-D signals, two-channel filter bank is shown in Figure 3.6. Filters of different cut off frequencies are used to analyze data here. The data is passed through a series of low pass filter and high pass filter (G_0 and H_0respectively) to separate high and low frequency contents. Each filtered output is down sampled by a factor 2. It is done by removing alternate samples. The results obtained are low pass and high pass signals each containing half as much sample as the input signal. This process halves the resolution and the scale is unchanged. The subsequent down sampling by a factor 2 changes the scale or resolution. In the synthesis part, the outputs $d_1[n]$ and $a_1[n]$ are up sampled by a factor of 2 (upsampling is done by adding zeros in alternate rows and columns) and are passed through two filters G_1 and H_1. The outputs obtained are combined to construct the original data. It is called the one level decomposition using wavelet transform.

Multi level decmposition is possible with wavelets. Computation of DWT through the one level decomposition can be iterated to obtain further analysis of the approximation coefficients $a_1[n]$. Figure3.7 shows 3 level wavelet decomposition tree. Here the signal is denoted by the sequence x[n], where n is an integer. At

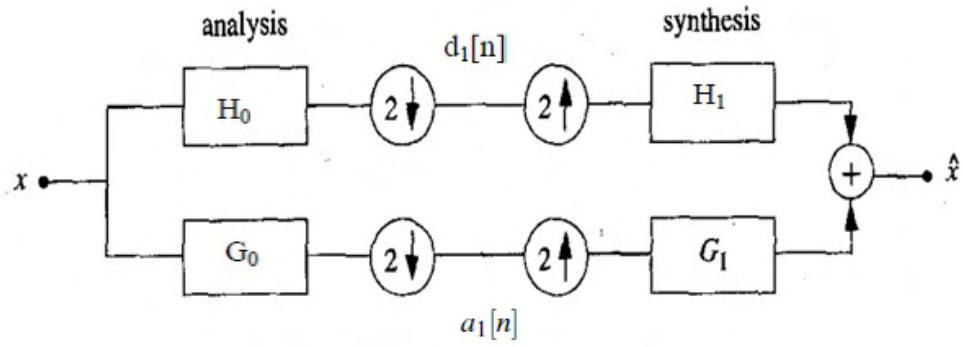


Figure 3.6: Two channel filter bank

each level, the high pass filter produces detail information, d[n], while the low pass filter associated with scaling function produces coarse approximations, a[n]. The

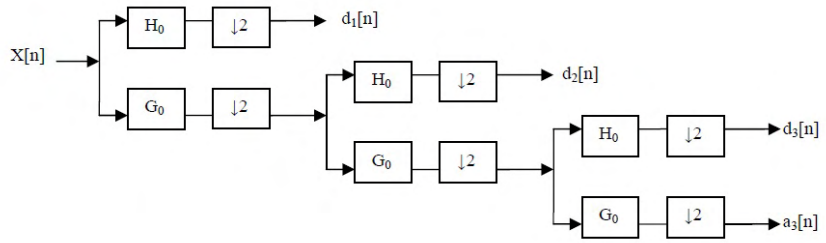


Figure 3.7: 3 level analysis filter bank

signal a[n] obtained by low pass filtering and sub sampling resembles the original signal very much. It can be considered as representation of original signal at low resolution. The low frequency information is considered as a new signal and filter bank can be applied to the new signal.This process can be repeated a number of times. After some iterations very low frequency signal with detail information for different resolution levels are obtained. Figure3.8 shows the reconstruction of the original signal from the wavelet coefficients. Basically, the reconstruction is

the reverse process of decomposition. The approximation and detail coefficients at every level are upsampled by two, passed through the low pass and high pass synthesis filters and then added. This process is continued through the same number of levels as in the decomposition process to obtain the original signal. The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is then obtained by concatenating all the coefficients, a[n] and d[n], starting from the last level of decomposition.

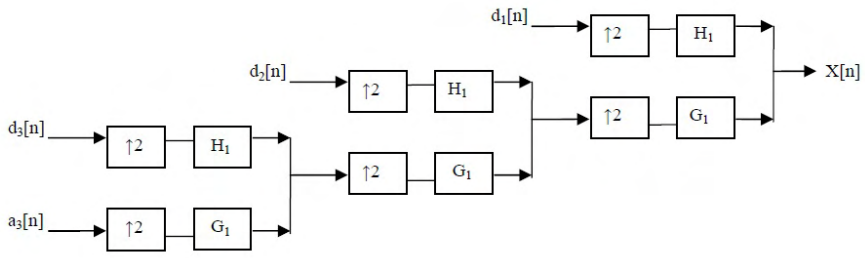


Figure 3.8: 3 level synthesis filter bank

At each decomposition level, the half band filters produce signals spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in frequency is reduced by half. In accordance with Nyquist's rule if the original signal has a highest frequency of $'\omega'$, which requires a sampling frequency of 2ω radians, then it now has a highest frequency of $\omega/2$ radians. It can now be sampled at a frequency of ω radians thus discarding half the samples with no loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. Thus, while the half band low pass filtering removes half of the frequencies and thus halves the resolution, the decimation by 2 doubles the scale. Figure3.9 shows the splitting the signal spectrum with an iterated filter bank.

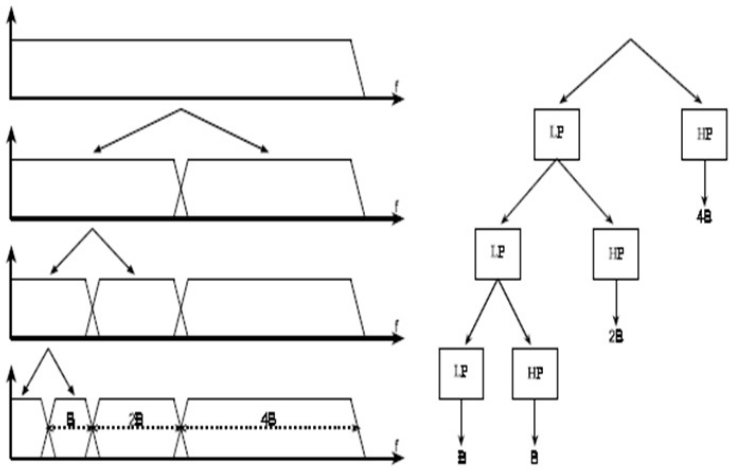


Figure 3.9: Splitting the signal spectrum with an iterated filter bank.

In most Wavelet Transform applications, it is required that the original signal be synthesized from the wavelet coefficients. To achieve perfect reconstruction the analysis and synthesis filters have to satisfy certain conditions. Let $G_0(z)$ and $G_1(z)$ be the low pass analysis and synthesis filters, respectively and $H_0(z)$ and $H_1(z)$ the high pass analysis and synthesis filters respectively. Then the filters have to satisfy the following two conditions as given by;

$$G_0(-z)G_1(z)+H_0(z)H_1(z)=0 \tag{3.17}$$

$$G_0(z)G_1(z)+H_0(z)H_1(z)=2z^{-d} \tag{3.18}$$

The first condition implies that the reconstruction is aliasing-free and the second condition implies that the amplitude distortion has amplitude of one. It can be observed that the perfect reconstruction condition does not change if the analysis and synthesis filters are changed.

2-D Wavelet Transform

The concepts of one-dimensional DWT and its implementation through filter bank can be easily extended to two-dimensional signals for digital images. For two dimensional signals like images, separable WT can be used. Separable WT means a 1-D filter bank is applied to the rows of the image and the same transform is applied to the columns of each channel of the result. Separable wavelet transform needs only one dimensional wavelet transform. Therefore, 3 high pass channels corresponding to vertical V (LH), horizontal H (HL), and diagonal D (HH), and one approximation image A (LL) are obtained. The above procedure is repeated on the low pass channel to obtain different levels. 2D analysis and synthesis filter bank is shown in figure 3.10 and figure 3.11.

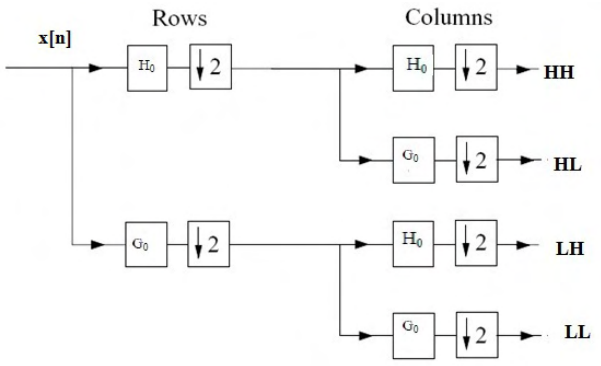


Figure 3.10: 2D WT analysis filter bank

Figure 3.12 shows frequency decomposition of 1-level and 3-level transform. Figure 3.13 shows an image undergone into 3-level decomposition.

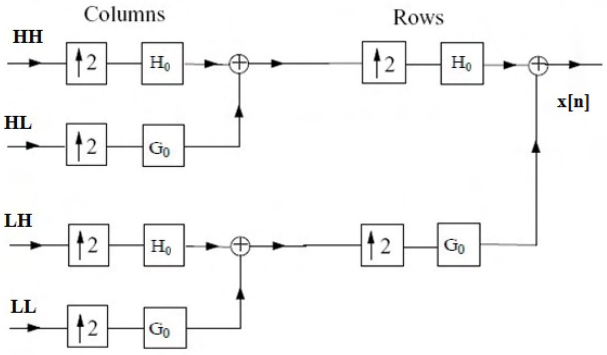


Figure 3.11: 2D WT synthsis filter bank

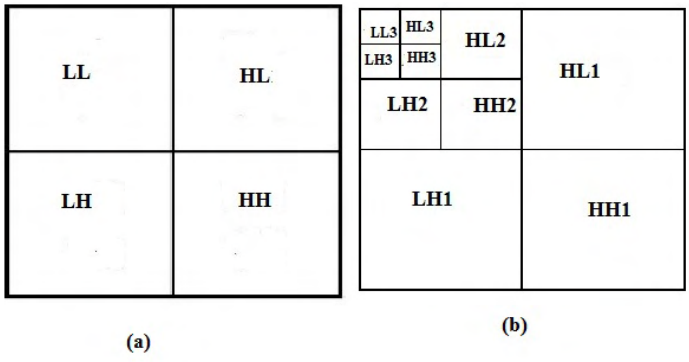


Figure 3.12: Frequency decomposition of (a) 1-level 2D WT (b) 3-level 2D WT

Figure 3.13: WT in 3 level

Classification of wavelets

Wavelets can be classified into two classes: (a) orthogonal and (b) biorthogonal. Based on the application, either of them can be used.

Orthogonal wavelet filter banks

The coefficients of orthogonal filters are real numbers. The filters are of the same length and are not symmetric. The low pass filter, G_0 and the high pass filter, H_0 are related to each other. Also, for perfect reconstruction, the synthesis filters are identical to the analysis filters except for a time reversal. Orthogonal filters offer a high number of vanishing moments.

Bi-orthogonal wavelet filter banks

In the case of the biorthogonal wavelet filters, the low pass and the high pass filters do not have the same length. The low pass filter is always symmetric, while the high pass filter could be either symmetric or anti-symmetric. The coefficients of the filters are either real numbers or integers. For perfect reconstruction, biorthogonal filter bank has all odd length or all even length filters. The two analysis filters can be symmetric with odd length or one symmetric and the other antisymmetric with even length. Also, the two sets of analysis and synthesis filters must be dual.

Wavelet families

There are a number of basis functions that can be used as the mother wavelet for Wavelet Transformation. Since the mother wavelet produces all wavelet functions used in the transformation through translation and scaling, it determines the characteristics of the resulting Wavelet Transform. Therefore, the details of the particular application should be taken into account and the appropriate mother wavelet should be chosen in order to use the Wavelet Transform effectively.

Figure 3.14 illustrates some of the commonly used wavelet functions. Haar wavelet is one of the oldest and simplest wavelet. Therefore, any discussion of wavelets starts with the Haar wavelet. Daubechies wavelets are the most popular

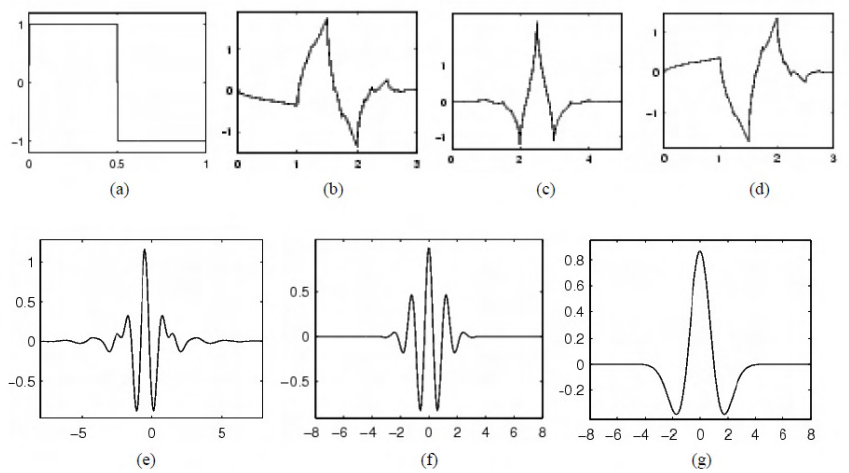


Figure 3.14: Wavelet families (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat

.

wavelets. They represent the foundations of wavelet signal processing and are used in numerous applications. These are also called Maxflat wavelets as their frequency responses have maximum flatness at frequencies 0 and π. The Haar, Daubechies, Symlets and Coiflets are compactly supported orthogonal wavelets. These wavelets along with Meyer wavelets are capable of perfect reconstruction. The Meyer, Morlet and Mexican Hat wavelets are symmetric in shape. The wavelets are chosen based on their shape and their ability to analyze the signal in a particular application.

3.4 Study of Freemann's Onepass algorithm

A learning-based approach which is also called single frame image super resolution is used here for obtaining high quality enlarged images. The algorithm learns the correspondence of different image regions of an input low-resolution image with those in the training set and then uses those learned relationships to predict the fine details of input LR region using the selected HR details from

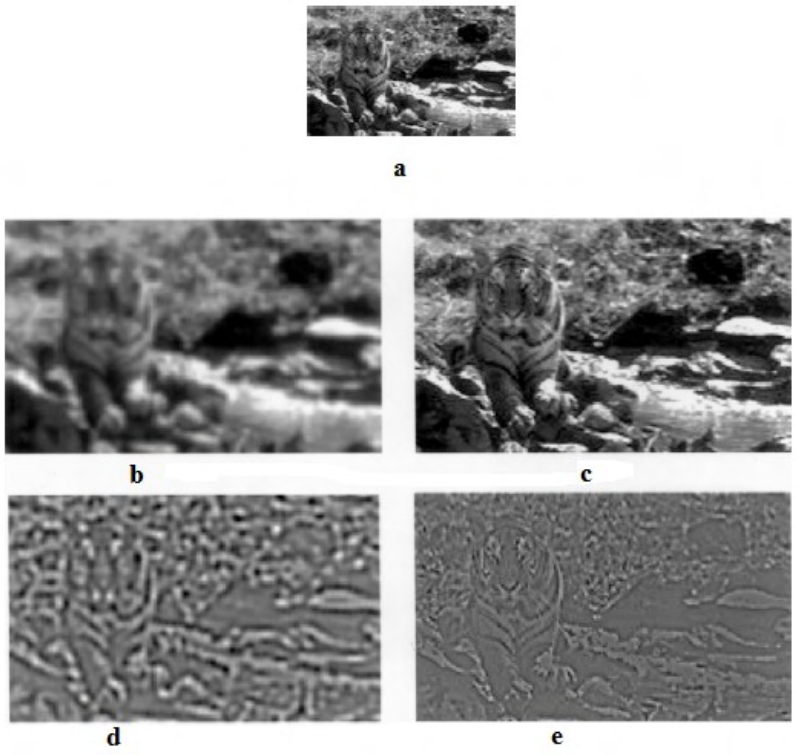


Figure 3.15: a) Low resolution image (b)Cubic spline interpolated low resolution image (c)The original image (d)Midband (e)High frequency band

the training set. Patch based approach is used here. All training set images and input low resolution images are undergone preprocessing for generalization of the different training set images.

SR problem is to estimate the original image Figure 3.15(c) from the low resolution image(a). The low resolution image (a) is the blurred subsampled version of Figure 3.15(c). Figure 3.15(a) is interpolated back up to the original sampling rate to form (b). The missing high frequency detail, (c) minus (b), is the high frequency to be estimated,(e). Two pre processing steps are taken for

efficiency: the low frequencies of (b) are removed to form the input mid band and high frequency bands are contrast normalized by the local contrast of the input band, yielding (d) and (e).

Freemann's one pass algorithm to super resolve a low resolution image involves three steps : 1.Pre-processing 2. Training set generation 3. Learning. These steps are explained below.

Preprocessing

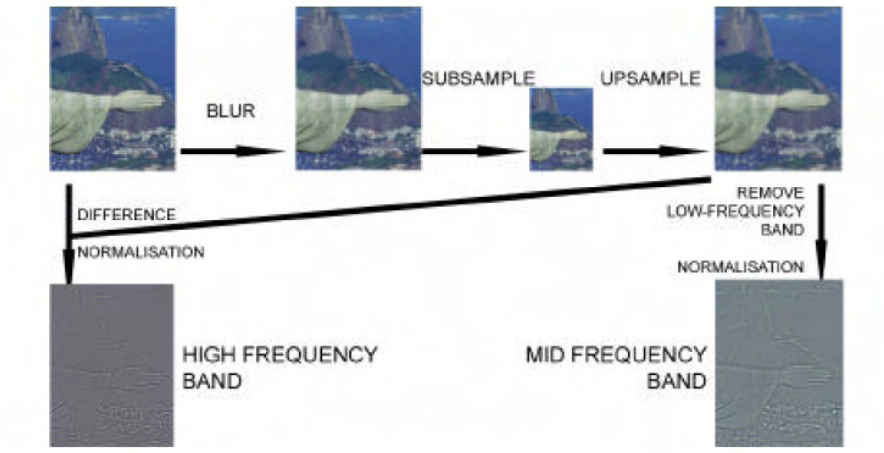


Figure 3.16: Preprocessing steps and formation of mid and high frequency bands

All the high resolution training set images and the input low resolution image have to be pre-processed before they have undergone frequency decomposition. To increase the efficiency of the training set, two pre-processing steps are done. The first step is based on the assumption that images can be represented using a bandpass image representation, such as the Laplacian pyramid image decomposition, where they can be divided into different frequency bands. Let H be the high spatial frequency pixel values, and M be the values of the next-highest spatial frequency band, which will also be called the mid-frequency band, and L be

the pixel values of all lower spatial frequencies in the image. The highest resolution frequency band is conditionally independent of the lowest frequency band, L, but only depends on second highest resolution frequency band(mid frequency band). Based on this assumption, to predict the highest frequency band, only the mid-frequency band, M is needed. This greatly reduces the variability of data stored in the training set.

The next pre-processing step is based on the assumption that the relationship between the high and mid frequency bands are independent of the image contrast. Contrast normalization does not affect the learning process. This saves replication of the training set for all possible values of image contrast. So the high and mid frequency bands are contrast normalized. The resulting high pass filtered and contrast normalized band pairs are used for training. The contrast normalization step is undone when the high-resolution image is reconstructed.

The above mentioned steps are shown in Figure 3.16. In this figure a high resolution image is blurred and subsampled to obtain a low resolution image. The LR image is cubic spline interpolated back to the original resolution. Though this interpolated image has the same size as that of the original high resolution image, it lacks high frequency details. It contains only low and mid frequencies. It is high pass filtered to obtain the mid frequency band. Also, the difference between original HR image and the interpolated image gives high frequency details of the HR image (high frequency band). Both high and mid frequency bands are contrast normalized. They form the elements of training set.

Training Set generation

After the pre-processing steps, high frequency band of high resolution image and mid frequency band of its corresponding low resolution image are obtained. The mid frequency band and high frequency band of training set images are then divided into small patches and are stored as the training set. The patch size can not be too large and can not be too small. The local image patch would not give sufficient information for analyzing the underlying scene variable when the patch size is too

small. But, if the patch size is small it can be easily stored in the database. On the other hand, a large patch size would solve the above problem by disambiguating the underlying scene variables, but it would take prohibitive amount of time to learn the relationship between local image and scene patches. Also the memory requirement increases exponentially as the size of image patches increases. As a compromise, patch size must have enough dimension to give some useful information about the underlying scene, and must be small enough to allow quick learning of the relationship between low resolution and high resolution patches. The vector for a patch pair in the training set is obtained by the concatenation of the mid frequency patch and the region in the corresponding high frequency band which overlaps with the neighboring left and top patches. It is not required that the high frequency and mid frequency patches are of same size but they must be center aligned. That means they must have same pixel as center. High resolution patches are taken with one pixel overlap to preserve the spatial continuity or neighborhood effect.

Learning Using Single pass algorithm

The single-pass algorithm predicts the missing high band of a cubic spline interpolated input low resolution image, sequentially. The given input image, once preprocessed is broken into patches. Then it is scanned in a raster scan order predicting at every step the high resolution patch as shown in Figure 3.17.

A block diagram representation of one pass algorithm is shown Figure 3.18. To obtain the input search vector, mid frequency patch corresponding to the input low resolution image is lexicographically arranged and concatenated with the overlap pixels from previously determined neighboring top and left high resolution patches. The search vector is contrast normalised as shown in figure and is used to find the best match from the training set. From the training set mid frequency patches with the same variance as the input mid frequency patch are to be selected. Matches are searched using L2 norm. Finding the absolute best match considering all the training set vectors would be a tiresome process due to the high dimension of the search space of training set. Hence a binary tree-based approach is used to find the approximate nearest neighbor.

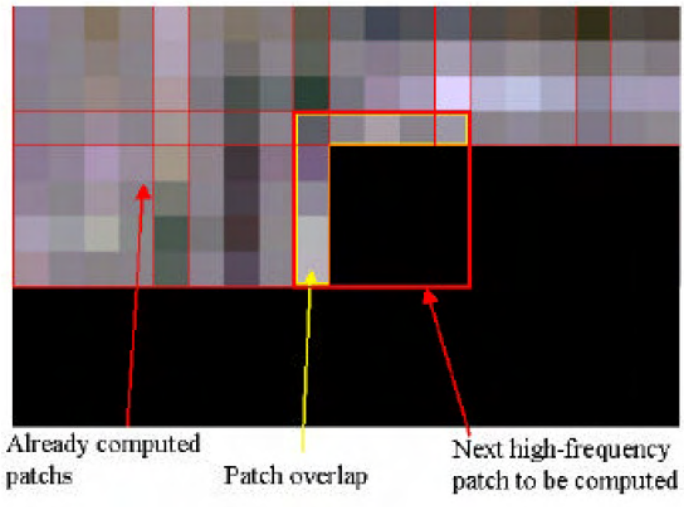


Figure 3.17: High frequency band patch overlap

The tree is built by recursively splitting the training set in the direction of higher variance. At each step, the set of tiles is divided into half based on variance, to maintain a balanced tree. Using this tree, a set of mid frequency vectors with same variance as the input mid frequency patch vector is chosen. Then by checking the overlap pixels using the L2 norm the mid frequency patch most similar to the input patch is selected. The high frequency patch corresponding to the selected mid frequency patch can be used as the missing high frequency patch of the input low resolution patch. That is, the missing frequency patches are selected based on the local mid frequency details and previously determined adjacent high frequency patches. When a match is found, the contrast normalization is undone on the high frequency patch. This is repeated for all mid frequency patches. The reconstructed high frequency band is added to the up sampled input image to obtain high resolution image.

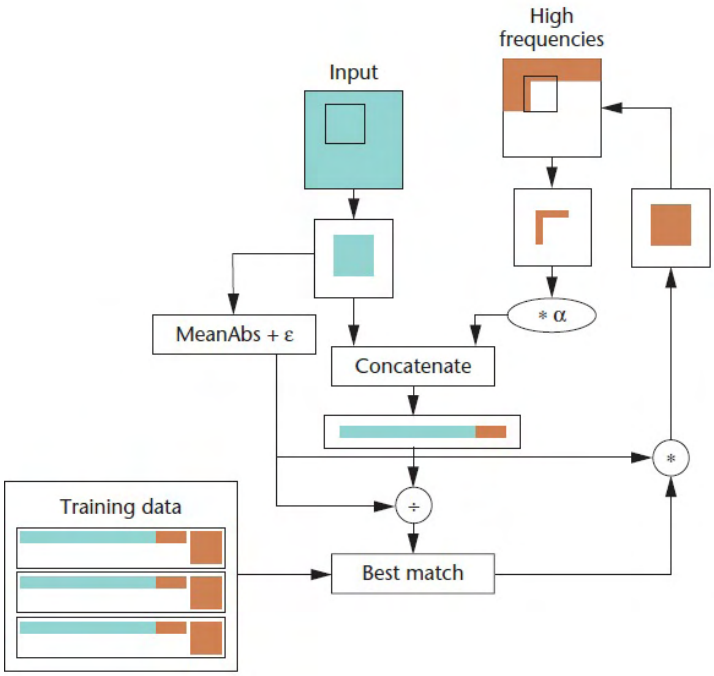


Figure 3.18: Block diagram representation of Freemann's method

3.4.1 Implementation of Freemann's method

Details of images used for training and learning

Images having different sizes like 256x256, 512x512, 230x333 etc represented by 256 grey levels are downloaded from the Internet. They are of Tiff format. All the works presented in this thesis including this chapter uses a magnification factor of 2 and 4. Images of size NxN is super resolved to an image of size 2Nx2N and 4N*4N.

Implementation of the method

The implementation of this method is done to study the effect of the images used in the training set on super resolved image and to compare Freeman's super

resolved image to super resolved images by other methods presented later portion of this thesis.

Input images used

For the purpose of evaluation of efficiency of any algorithm, one has to compare the reconstructed image with its actual original image. Hence in this study, low resolution images are generated from the available high resolution images. The input low resolution images of size N x N are obtained by blurring and subsampling the high resolution images of size 2Nx2N.

Training set generation

The low resolution (LR) images of the high resolution (HR) images are obtained by convolving them with a blurring (low pass) filter with coefficients [0.25 0.5 0.25] and then sub sampling (removing pixels alternatively) the resulting blurred image. The resulting image is then up sampled by cubic spline interpolation method to obtain the same size as the original high resolution image. To separate the mid frequency band from the cubic spline interpolated image, the lowest frequency band is removed, by applying a high pass filter with coefficients [-1/9 -1/9 -1/9;-1/9 +8/9 -1/9; -1/9 -1/9 -1/9]. The difference between original high resolution image and cubic spline interpolated low resolution image is the high frequency content present in the sharp image. After these preprocessing steps, two bands are obtained - mid frequency band and high frequency band. Spatially corresponding $M \times M$ mid frequency and $N \times N$ high frequency patches are taken from mid and high frequency bands respectively. M=7 and N=5 were used for implementation. Patch pairs are both contrast normalized by energy of the mid frequency patch. As explained earlier, they are center aligned, so that the image patch centered at pixels (i, j)covered all pixels (i$\pm$3, j$\pm$3) of mid frequency band and the corresponding high frequency patch covered all pixels (i$\pm$2, j $\pm$2). The search vector for a patch pair is created by concatenating the mid-frequency patch of size 7×7 and the one pixel width overlap region from the neighboring high-frequency patches as mentioned in previous section. The overlapping region is adjusted by the weighting factor α. α is a controlling parameter which balances the matching between the

mid resolution patch data and the predicted overlapped pixels and helps to find a high-resolution patch that is compatible with its neighbors. This weighting factor compensates for the different relative areas of the mid-frequency patches and overlapped high-frequency pixels as a function of M and N[108]. The value of α is determined by the equation,

$$\alpha = 0.1 * \frac{M^2}{2N-1} \tag{3.19}$$

Learning Of High Frequency patches

The input LR image is interpolated and high pass filtered to get the mid frequency band. The mid frequency band is divided into patches of size 7×7, in raster scan order, and contrast normalised. Each patch is compared with mid frequency patches in the training set to get the best match (as mentioned in section 3.4.3). The corresponding high frequency patch of size 5×5 is selected, and contrast normalisation is undone by the energy of the corresponding mid frequency patch. Care is taken to see that the center of high frequency patch coincides with that of the mid frequency patch. This is repeated for all the patches in the input mid frequency band, where consecutive mid frequency patches overlap by . The learned high frequency band is added to the upsampled input LR image (which contains both low and mid frequency bands) to obtain high quality super resolved images. It is compared with original image to study the performance of the algorithm.

3.4.2 Results and discussion

One of The training set image used for this work is shown Figure 3.19. It is of size 644x800 and in tiff format.

Figures 3.20(a) and (c) show high resolution images of size 480x640 and 256x256 respectively and their corresponding low resolution images of size 240x320 and 128x128 respectively are created as mentioned in section 3.3.3. Figure 3.21 and Figure 3.22 show reconstructed images using Freeman's method. Figure3.21(b) and Figure 3.22(b) show original high resolution images, Figures 3.21(c) and 3.22(c)show cubic spline interpolated images, Figures 3.21(d) and

Figure 3.19: Training set image

3.22(d)show learned high frequency bands and Figure3.21(e) and Figure 3.22(e) show super resolved images. The original images are shown here to compare the super resolved images. When the learned high frequency band is added to the cubic spline interpolated image, the final SR images become better, but still there are artifacts as can be seen in the region of hair of 'Lena' in 3.22(e).

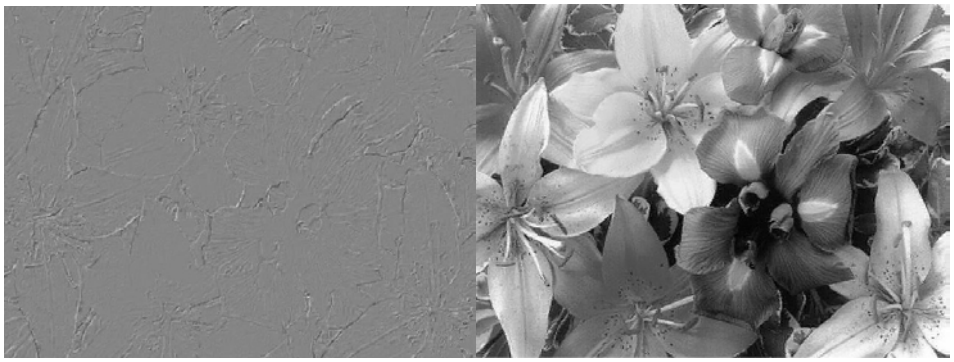
(a) (b)

(c) (d)

Figure 3.20: (a),(c)Original images(b),(d)Their low resolution images

(a)

(b) (c)



(d) (e)

Figure 3.21: (a)Low resolution image (b)Original image(c)Cubic spline interpolated image(d)Learned high frequency content(e)super resolved image

(a)

(b) (c)

(d) (e)

Figure 3.22: (a) Low resolution image (b)Original image(c)cubic spline interpolated image(d)learned high frequency content(e)super resolved image

Effect of training set on Super-resolved image

To study the effect of various types of images used to create the training set images, in the quality of super resolved images, training set images are grouped into three groups group1, group2 and group3. They are used to evaluate the super resolution quality as mentioned below. These groups contain images with low, medium, high information content respectively. A few of images in the 3 groups are shown in Figure 3.23. Signal to noise ratio (SNR) values are used as an initial check on performance of algorithm. Signal to Noise Ratio (SNR) is calculated using the equation;

$$SNR = \sum_{i,j} \frac{{z_{i,j}}^2}{\sum_{i,j} [z(i,j) - z'(i,j)]^2} \tag{3.20}$$

where $z(i,j)$ and $z'(i,j)$ are the $(i,j)^{th}$ the pixel intensity of the original image and reconstructed image respectively.

Table 3.1: SNR values obtained with training set containing different images

Input Images)	SNR(dB) obtained with training set images containing					
	low1	low2	mid1	mid2	high1	high2
low1	infinity	32.2698	33.2365	33.3307	37.0712	34.0803
low2	29.2651	infinity	30.4137	30.4419	33.2799	31.0253
mid1	30.5833	29.5835	infinity	30.6933	34.1935	32.0169
mid2	28.5607	28.5149	30.3485	infinity	33.6794	31.5954
high1	32.6133	31.4813	33.9488	34.357	infinity	35.1812
high2	26.0506	28.9804	26.7547	28.8065	29.9141	infinity

Table 3.1 shows the SNR values obtained when different training sets are used. Row 1 shows the effect of super resolving the image low1 of group 1. SNR obtained is 32.2698dB when low2 is used as the training set image. Using images with higher information content from group2 and group3 from the training set, it is seen that the SNR is increased to 33.3307dB to 37.0712dB for input image low1. The quality of images is independent of the class of the training set images. The

same effect is noticed when group 2 and 3 images are super resolved. Thus quality of super resolved image increases as with the quality of images in the training set.

(a) low1 (b) low2 (c) low3

(d) med1 (e) med2 (f) med3

(g) high1 (h) high2 (i) high3

Figure 3.23: Training set images-(a),(b),(c)low information content images(d),(e),(f)medium information content images(g),(h),(i) high information content images

From Table3.1, it is clear that the quality of the super resolved image depends on the training set. The training set with content rich group 3 images (high1 and high2) has more effect than the images from other two groups. It is obvious that group1

low information images (low1 and low2) produce low SNR values compared to that of other group images. As the information content of training set images increases the SNR values and the quality of super resolved images increases. The image high1 from group3 of the training set gave the best result for all the other images, compared with other training set images. So one can use carefully selected high information content image instead of using a number of low information content images. The size of training set can be reduced by selecting appropriate high information content image as training set image. For the method proposed in the fourth chapter carefully selected single image is used for training set formation.

It is also clear that quality of super resolved image is independent of class of images used in the training set. This means that the training set images need not be of same class of input low resolution image.

3.5 Wavelet methods

Wavelets are very good tools for multi resolution analysis. They enable analysis of data at multiple levels of resolution or scale. This means that the view of image at different "scales" or "resolutions" is obtained using wavelet transform. A rough approximation of the signal might look stationary, while at a detailed level discontinuities become apparent [26], [41].

This section describes two wavelet transform based methods. The first method was proposed by Jiji et al [15] and it is also based on the assumption that images can be decomposed into low and high frequency bands. Here decomposition process is done using wavelet transform. The super resolution is considered as a problem of obtaining high resolution image, if the low band is given. For that one has to first estimate the high frequency bands, so that using inverse wavelet transform of these bands gives the original signal.

To super resolve a LR image using wavelet transform, the different frequency bands (LL, LH, HL, HH) corresponding to high resolution image need to be known.

The input LR image itself is considered as approximation or low frequency band (LL) of the unknown HR image. For obtaining high frequency bands (LH, HL, HH), a training set which contains wavelet coefficients of available HR images is used here. Three-level wavelet decomposition of the high resolution images are taken and are used as the training set. Two level decomposition of input low resolution image is taken. The absolute difference between first and second level wavelet transform coefficients in the LR image and second and third level coefficients for each of training set images is taken. For comparing the coefficients, parent child relationship is used here. If a match is found, corresponding first level coefficients of high resolution image are taken as the first level coefficients of unknown high resolution image. The HR image is obtained by taking the inverse wavelet transform of the approximation and learned wavelet coefficients. In order to obtain a spatial coherence during the HR reconstruction, a smoothness constraint is needed. The wavelet method without smoothing is implemented here.

It is found that the above explained wavelet-based method has some disadvantages. One problem is that it needs regularization to bring spatial coherence. Another problem is that it is highly resolution dependent. Since it follows parent child relationship, using one training set low resolution image of fixed size can be super resolved. To super resolve an image of different size, another training set is needed. This problem is solved by modifying the first method using block based approach. But this method also fails to eliminate aliasing effects.

3.6 Single image super resolution using Wavelet Transform

Two methods are considered in the following section to super resolve a LR image using wavelet transform.

3.6.1 Jiji et al's Single Frame Image Super resolution Using Learnt Wavelets: Wavelet method-1

Methodology

Here super resolution method [15] is implemented.(Regularization step to smoothen the resultant image is not implemented here)

Wavelets can be used in super resolution problem, since super resolution involves processing of data at different resolutions. Here the concept used is that any image can be decomposed using wavelet transform into low frequency band LL and high frequency bands LH, HL, HH using discrete wavelet transform. Now the problem is to generate the original HR image, given the low frequency band LL (LR image). To generate the original image, the high frequency bands have to be found out and inverse wavelet transform of these bands (LH, HL, HH) and the known low frequency band (LL) is to be taken to obtain the original high resolution image. Then the actual problem is finding out the missing high frequency bands. In learning based super resolution method, these missing bands are learned from a training set.

Training set generation

The training set contains three level wavelet coefficients of different high resolution images. Here parent child relation(zero tree concept) is used to find out the missing high frequency details.

Contrast normalised high resolution images are subjected to 3 level decomposition using wavelet transform. The decomposition coefficients are used as the training set. The concept of parent child relation is that in a multi resolution system, every coefficient at a scale can be related to a set of coefficients at the coarser scale having similar orientation [29]. The parent-child relationship shows that a 4 x 4 pixel area in the first level decomposition of a high resolution image has a related set of elements of size 2 x 2 pixel area in the next level and a single corresponding element in the third level as shown in Figure 3.24. If the

magnification factor is two, then each training set image must have size 2M x 2M, for a low resolution image of the size M x M. Here a training set images of size 256 x 256 are used to super resolve images of size 128 x 128.

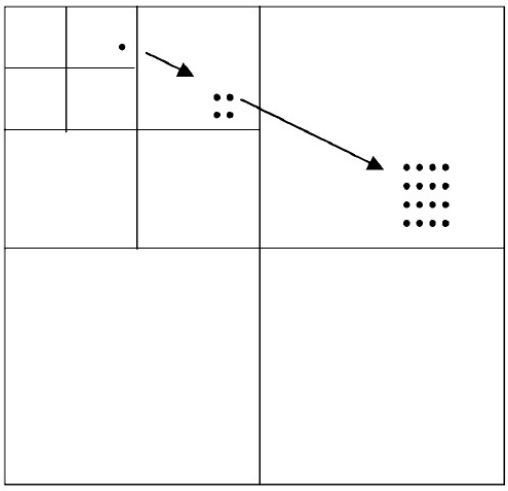


Figure 3.24: Parent child relation

Learning Wavelet coefficients

The input low resolution image has undergone two level decomposition. The low resolution image is formed using the equation 3.1. Figure 3.25 shows how the missing high resolution coefficients in blocks VII, VIII, IX are learnt from a set of N training images. The second (IV, V, VI) and third (I, II, III) level coefficients of training set images and the two level coefficients of the test input image are compared. As shown in Figure 3.25, for each coefficient in I, II, III sub bands and the corresponding 2x2 blocks in the sub bands IV to VI of the test image, a block of 4x4 wavelet coefficients in each of the high frequency sub bands VII, VIII, IX can be learned from the training set using zero tree concept. The absolute difference between the wavelet coefficients in the low resolution input image and the corresponding coefficients in each of the high resolution training set images

are taken to select the best matching image, using Minimum Absolute Difference (MAD) criteria. The high frequency coefficients of this related training set image gives the missing wavelet coefficients of the input test image. The lowest sub band '0'corresponding to LL, the low resolution portion is not used in the process since they have different brightness averages. The process is repeated for each coefficient in the sub bands in I, II, III to obtain the coefficients in the VII, VIII,IX sub bands of low resolution image. The inverse wavelet transform of these learned bands together with low resolution image as approximation or LL gives the original high resolution image.

The wavelet based method can be summarized as follows:

1. The training set high resolution images of size 2M x 2M and input low resolution image of size MxM undergo three level and two level decomposition respectively using DWT.

2. Absolute difference between I^{st} and II^{nd} level wavelet coefficients of low resolution image and the corresponding coefficients of each of the training set images is taken.

3. 4x4 third level coefficients of selected training set image with Minimum absolute difference(MAD) is used as the first level 4x4 high resolution wavelet coefficients of the unknown high resolution image.

4. This process is repeated for every coefficient in band I of low resolution image.

5. The inverse discrete wavelet transform of learned high frequency wavelet coefficients of bands VII, VIII, IX and the low resolution image itself(approximation or low frequency band) gives the high resolution image.

.

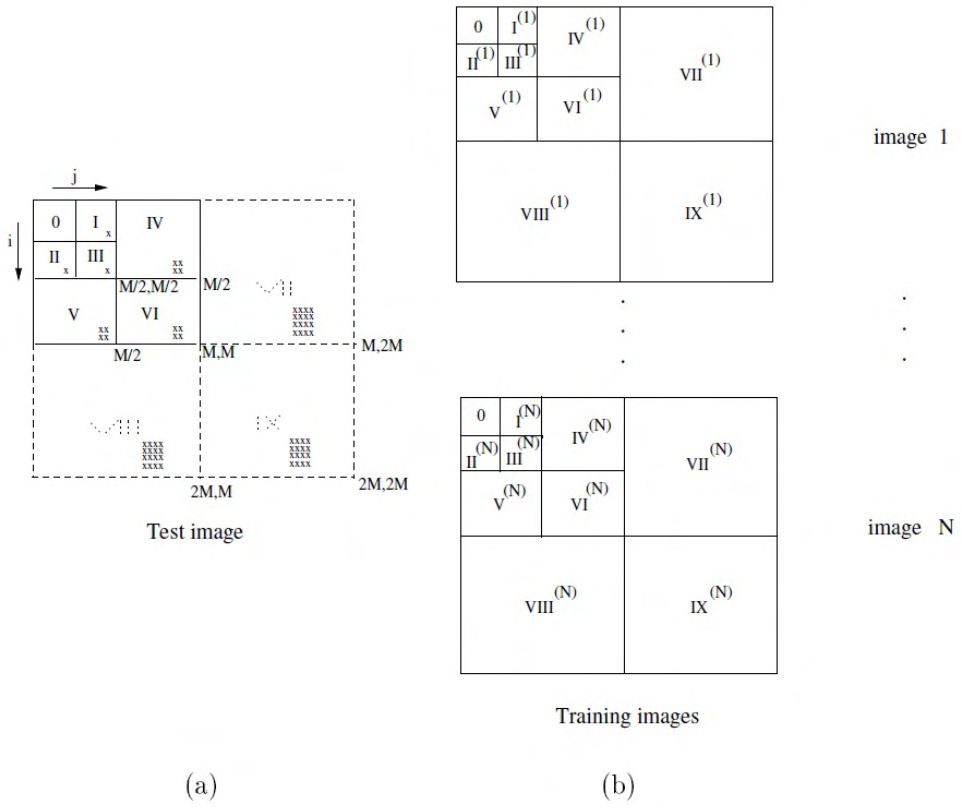


Figure 3.25: Learning of wavelet coefficients

Implementation

A number of different high resolution images downloaded from Internet is used here as training set images. Training images of size 256×256 and 256 grey levels are used here. LR image is obtained from HR image by averaging every non overlapping 2x2 pixels (here magnification factor q=2) in raster scan order (using the equation 3.1). Using this training set one can super resolve low resolution images of size 128x128 to obtain the SR image of size 256x256.

3.6.2 Single image super resolution using learned wavelets-Block wavelet method

Methodology

The method explained above needs different training set for super resolving low resolution images of different sizes. For example, to super resolve an image of size 128×128 into 256×256, one needs training set images of size 256×256. But, training set images of size 128×128 are needed to super resolve low resolution images of size 64×64, the 256×256 images can not be used. Another problem is that due to spurious learning, super resolved images have artifacts. So smoothness constraint is needed to solve this problem, which increases computational complexity of the method. Hence wavelet method I is modified using patch based approach to obtain block wavelet method. In this new method, single training set with suitably selected high resolution images is used to super resolve different sized input low resolution images.

In the block wavelet method, which is a novel approach, instead of applying wavelet transform to the full image, high resolution images and their low resolution images are divided into small overlapping blocks. Both high and low resolution patches are contrast normalized using energy of the input low resolution patch and 1-level wavelet transform is applied to these patches. The wavelet coefficients of these patches are stored in the training set. The LR image to be super resolved is also divided into patches and wavelet coefficients of these patches are compared with those of the patches in the training set. The training set patch with minimum absolute difference is found out and the wavelet coefficients of its HR patch are used as the HR coefficients of the input LR patch. Here also, the LR patch is considered as the approximation. Contrast normalization is undone at the end of the entire reconstruction process. The process is repeated for all the patches in the LR image in raster scan order to obtain the unknown HR image.

Implementation

HR images of size 256×256 and 256 grey levels are used for training. LR images are obtained from the corresponding HR images as mentioned earlier, for

the purpose of evaluation. The high resolution images and their corresponding low resolution images are divided into small overlapping patches of size 8×8 and 4×4 respectively. One level wavelet transform is applied on these patches and coefficients corresponding to the high frequency bands H, V and D of high and low resolution patches are stored as the training set. The input low resolution image is divided into small overlapping patches of size 4×4 and one level wavelet transform is applied to each patch. To find the missing high frequency bands of its unknown high resolution patch, wavelet coefficients of low resolution patch is compared with the corresponding coefficients of low resolution patches in the training set, by taking the absolute difference. The low resolution patch with minimum absolute difference is selected and it's corresponding HR high frequency coefficients are used as the missing HR coefficients. This is repeated for all the patches. The low resolution patch was used as the approximation. The inverse transform of the approximation and learned coefficients corresponding to H, V, D gives the unknown HR patch.

Implementation of Sapan et al method "Single image super resolution in spatial and wavelet domain"

In this method [86] single image super resolution algorithm is proposed which uses both spatial and wavelet domain. For up sampling and down sampling of an image in spatial domain, respectively bicubic smoother and bicubic sharper method of Adobe Photoshop cs5 is used. For removing blur and get smoother result back projection method is used. Wavelet based denoising method is also used to remove noise from image.

3.6.3 Results and Discussion

Figures 3.26(a), (b) show low resolution images which are obtained from original high resolution images of butterfly and Lena shown in Figures3.26(c) and (d).

(a) (b)

(c) (d)

Figure 3.26: (a),(b)Low resolution images(c),(d)Original images

Table 3.2: SNR values with different methods

Method	SNR in DB		
	Barbara	Butterfly	Tiger
Freeman method	13.3849	22.3943	19.9744
Wavelet method 1	8.7688	3.1185	3.6367
Sapan et al's wavelet based method	16.1121	20.9643	19.1132
block Wavelet method	17.4133	24.8016	24.1414

SNR values of different methods are shown in table 3.2. From table it is clear that block wavelet method is better than other super resolution methods. For the image Barbara SNR with new block wavelet method is 17.4133dB, while it is low for other methods.

The results are shown in Figure 3.27. Here Figure 3.27(a)low resolution image (b)original image. Figures3.27(c), (d), (e) and (f) are super resolved images using Freeman method, wavelet method 1, super resolved image using Sapan et al wavelet based super resolution method [86], block wavelet method.

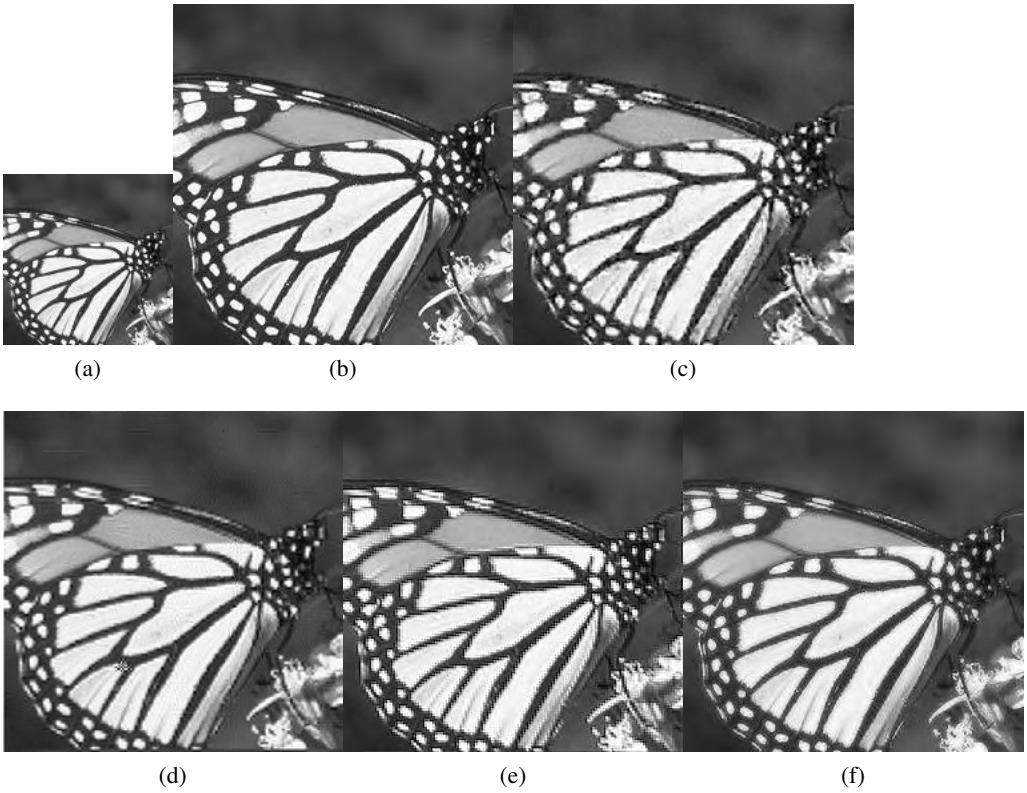


Figure 3.27: (a)low resolution image (b)Original image (c)Freeman method (d)Wavelet method 1 (e)Sapan's wavelet method (f)Block wavelet method

Another result is shown in Figure 3.28. Here Figure 3.28(a)is the low resolution image, (b)original image. Figures3.28(c), (d), (e) and (f) are super resolved images using Freeman method, wavelet method 1, Super resolved image using Sapan et al wavelet based super resolution method, block wavelet method. It is clear that

the block wavelet method removes artifacts much better compared with the super resolved images. There is no need of regularization step. There is no constraint about the size of training set images and SR images of any size can be super resolved using this method.

(a) (b) (c)

(d) (e) (f)

Figure 3.28: (a)Low resolution image (b)Original image (c)Freeman method(d)wavelet method 1(e) Sapan's wavelet method (f)block wavelet method

3.7 Conclusion

This chapter explains three single image super resolution methods : Freemann's single image super resolution method and two single image super resolution methods using wavelet transform. The first method uses Laplacian pyramid

decomposition and cubic spline interpolation to prepare the training set and one pass algorithm for learning process. From the results it is clear that quality of resultant images depends on training set images used. Images with high information content can produce better result than less information image. One single high information content image can give good results. So instead of using a group of less information content images, carefully selected information rich single image is sufficient.

The other methods are based on wavelet transform and block wavelet method is a modified form of wavelet method 1. In the first wavelet method coefficients of finer scales are learned from a training set, which contains wavelet coefficients of different high resolution images. The learning process is done according to parent child relationship. This method produces artefacts in the super resolved images. This method needs different training sets for different sized low resolution images. The block wavelet method over comes this limitation by using patch based approach. Some artefacts are removed in this method. Sapan et al method, a single image super resolution algorithm is implemented here and it is based on both spatial and wavelet domain and take the advantages of both. All these methods fail to remove aliasing. A better solution can be obtained with directionally adaptive super resolution method using multiple direction wavelet transform. Next chapter describes a new directionally adaptive method using directionlets.

Chapter 4

Single Image Super resolution using Directionlets

In this chapter a novel directionally adaptive single image super-resolution method based on a multiple-directional wavelet transform, called directionlet transform is, presented. The directionlet transform, a new lattice-based perfect reconstruction transform retains the separable filtering and simple filter design from the standard two-dimensional discrete wavelet transform [102]. The new algorithm using directionlet transform efficiently captures directional features and extract edge information along different directions, from high resolution images. This directional information is stored in a training set and is used to super-resolve a low resolution image which is entirely different from training set images. This new algorithm outperforms the state-of-the-art methods in terms of both numeric and visual quality of the super-resolved image. This work is called directionally adaptive single image super-resolution because patch wise approach is used here and prominent edge directions of each patch are determined first and transform is taken along these selected directions. Since directional variations change locally in images, transform directions also vary from patch to patch. That is transform directions get updated according to each patch.

4.1 Introduction

The previous chapter discussed three single image super-resolution methods - Freemann's learning based single image super-resolution method and two wavelet transform based approaches. The idea used in these methods is that a high resolution image can be decomposed into different frequency bands (using multi resolution methods like pyramid decomposition and wavelet transform). The low resolution image to be super-resolved contains only low and mid frequency bands. The missing high frequency information is learned from a training set which contains mid and high frequency bands of high resolution images. When compared to standard interpolation methods they give better results. But still there exists certain artifacts like ringing effect in edges, aliasing etc. So a new method is introduced here which uses skewed anisotropic wavelet transform called Directionlet transform.

4.2 Directionlet Transform:An Overview

It is already proved that the standard wavelet transform is an efficient tool for analyzing one dimensional signal. But two dimensional signals like images contain multiple direction oriented and elongated edges and wavelet transform fails to provide an efficient representation of such signals. This is due to the isotropic property of Wavelet transform. Wavelet transform is isotropic in the sense that equal number of filtering and sub-sampling operations are applied along both horizontal and vertical directions at each scale. Since directional features in synthetic and natural images involve more than two standard directions, multi-directionality and Directional Vanishing Moment (DVM) play an important role in pursuing sparse representations. Isotropic transform can not properly capture the anisotropic discontinuities present in the two dimensional signals like images. This is because the directions of the transforms and discontinuities in images are not matched and the transforms fail to provide a compact representation of two-dimensional signals.

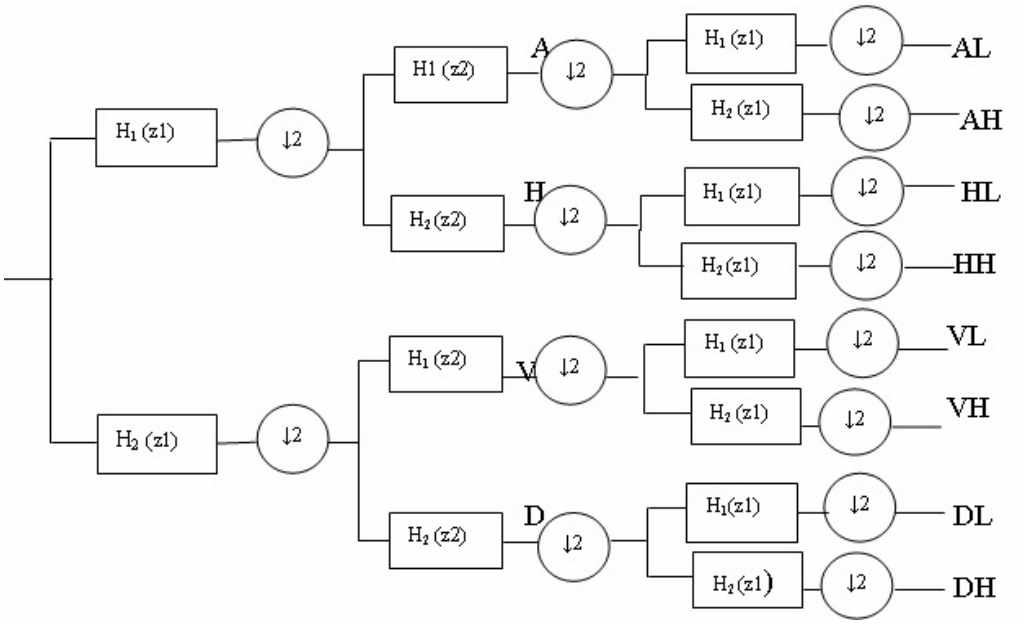


Figure 4.1: Filtering scheme for the AWT (2,1), where one step of iteration is shown

As different from wavelet transform, the directionlet transform is anisotropic. The difference between isotropic and anisotropic wavelet transform is that in the anisotropic wavelet transform, the number of transforms applied along the horizontal and vertical directions is unequal, that is, there are n_1 horizontal and n_2 vertical transforms at a scale, where n_1 is not necessarily equal to n_2. The iteration process is continued in the lower sub-band, as in the standard wavelet transform to obtain multi level transform. Anisotropic transform is represented as AWT(n_1,n_2). The DWT can be represented as a special form of the anisotropic wavelet transform, with $n_1 = n_2 = 1$. The standard DWT is simply given by AWT(1,1) and it decomposes a signal into four bands A, H, V and D. The ratio n = n_1 / n_2 is known as anisotropic ratio and it determines elongation of the basis functions of the AWT(n_1, n_2). When n_1 = 2, n_2 = 1, the AWT(2,1) produces eight bands AL, AH, HL, HH, VL, VH, DL and DH as in Figure 4.1. To obtain the

next level decomposition the process are repeated in the low sub band AL. Figure 4.2(a) and Figure 4.2(b) shows 3 level frequency decomposition of 2D DWT and AWT. This figure shows that a simple image with one discontinuity along a smooth curve is represented by the two types of basis functions: isotropic and anisotropic. The support of these basis functions is shown schematically as black rectangles, (a) isotropic basis functions generate a large number of significant coefficients around the discontinuity, (b)anisotropic basis functions trace the discontinuity line and produce just a few significant coefficients. The skewed AWT can follow discontinuity efficiently with fewer significant coefficients as compared to wavelet as shown in Figure 4.2(c) and Figure 4.2(d) [10].

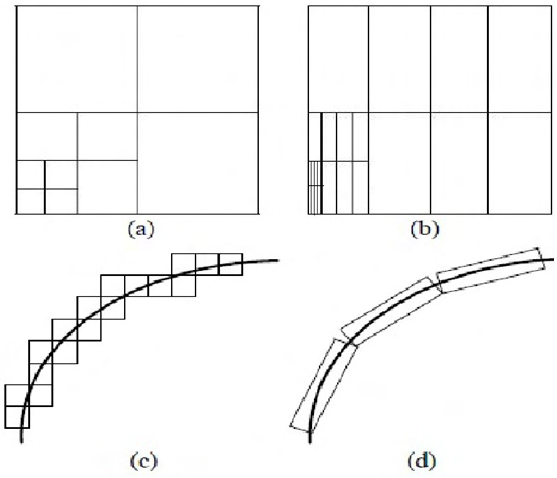


Figure 4.2: Frequency decomposition of (a)3-level 2-D DWT (b)3-level AWT (c)Isotropic basis function (d) Multi-directional and anisotropic basis function

The directionlets are Skewed Anisotropic Wavelet Transform (S-AWT). That means, scaling and filtering operations are along a selected pair of directions, not necessarily horizontal and vertical. The operations of standard wavelet transform like 1D filtering and sub-sampling are retained here and can provide anisotropic perfect reconstruction. This property makes the main difference between directionlets and some other directional filter banks, like Gabor filters. Other directional transforms (like curvelets, contourlets and bandlets) require over-sampling, and have higher complexity than the standard 2D-DWT. They need non-separable convolution and filter design which includes computationally more complex operations. The standard WT uses only horizontal and vertical directions and the high pass filters in this transform have vanishing moments only along these standard directions. The directionlet transform construction based on partitioning of the discrete space using integer lattices, where the 1-D filtering is performed along co-lines across the coset of lattice. The corresponding anisotropic basis functions are called directionlets. Lattice based approach forms a new transform construction method to include separable (1-D) filtering and sub sampling across multiple directions, not necessarily on horizontal and vertical.

4.2.1 Problem of directional interaction

As already stated directionlet transform or the skewed anisotropic wavelet transform is obtained by applying transform in two random directions, not necessarily along horizontal and vertical directions. For this, the transform can be taken on two random digital lines. Applying transform in two random directions causes a problem called directional interaction. A Digital line L(m, c) is obtained by discrete approximation of continuous line having slope m and intercept c and can be represented using equation 4.1[105]. Figure 4.3 represents two digital lines L(1/2,0), L(1/2,3) with slope 1/2 and different intercepts '0' and '3'.

$$y = mx + c \tag{4.1}$$

To apply a discrete transform in the discrete space in a certain direction, the pixels are to be defined in that chosen direction itself. The condition for critical

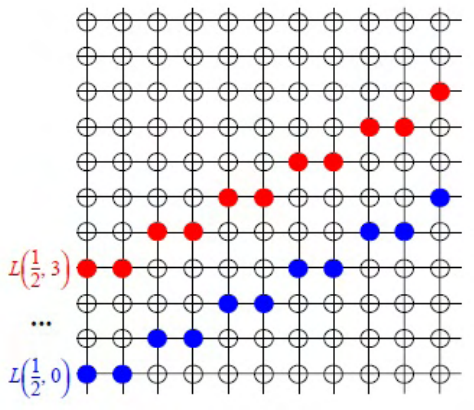


Figure 4.3: Two digital lines L(1/2,0),L(1/2,3)

sampling of the transform, is that every pixel belongs to one and only one digital line L(m,n). Another condition is that after taking 1-D transform along first direction, the remaining pixels must be aligned in second direction. A set of digital lines, with slope m, partitions the discrete space. The problem of directional interaction can be explained using an example.

In Figure 4.4, transform is applied in two random directions with slope -1/2 and 2/3. 1-D WT is applied along the digital lines L(-1/2, n). The result is that high pass filtering vanishes the digital line with the slope -1/2. The coefficients along this direction are annihilated in the high pass sub band. The coefficients along the second direction with the slope 2/3 are retained after filtering. However, after sub sampling, unlike in the case of the standard directions, the coefficients along the second direction are not aligned, that is, they cannot be clustered in the digital lines with the slope 2/3. This is against condition of critical sampling, that after sub sampling pixels should be aligned along the second transform direction. Therefore, the following 1-DWT applied along the digital lines with the slope 2/3 does not annihilate the coefficients along the second direction and, hence, it yields a non-sparse representation. This phenomenon is called directional interaction[102].

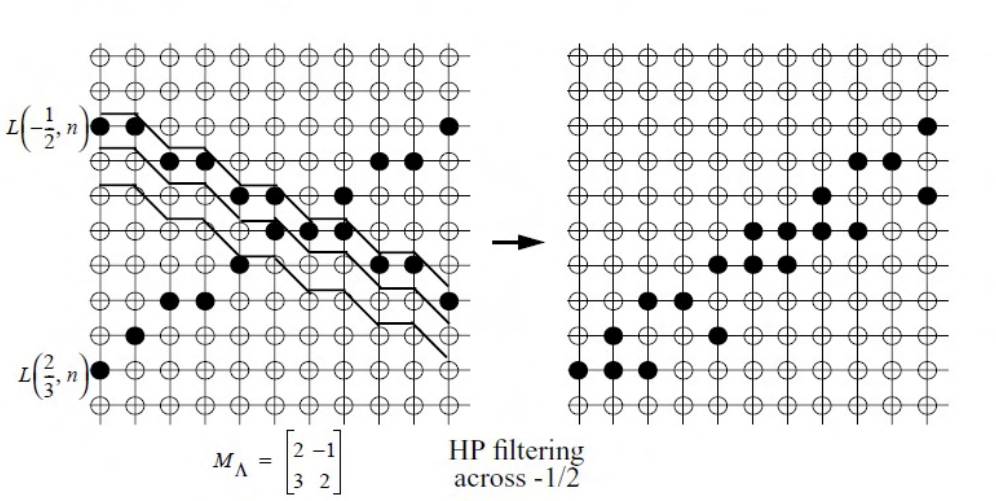


Figure 4.4: Directional interaction

It is clear that the transform along random digital lines is efficient when applied in over sampled schemes but it fails to provide a systematic sub sampling method when critical sampling is carried out. Hence the concept of digital lines is not sufficient to provide a systematic rule for sub sampling in the case of critical sampling. To overcome the problem of directional interaction and to propose an organized iterated subsampling method, the concept of integer lattices is proposed by Velisavljevic et al [102].

4.2.2 Lattice based Transform

Lattice

To avoid directional interaction, lattice based sub sampling method can be used. A full rank integer lattice A can be considered as a collection of points obtained by taking linear combinations of two linearly independent vectors or digital lines where both the components of the vectors and the coefficients are integers. A lattice can be considered as sublattice of the ordinary cubic integer lattice. The

lattice A can be represented by a generator matrix M_A given by;

$$M_A = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \tag{4.2}$$

where a_1, a_2, b_1 and b_2 are integers. The slopes of individual vectors (digital lines) which constitute the lattice are obtained from these integers. They are slope $m_1 = a_2/a_1$and slope $m_2 = b_2/b_1$. An example is shown in Figure 4.5. Its generator matrix M_1 is given by

$$M_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \tag{4.3}$$

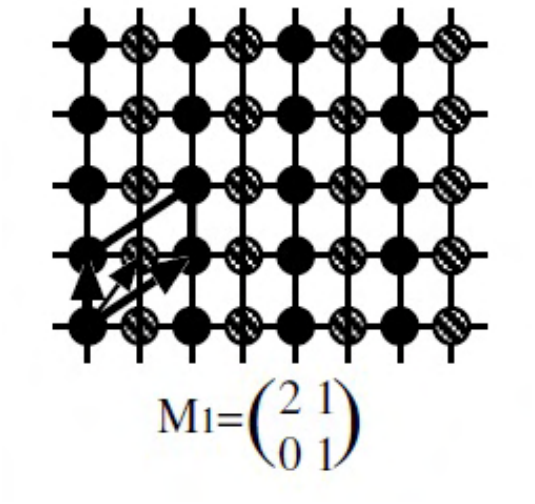


Figure 4.5: Lattice with generator matrix M_1

It is shown in lattice theory [35] that, given an integer lattice A with a generator matrix having determinant $\det(M_A)$, the lattice can be partitioned into $\det(M_A)$ cosets. Each coset is determined by shift vectors s_k, where $k = 0, 1, ..., \det M_A - 1$.

Shift vector determines how much the coset is shifted from origin. Each coset is a shifted version of lattice A. A co-line is defined as intersection of each coset and digital line. Therefore, the lattice A with the corresponding generator matrix M_A, partitions each digital line into co-lines.

For example, consider lattice A with generator matrix M_1 shown in Figure 4.5. The lattice A is divided into 2 cosets ($\det M_1 = 2$). The black dots form the first coset with shift vector $s_0 = (0,0)$, shaded dots form the second coset with shift vector $s_1 = (1,1)$.

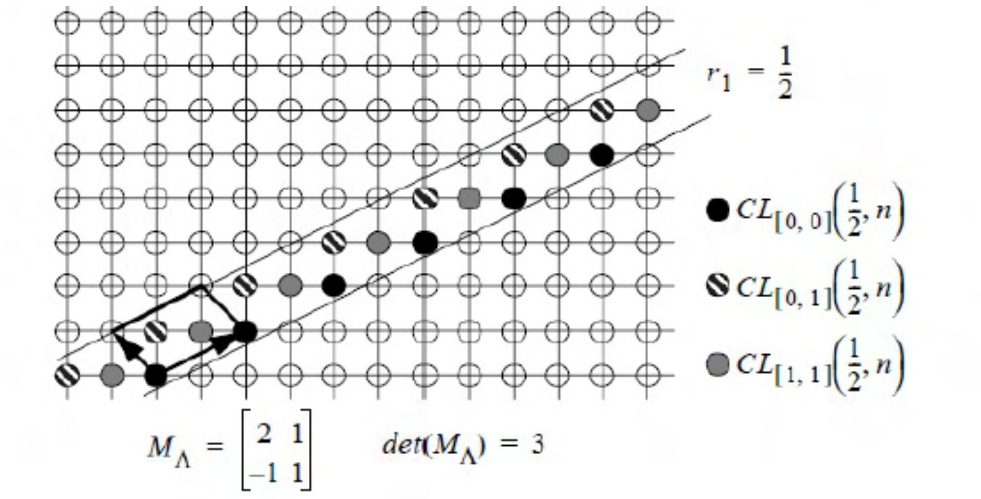


Figure 4.6: The intersections between the 3 cosets of the lattice A given by the generator matrix M_A.

Another example is shown in Figure4.6. Here generator matrix is given by the equation;

$$M_A = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \tag{4.4}$$

Since $\det(M_A) = 3$, there are 3 cosets. The 3 different shades of circles show 3 set of co-lines $CL_{[0,0]}(1/2,n)$, $CL_{[0,1]}(1/2,n)$, $CL_{[1,1]}(1/2,n)$ with digital line with slope=1/2.

Lattice based Transform

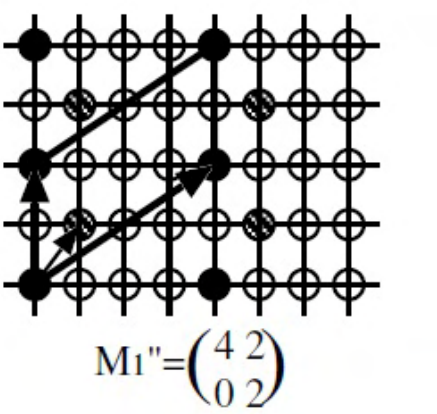


Figure 4.7: Subsampled version

To obtain lattice based transform, a 1-D wavelet transform is applied along co-lines with the first slope m_1 for all cosets of the lattice. Subsampling along that direction is done independently in each coset and the set of points obtained after the first subsampling consists of a lattice with generator matrix;

$$M_A^{\backslash} = \begin{bmatrix} 2v_1 \\ v_2 \end{bmatrix} \tag{4.5}$$

Discarding each second sample along each transform co-line secures a valid subsampling in the sense that perfect reconstruction condition is satisfied. The process is continued in a similar way along the second slope m_2 and the final

generator matrix is simply given by;

$$M_A^{\backslash\backslash} = \begin{bmatrix} 2v_1 \\ 2v_2 \end{bmatrix} \tag{4.6}$$

The direction along the first vector d_1 (with the slope $m_1 = a_2/a_1$) is called the transform direction. Similarly, the direction along the second vector d_2 (with slope m_2=b_2/b_1) is called the alignment direction. The sub sampled corresponding lattice is clearly a sublattice of the initial one containing a quarter of the samples[103]. Subsampled version of the lattice with generator matrix M_1 is given in Figure 4.7.

Another example of the lattice-based sub-sampling is shown in Figure4.8 for the direction 45^0 and -45^0. Here the lattice is divided into two cosets and transform is applied on each coset separately along 45^0.

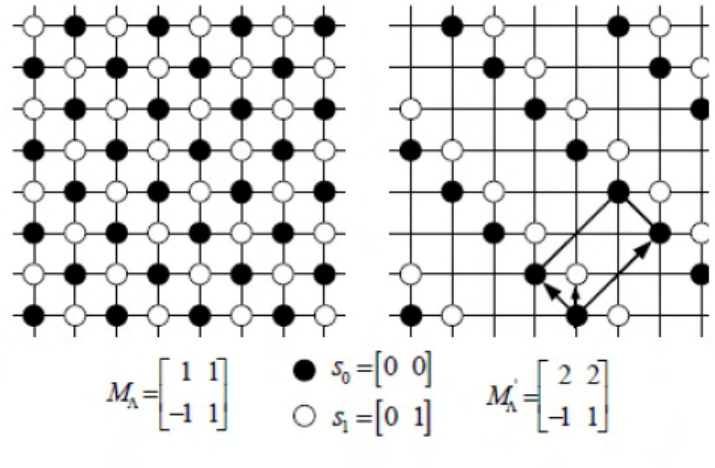


Figure 4.8: Lattice partitions the cubic lattice into cosets along 45° and –45°, then the sub-sampling are applied separately in two cosets

Figure4.9(b) shows the resultant lattice after applying lattice based filtering and subsampling along co-lines with slope $\frac{1}{2}$ in each coset separately. It is clear that

the lattice-based filtering and sub sampling does not create directional interaction. Since the filtering and sub sampling are applied in each coset separately, the pixels retained after the subsampling are clustered in co-lines along the alignment direction ($-\frac{2}{3}$). This property is crucial to avoid directional interaction.

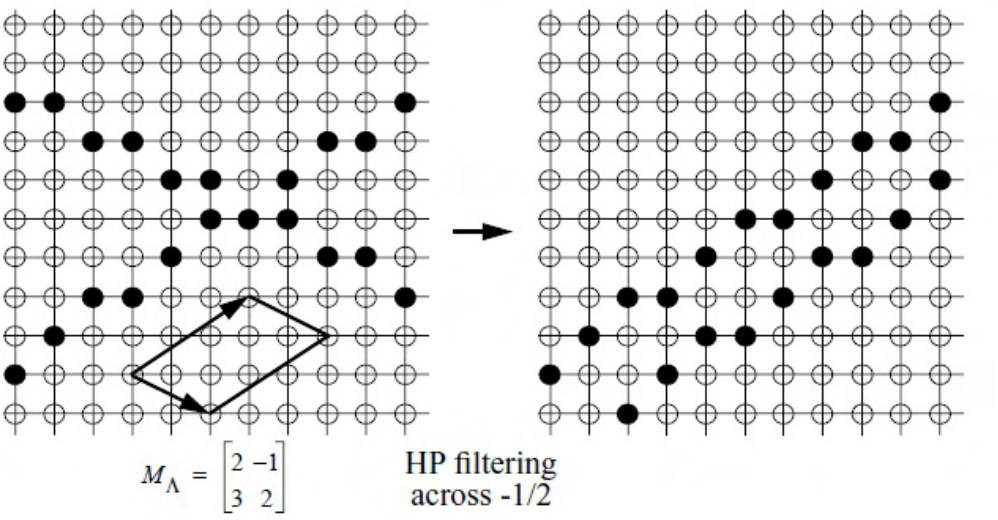


Figure 4.9: No directional interaction

For constructing the Directionlets, the discrete space containing the image is partitioned into integer lattices, where the 1D filtering is performed along co-lines across the lattice. For a lattice A, the skewed transforms are applied along co-lines in the transform and alignment directions of the lattice A. The basis functions of the S-AWT are called directionlets, since they are anisotropic and have a specific direction.

4.2.3 Directionlet transform along (0^0, 45^0)

Directionlet transform filtering and subsampling are done along different pair of directions. The pair of directions used in this work are (0^0, 45^0), (0^0, -45^0),

$(90^0, 45^0)$, $(90^0, -45^0)$, $(0^0, 90^0)$. The speciality of these pair of directions is that they produce only one coset. As the number of coset increases computational complexity also increases. Implementation of directionlet transform in $(0^0, 45^0)$ is explained below. First the transform is applied along 0^0 or horizontal direction and down sampling (removing column of pixels in 45^0). Then the co-lines along 45^0 directions are selected and transform is applied on these. It is clear that after the transform along 0^0 directions, the remaining pixels are aligned along 45^0. Transform is applied along 45^0 and then resulting pixels are aligned along 0^0.

Figure 4.10 shows a lattice formed with vectors in pair of directions $(0^0, 45^0)$. Figure4.11, 4.12, 4.13 show lattices after filtering and subsampling in 0^0, 45^0, 0^0 directions respectively. For example Figure4.11(a) shows the lattice after applying directional transform along 0^0 direction. Here the black dots are the ones which are to be removed in subsampling process. Figure4.11(b)shows the lattice after subsampling along 0^0 direction.

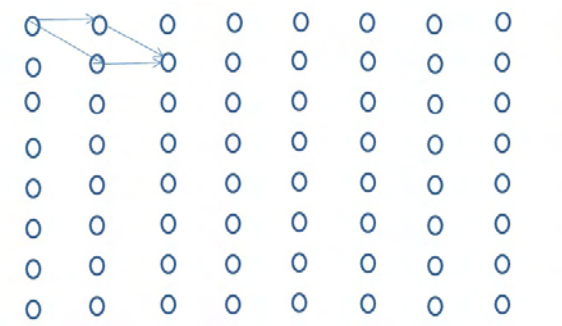
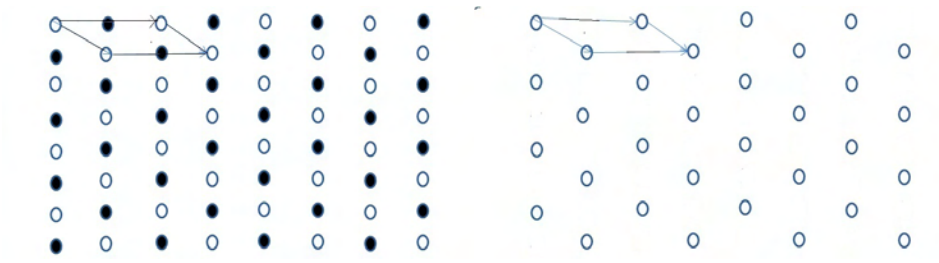
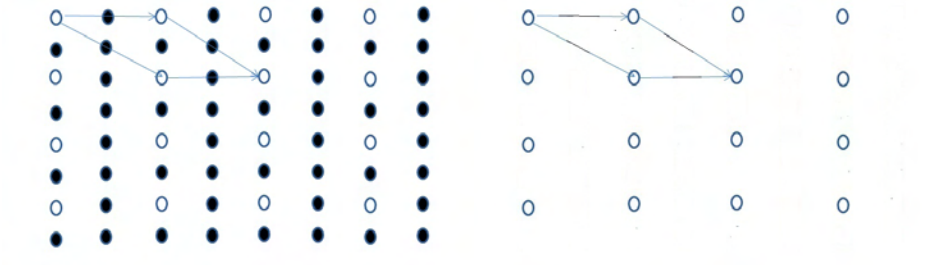


Figure 4.10: Lattice in (0,45) direction)



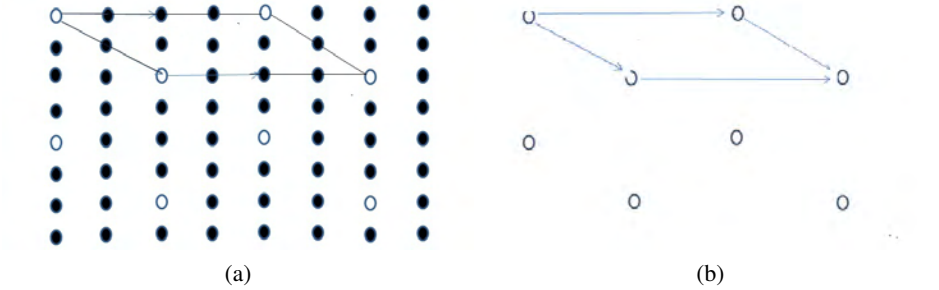
(a) (b)

Figure 4.11: (a) Lattice obtained after filtering and subsampling in 0 direction (b) Lattice obtained after removing subsampled pixels (black dots) in (a)



(a) (b)

Figure 4.12: (a) Lattice obtained after filtering and subsampling in (0,45) directions (b) Lattice obtained after removing subsampled pixels (black dots) in (a)



(a) (b)

Figure 4.13: (a) Lattice obtained after filtering and subsampling in (0,45,0) directions respectively (b) Lattice obtained after removing subsampled pixels (black dots) in (a)

4.3 Single image super resolution using Directionlet Transform

The new method presented in this chapter also uses a training set which contains different frequency features of available high resolution images. These features are used to obtain the high resolution version of an image captured using a low resolution camera. The idea used here is to learn the HR representation mapping of an LR edge from the training dataset during up sampling. Directionlet transform is used here to decompose images into different frequency bands. In the training set, information about patches or small blocks of high resolution images and their corresponding low resolution images are stored in terms of the coefficients of directionlet transform. Two methods are presented here. The first method uses critically sampled (subsampling followed by filtering) directionlets. The second method uses over sampled directionlets.

4.3.1 Single image super resolution method using critically sampled directionlets

Critically sampled directionlet transform involves filtering followed by sub sampling. Critically sampled directionlets are used to extract directional features present in high resolution and their corresponding low resolution images. This information is stored in the training set. Using this training set low resolution images is super resolved to obtain their high resolution images.

Training set generation

To generate a training set, a collection of high resolution images and their low resolution (LR) images are used. LR images are formed by averaging the intensities of non overlapping block of size 2×2 pixels from HR image, (where 2 is the decimation factor) using the equation 3.1. Directional information other than the standard (horizontal and vertical) directions can be extracted using directionlet transform. In the case of images, the directional information varies over space. Thus, directionality can be considered as a local feature, defined in a small neighborhood. Therefore, to extract directional variations of an image it has

to be analysed locally. For this, the HR and LR images are subdivided into patches of size 8×8 and 4×4, respectively in raster scan order. To avoid a blocking effect in the transform caused by the small patches, overlapping blocks are used. ie; extra pixels from neighboring patches are added on four sides to avoid errors in borders.

The super resolution algorithm also operates under the assumption that the predictive relationship between low and high resolution images is independent of local image contrast. Because of this, patch pairs of high and low resolution images are contrast normalized by the energy of the low-resolution patch. This energy is the average absolute value of the LR patch given by:

$$energy = 0.01 + \Sigma\sqrt{y_k{}^2} \tag{4.7}$$

where k is the value of the pixel number in the LR patch. The constant 0.01 is added to prevent division by zero.

Determination of best pair of direction for a patch

To apply the directionlet transform, optimum pair of direction must be selected. The best pair of directions for each patch is to be chosen from five sets of directions D = (0^0,45^0), (0^0,-45^0), (90^0,45^0), (90^0,-45^0), (0^0,90^0). The assigned best pair of transform directions of each patch form a directional map of that LR image and its corresponding HR image. Figures 4.14 (b) and (d) show directional map for the Figures 4.14 (a) and (c). The directional map gives best suitable pair of directions for each patch in the low resolution image.

To find the best pair of directions, the directionlet transform is applied in each patch along these five set of directions. The best pair of directions d_n is chosen for each patch indexed by n using equation 4.8

$$d_n = \arg\min \sum_{n,j} |W_{n,i}|^2 \tag{4.8}$$

where the directionlet coefficients $W_{n,i}$ are produced by applying directionlets to the n^{th} patch along the pair of directions d_n. Using directional map determined

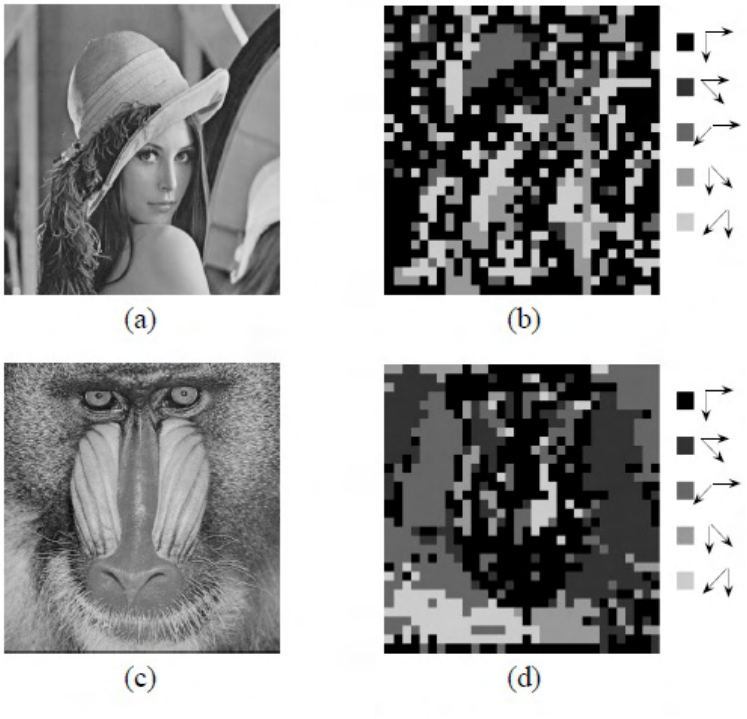


Figure 4.14: (a) The original image Lena.(b) The corresponding directional map.(c) The original image Baboon.(d)The corresponding directional map.

by the set D minimizes the energy in the high-pass sub-bands and provides the best matching between transform and locally dominant directions across the image patches.

Directionlet transform is then applied in each patch of the LR and HR image along the selected best pair of directions d_n using the db4 basis functions. In the training set, the coefficients of six sub-bands HL, HH, VL, VH, DL, and DH of LR and HR image patches are stored in lexicographical order. The patch coefficients are grouped and stored according to the patch direction, so that the searching time can be considerably reduced. There are five groups of coefficients corresponding

to the 5 pair of directions.

Prediction (learning) of HR image for a LR image

The given LR image is divided into overlapping patches of size 4×4 best pair of direction is selected as in the case of training set images. Along these directions Directionlet transform coefficients are found out. Each patch is energy normalized using the equation 4.7. To predict the high resolution image, the high resolution patches corresponding to each input low resolution patch must be found out. For this eight frequency bands – AL, AH, HL, HH, VL, VH, DL and DH corresponding to each high resolution patch must be known. The LR input image itself is used to obtain the low frequency band (AL and AH). For this the low resolution patch is upsampled by padding zeros in between the pixels to obtain the size $2M \times 2M$, where $M \times M$ is the size of low resolution patch. Directionlet transform is applied to it and the resulting AL and AH are used as AL and AH of unknown high resolution patch. Remaining frequency bands HL, HH, VL, VH, DL and DH are to be learned from the training set. The inverse directionlet transform of the learned six bands and AL and AH together gives the high resolution patch. Finally the predicted high resolution patch is denormalised using the energy of the LR input patch. The high resolution image is obtained by repeating these processes in the remaining low resolution patches.

It may be noted that like the wavelet method 2 image of any size can $M \times N$ can be super resolved to $2M \times 2N$, using the LR and HR patches coefficients in the training set.

Learning directionlet coefficients

As explained above, the given contrast normalised LR image to be super resolved is subdivided into patches of size 4×4 with suitable overlap of all sides. For db4 wavelet the overlap is three, 10×10 overlapping blocks are used. Each patch is divided into eight directional subbands using directionlet transform. Here also, the best direction pair for each patch is found out as in the case of training set images. The directionlet coefficients of the six bands- HL, HH, VL, VH, DL,

and DH corresponding to the finest level are learned from the training set. Here, the minimum absolute difference (MAD) criterion is used to select the matching directionlet coefficients. The absolute difference between directionlet coefficients of the 6 bands HL, HH, VL, VH, DL, DH of the input low resolution patch and those corresponding coefficients of the several low resolution patches in the training set is taken. The LR patch in the training set with minimum absolute difference is selected. The high frequency bands – HL, HH, VL, VH, DL, DH of its corresponding HR patch is taken as the missing high frequency bands of input patch. These learned high frequency bands and the LR input patch as approximation (AL and AH) are used to reconstruct the HR equivalent of the LR input patch. By taking the inverse directionlet transform, the high resolution patch is obtained. In effect, best matching 8×8 patch is obtained with the training data for a given 4×4 patch in the LR input image. Next the contrast normalisation of the patch is undone by multiplying the patch with energy of the LR patch. This is repeated for the entire input patches to obtain full HR image. LR image of any size can be super resolved with this training set, with a magnification factor of 2. The proposed method is effective in the case of aliased images also.

The new directionlet method can be summarized as follows

1. The LR image of HR image is obtained by averaging every 2×2 block pixels in HR image.

2. The low and high resolution images are partitioned into small blocks of size 4×4 and 8×8 (with suitable overlapping with four neighboring patches) respectively and patches are contrast normalised by energy of LR patch.

3. Suitable pair of directions for each LR patch is found out and Directionlet transform coefficients of that patch and its HR patch, along the selected pair of directions, are saved as the training set in five groups according to five pairs of directions.

4. The given LR input image is also partitioned into small blocks and

directionlet transform is applied to it, along its selected pair of directions.

5. Directionlet coefficients of this input patch are compared with those of the LR patches in the training set, using minimum absolute difference criterion.

6. LR patch with minimum absolute difference is selected and coefficients of the corresponding high frequency bands are used as the missing bands for the unknown HR patch.

7. Take the inverse directionlet transform to obtain the HR patch and undo the contrast normalisation.

8. Repeat the above process for the remaining input patches.

.

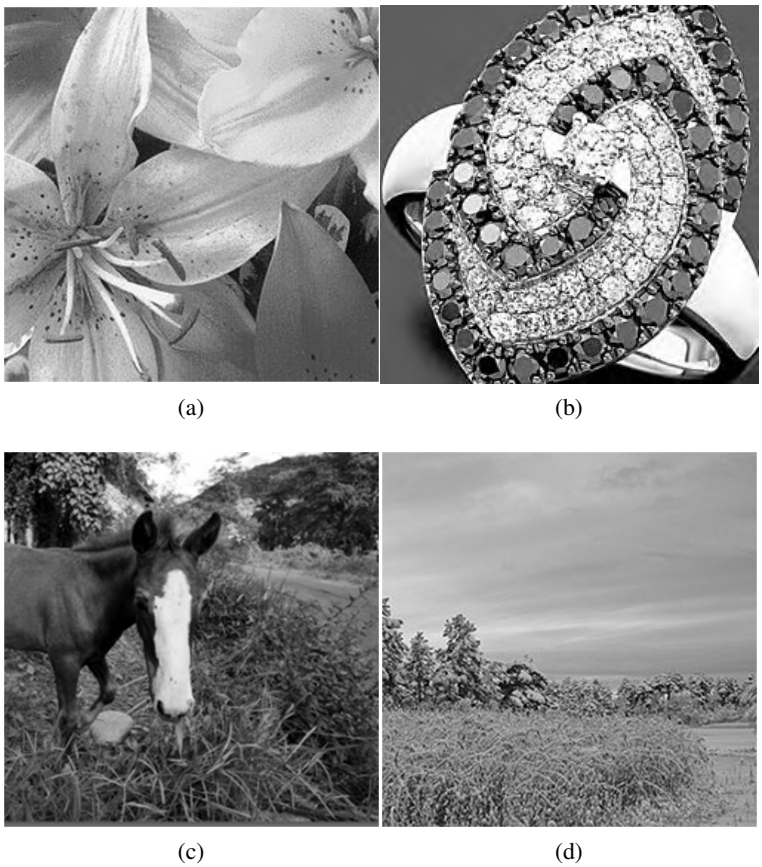
Implementation and results

Experiments are performed for various types of grey images, using a training set which contained good quality images having various levels of information content. The important thing is that training set is not specific to the class of objects to be super resolved. Some of the training set images are shown in Figure 4.15 and they are of size 256×256. The LR image to be super resolved was obtained by downsampling an available HR image using the averaging using equation 3.1. For finding the directionlet transform db4 wavelet basis is used. The low resolution images are super resolved to double size. That is the magnification factor is 2.

Implementation

Using only the same (its own original) HR image in the training set

To study the comparative performance of the various super resolution methods, first the same original HR image of the input LR image alone is used in the training set. The low and high resolution images are divided into overlapping patches of size 8×8 and 4×4. The number of overlapping pixels is determined by the type of wavelets used. From experiments it is observed that the optimum size for high resolution patch is 20×20, when db4 is used. It is true for all wavelets with 8



(a) (b)

(c) (d)

Figure 4.15: Training set images

filter coefficients for example 'bior 3.3'. For LR image, patch size is 10×10. The extra number of pixels, 'P' to be added on each side of the 8×8 patch depends on wavelet type and is determined by the equation4.9.

$$P = L - 2 \tag{4.9}$$

where L is length of wavelet coefficients. For db4 'P' is six, since number of wavelet coefficients are eight and patch size is 20×20. For db5 which has 10 wavelet coefficients, P is 8 and patch size is 24×24. The directionlet transform of

the contrast normalised patches are taken and are stored in the training set.

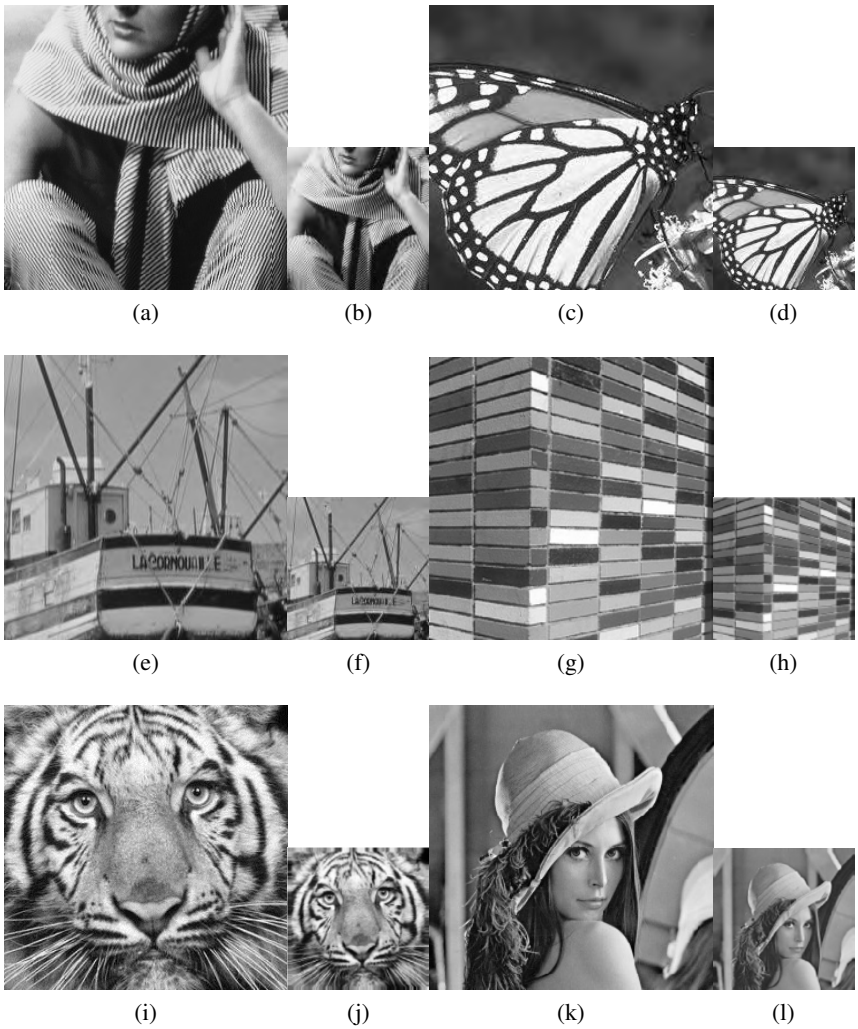
To super resolve a low resolution image, it is divided into over lapping patches of size 4×4. The number of overlapping pixels is '3' for db4 wavelet basis. Directionlet transform of each patch is taken and is compared with corresponding low resolution patches in the training set. The patch with minimum absolute difference is selected and high frequency bands of it's high resolution patch is used as the missing high frequency bands of unknown high resolution patch. The approximation (AL and AH) are obtained from the LR image patch itself.

The directionlet method has been compared with results obtained with the cubic spline interpolation method, the two wavelet transform-based methods mentioned in previous chapter. In wavelet method1, training set is formed with high resolution image of size 256×256. Using this training set, only low resolution images of size 128×128 can be super resolved. Hence, for comparison, in the new directionlet method also, low resolution images of size 128×128 are considered initially. They are super resolved to the size of 256x256. Figures 4.16 show a few of the high resolution images and their corresponding low resolution images used for super resolution experiments.

Signal to Noise Ratios (SNR) are calculated using equation4.10.

$$SNR = \sum_{i,j} \frac{{Z_{i,j}}^2}{\sum_{i,j} [Z(i,j) - Z'(i,j)]^2} \tag{4.10}$$

where $z(i,j)$ and $z'(i,j)$ is the (i,j) the pixel intensity of original image and reconstructed image respectively.



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

(e) (f) (g) (h)

(i) (j) (k) (l)

Figure 4.16: (a), (c), (e),(g),(i),(k) Original images (256x256)(b), (d), (f),(h),(j),(l)Low resolution images 128x128

SNR obtained for the different methods are shown in Table4.1. It may be noted that SNR for directionlet method is more than the other methods by 10 to 15 dB. The SNR values of cubic spline interpolated image and block based

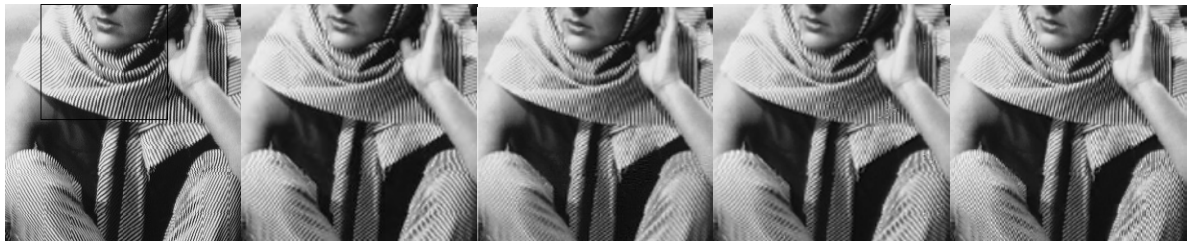
Table 4.1: SNR values with the same original image alone in the training set

Method	SNR in dB			
	Barbara	Butterfly	Tiger	Boat
Cubic spline	15.5907	21.1581	20.6350	23.2912
Wavelet method 1	14.7646	10.2796	10.8505	11.9395
block Wavelet method	29.5342	34.0212	33.6059	36.7257
Directionlet method	33.1621	37.3297	34.7294	40.1452

wavelet method are 15.5907dB and 29.5342dB while it is 33.1621dB when using directionlet transform for the image Barbara. Thus from the SNR values it is clear that directionlet method outperforms wavelet methods and cubic spline method.

Figure 4.17 shows the results obtained using the low resolution image in Figure 4.16(b) of the image Barbara. Figures 4.17(b), (c), (d) and (e) are the images super resolved using cubic spline method, wavelet method 1, block wavelet method and the new directionlet method respectively. The training set contains only its own original HR image as shown in Figure 4.16(a). Figure 4.17 (f), (g), (h), (i), (j) show the zoomed marked area in Figure 4.17(a) with respect to Figure 4.17(b), (c), (d), (e). It can be noted that cross lines have appeared on the scarf as well as the region below the knee in images (b), (c), and (d). These cross lines are completely eliminated in the Figure 4.17(e) in the scarf region which is obtained by new directionlet method. This can be visualised clearly in the corresponding zoomed figure 4.17(j).

The results of the experiments done on another image are shown in Figure 4.18. The low resolution image is that of the butterfly shown in Figure 4.16(d). Figures 4.18(b), (c), (d), (e) show the results obtained using standard cubic spline interpolation, wavelet methods and newly developed directionlet method respectively. Figures 4.18(f), (g), (h), (i), (j) are the zoomed portions of the marked portion 1 in Figure4.18 (a) with respect to Figures4.18(a), (b), (c), (d), (e). The ringing effects on the sharp dark edges are visible in Figures4.18(g), (h), (i) whereas, it is eliminated in Figures4.18(j). Figures4.18(k),(l) , (m), (n), (o) are the



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

(f) (g) (h) (i) (j)

Figure 4.17: (a)original image (b)cubic spline interpolated image (c)super resolved using wavelet method 1 (d)super resolved using block wavelet method(e)super resolved using directionlets (f), (g), (h), (i), (j)zoomed portion of the face (a),(b),(c),(d),(e)respectively.

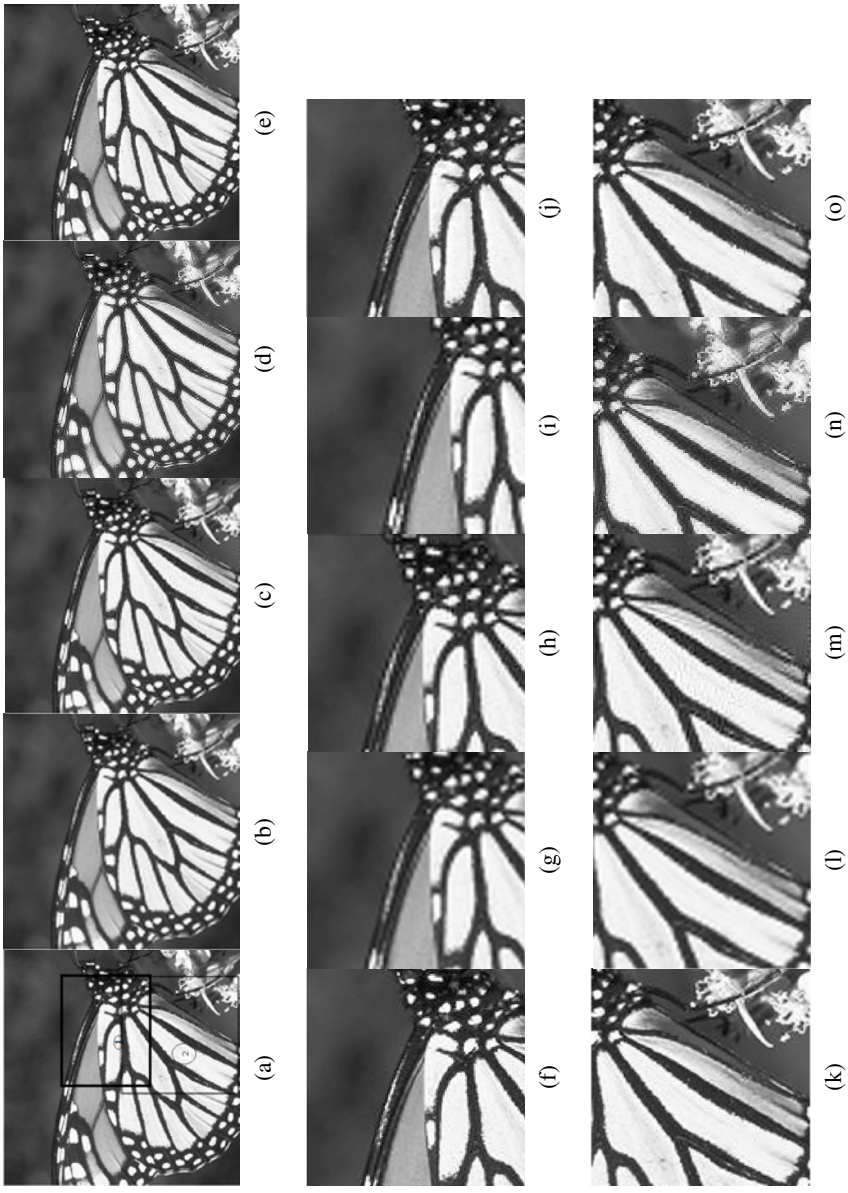


Figure 4.18: (a)original image(b)cubic spline interpolated image(c)(d)super resolved using wavelet methods(e)super resolved using directionlets(f),(g),(h),(i),(j)zoomed portion of the area marked 1 (k),(l),(m),(n),(o) are zoomed portion of area marked 2.

zoomed portions corresponds to the lower right portion of the wings of the butterfly (marked 2) in Figures4.18(a). The artifacts in the region in figure4.18(o) using new directionlet method are less.

Thus from the quantitative measurement using SNR values and visual subjective quality of the resultant images, it is clear that directionlets are far better than the wavelets and standard super resolution methods.

Implementation using entirely different images in the training set

To validate the performance of the directionlet method, the super resolution experiments are done using entirely different HR training set images. That is, the training set does not contain the HR image of the input LR image as in the above case.

To compare with existing super resolution methods, Yang et al sparse super resolution method [36] is implemented . In this method low-resolution image is viewed as down sampled version of a high-resolution image, whose patches are assumed to have a sparse representation. The principle of compressed sensing ensures that under mild conditions, the sparse representation can be correctly recovered from the downsampled signal.

Here the new block wavelet method and the directionlet method are compared with wavelet method 1, wavelet method proposed by Sapan et al [86], patch based methods of Freeman [108] and sparse method by Yang et al[36].

The SNR values are calculated for different methods and are shown in table 4.2. For the image Barbara, the SNR value of directionlet method is 20.3447dB which is much higher compared to other methods. SNR values of wavelet method 1 are low compared with other methods. For butterfly image, sparse method by Yang gives SNR value of 30.4005dB while with directionlet method SNR is 26.7953dB. But from the corresponding images in figure 4.20, it can be seen that there are artifacts present on the edges of the wings in the sparse method, while they are

almost eliminated in the new directionlet method. Since images with less artifacts are preferable in many applications, the new method is better than sparse method. One main observation from the results obtained from the new method is that this method is more effective for images with more edges.

Table 4.2: SNR values with entirely different HR images in the training set

Method	SNR in DB			
	Barbara	Butterfly	Tiger	bricks
Wavelet method 1	8.7688	3.1185	3.6367	3.2598
Sapan et al's wavelet based method	16.1121	20.9643	19.1132	14.3642
Yang et al's sparse method	16.9676	30.4005	26.5871	19.2689
Freeman method	13.3849	22.3943	19.9744	15.4277
Block wavelet method	17.4133	24.8016	24.1414	18.3143
Directionlet method	20.3447	26.7953	27.1543	19.5354

Figure 4.19(b), (c), (d) are HR images of low resolution image in Figure 4.16(b) obtained by Sapan et al wavelet based super resolution method , Yang et al sparse super resolution method and block wavelet method. In these images, there are cross lines in the stripes of the scarf which makes them aliased. Figure 4.19(e) shows the super resolved image obtained using the new directionlet method. It is seen that aliasing is reduced considerably here since there are less cross lines and directions of stripes are almost same as original image.

Figure 4.20 (b), (c), (d) are HR images obtained by Sapan et al wavelet method, sparse method and new directionlet methods. The zoomed portions of upper edge of wings shown in Figures 4.20 (f), (g), (h) show that the ringing effect is reduced in the new method.

Figure 4.19: (a)Original image(b)Super resolved image using Sapan et al's wavelet method (c)Super resolved image using Yang et al's sparse method(d)block wavelet method(e)Super resolved image using directionlets (f), (g), (h), (i) zoomed portion of the face of (a), (b), (c), (d),(e)respectively.

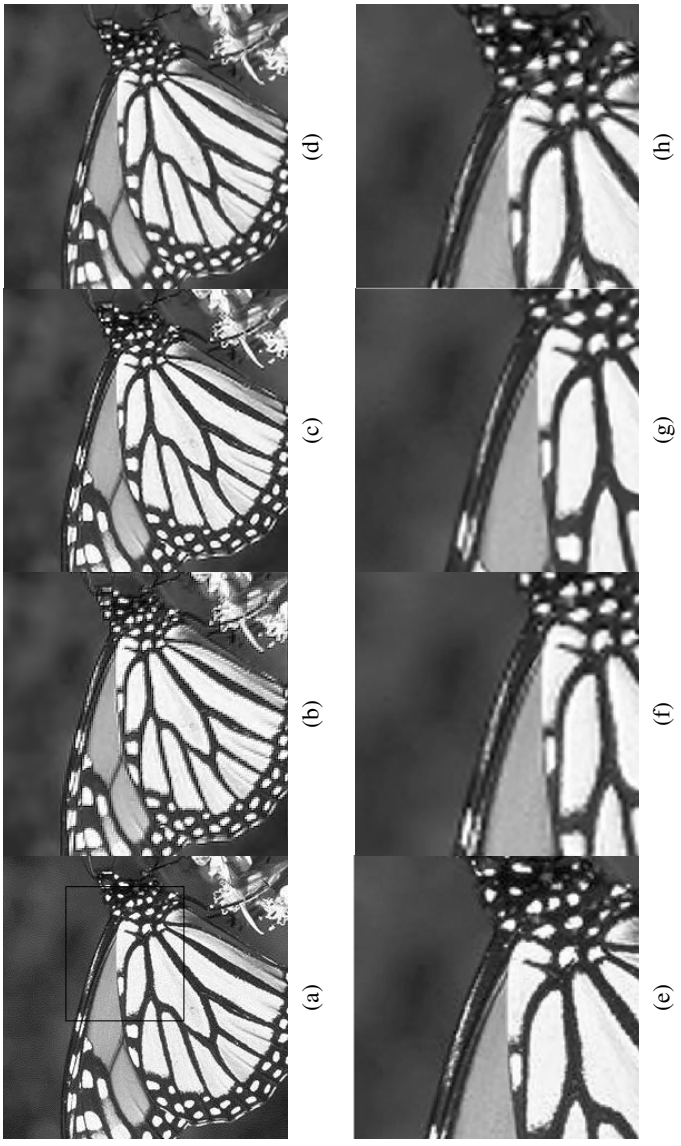


Figure 4.20: (a)original image (b)super resolved image using Sapan et al's wavelet method (c)super resolved image using Yang et al 's sparse method(d)super resolved image using directionlets (e), (f), (g), (h) zoomed portion of the face of (a), (b), (c),(d)respectively.

More results are shown below. The Figures 4.21(a), (b), (c), (d), (e), (f) are original image, super resolved images using wavelet method 1, wavelet based Sapan et al super resolution method , Yang et al sparse method , block wavelet method and directionlet method. The low resolution image is shown in Figure 4.16(f). The Figure 4.21(g), (h), (i), (j),(k),(l) are the zoomed portions of marked area in (a), (b), (c), (d), (e) and (f) respectively. There exist ringing effects on the black poles of boat in Figure 4.21(g), (h), (i), (j), (k) but these ringing effects are not present in Figure 4.21(l).

The Figures 4.22(a), (b), (c), (d), (e),(f) are original image, wavelet method 1, wavelet based Sapan et al method, super resolved image using yang et al method, block wavelet method, new directionlet method respectively. The low resolution image is shown in Figure 4.16(j). The Figures 4.22(g), (h), (i), (j),(k),(l) are the zoomed portions of (a), (b), (c), (d), (e) and (f) respectively. The ringing effects on the mustache of tiger in Figures 4.22(g), (h), (i),(j),(k) can be clearly seen but these ringing effects are not present in Figure4.22(l). It is more sharper than those with other methods.

The Figures 4.23(a), (b), (c), (d), (e), (f) are original image, wavelet method1, super resolved using wavelet based Sapan et al method, super resolved image using Yang et al method, block wavelet method and new directionlet method respectively. The low resolution image is shown in Figure 4.16(h). The Figure4.23 (g), (h), (i), (j), (k), (l) are the zoomed portions of (a), (b), (c), (d),(e) and (f) respectively. There exist artifacts on the edges of each block in Figures 4.23(g), (h), (i), (j), (k) but these artifacts are very much reduced in Figure4.23(l).

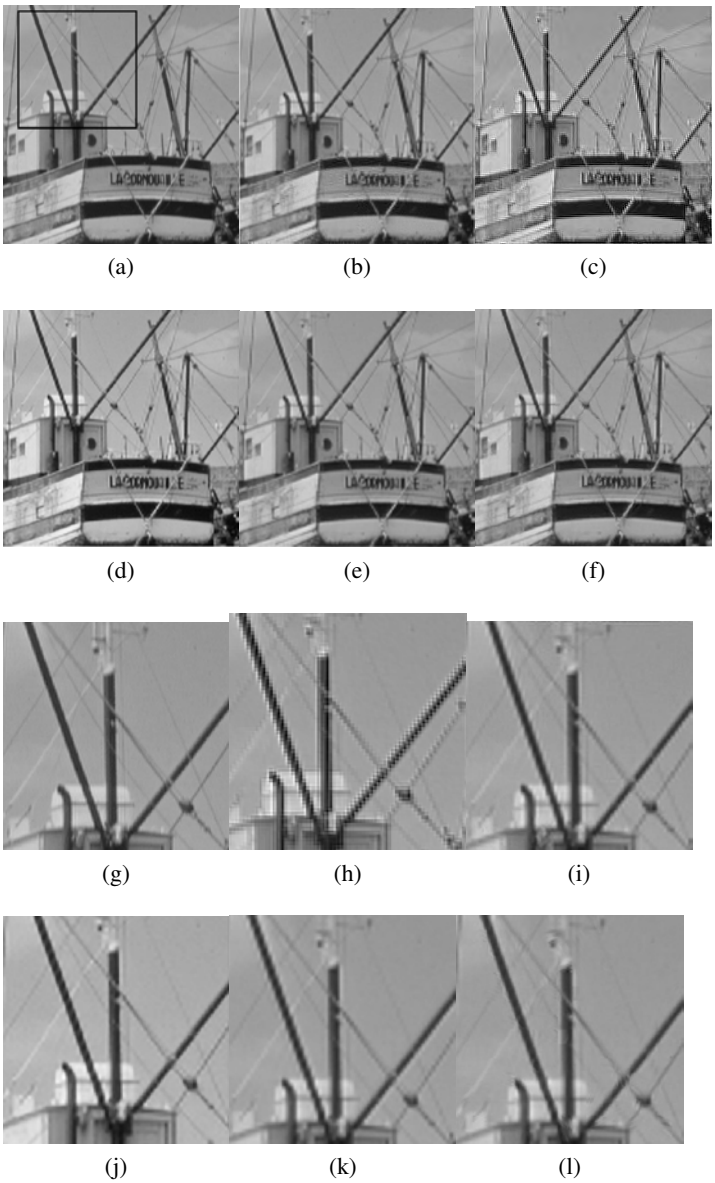


Figure 4.21: (a)original image(b) super resolved image using wavelet method1(c)super resolved sapan et al's wavelet method(d)super resolved using yang et al's sparse method (e)block wavelet method (f) super resolved using directionlets (g), (h), (i), (j),(k),(l) zoomed portion of the marked area of (a), (b), (c),(d), (e), (f)respectively.

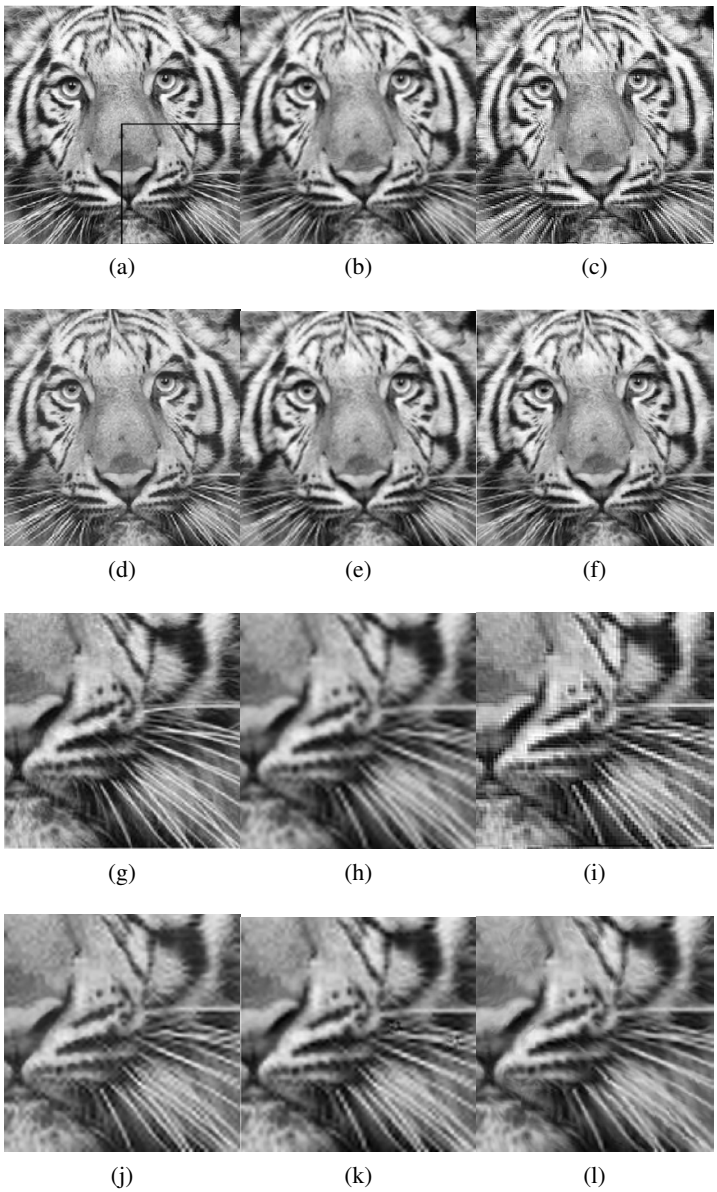
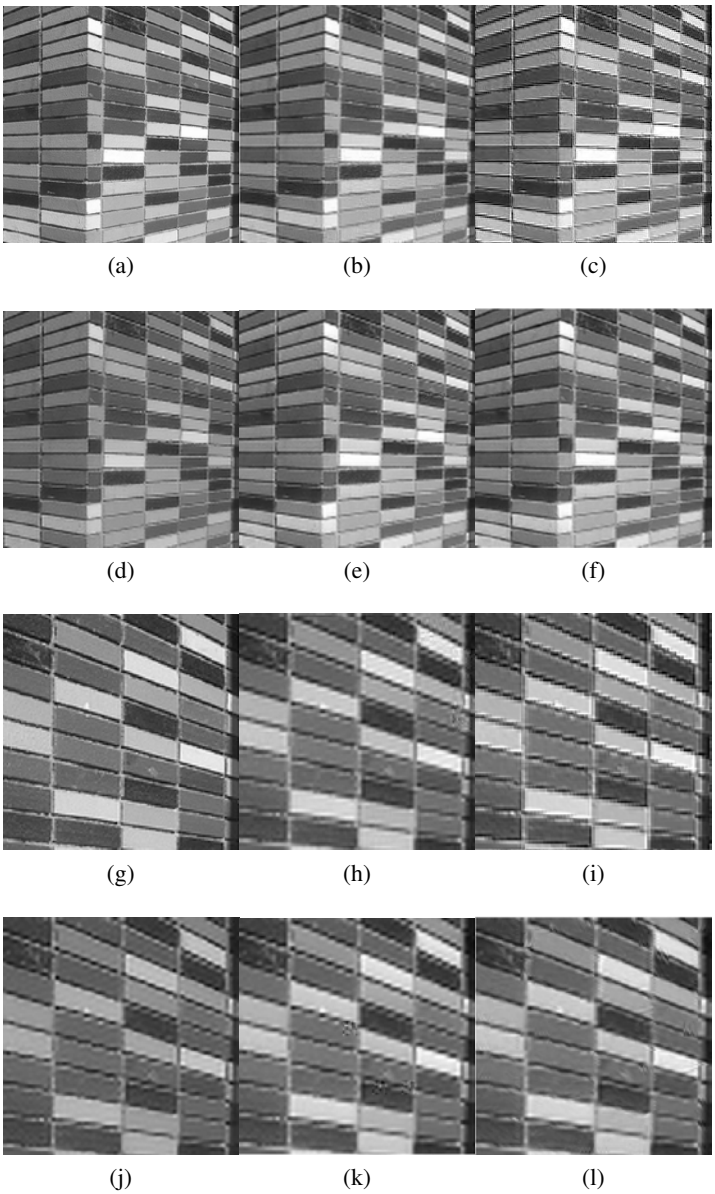


Figure 4.22: a)original image(b) super resolved image using wavelet method 1(c)super resolved sapan et al's wavelet method(d)super resolved using yang et al's sparse method (e)block wavelet method (f) super resolved using directionlets (g), (h), (i), (j),(k),(l) zoomed portion of the marked area of (a), (b), (c),(d), (e), (f)respectively.



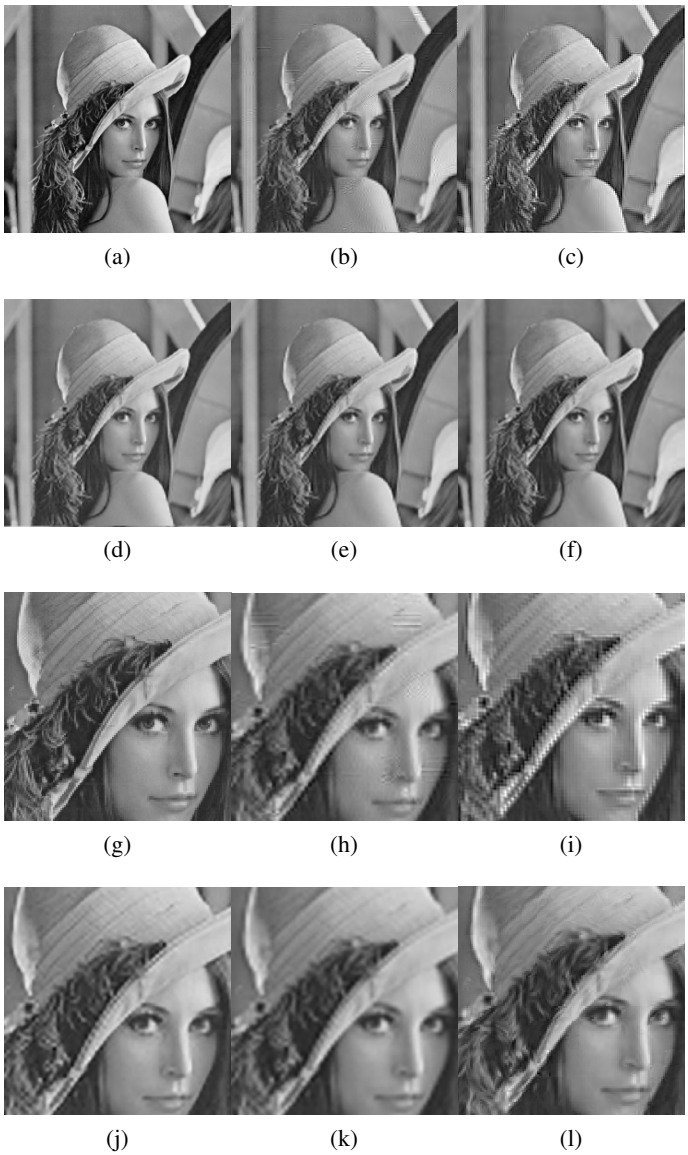
(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Figure 4.23: a)original image(b) super resolved image using wavelet method 1(c)super resolved sapan et al's wavelet method(d)super resolved using yang et al's sparse method (e)block wavelet method (f) super resolved using directionlets (g), (h), (i), (j),(k),(l) zoomed portion of the marked area of (a), (b), (c),(d), (e), (f)respectively



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Figure 4.24: (a)original image(b) super resolved image using wavelet method 1(c)super resolved sapan et al's wavelet method(d)super resolved using yang et al's sparse method (e)block wavelet method (f) super resolved using directionlets (g), (h), (i), (j),(k),(l) zoomed portion of the marked area of (a), (b), (c),(d), (e), (f)respectively

The Figures 4.24 (a), (b), (c), (d), (e), (f) are original image, wavelet method1, Sapan et al wavelet based method, using Yang et al method, block wavelet method and new directionlet method respectively. The low resolution image is shown in Figure 4.16(h). The Figure4.24 (g),(h), (i), (j), (k) and (l) are the zoomed portions of (a), (b), (c), (d), (e) and (f) respectively. There exist artifacts on the edge of cap in Figures 4.24(b),(c), (d), (e) but these artifacts are very much reduced in Figure4.24(f).

The visual quality of the super resolved images can be improved by adding more suitable images in the training set. It may be noted that the maximum quality obtained for a low resolution image, with the new directionlet based method, is with its own high resolution image in the training set. The more the images used to build training set is similar to the original image, the more will be the quality of the super resolved image. From SNR values and visual quality, it is clear that the new directionally adaptive algorithm using directionlets are more efficient than other traditional methods.

4.3.2 Single image super resolution method using over sampled directionlets

Over sampled directionlet method

The previous section described single image super resolution using critically sampled directionlets. In this section the above mentioned method is implemented with over sampled directionlets. Over sampled directionlets means there is no subsampling after filtering process. AL and AH are obtained from the low resolution image as in the way as the case of above Directionlet method.

Over sampled plus cubic spline interpolation method

The above mentioned method is modified using cubic spline interpolation. Here the cubic spline interpolated low resolution image is used to obtain low frequency bands AL and AH.

Implementation of both methods

Training set is formed with oversampled directionlets.

SNR values with critically sampled directionlet method, over sampled directionlet method, over sampled directionlet and cubic spline interpolation method are given in Table 4.3. SNR values show that critically sampled directionlet method and over sampled directionlet method give almost same result but the over sampled method needs more memory and computations. So critically sampled directionlet method is superior compared with other two methods.

Table 4.3: Comparison of SNR values with different methods

method	SNR in dB		
	Barbara	Butterfly	Tiger
Critically sampled Directionlet method	22.9467	26.8020	25.1543
Oversampled Directionlet method	22.5626	26.9598	25.253
Modified oversampled directionlet method with cubic spline interpolation	15.824	21.082	20.6603

Figures 4.25 (b), (c), (d), (e) show original image, super resolved image using critically sampled directionlet method, super resolved image using over sampled Directionlet method and super resloved image using modified over sampled directionlet method with cubic spline interpolation respectively. It is seen that there is not much difference between images obtained using critically sampled directionlet method and over sampled directionlet method. It is due to the fact that AL and AH are obtained in the same way as in the case of directionlet method. The image with third directionlet method is not much sharper than other two methods and its SNR value is low compared with other methods.

Figures 4.26 (a), (b), (c), (d), (e) show low resolution image, original image, super resolved image using Directionlet method , super resolved image using over sampled directionlet method, super resolved image using modified oversampled Directionlet method with cubic spline interpolation respectively.

From SNR values and visual quality, it is clear that critically sampled directionlet method is better than other over sampled methods. Since the oversampled methods

(a) (b)

(c) (d) (e)

Figure 4.25: (a)low resolution image (b)original image (c) super resolved using critically sampled directionlet method(d) super resolved using directionlet method with over sampled directionlet (e)modified over sampled directionlet method with cubic spline interpolation

need more computations and time compared to the critically sampled method, they are not considered for comparison in the coming chapters.

(a) (b)

(c) (d) (e)

Figure 4.26: (a)low resolution image (b) original image (c)super resolved using critically sampled directionlet method(d) super resolved using over sampled directionlet method(e)super resolved using modified oversampled directionlet method with cubic spline interpolation

4.3.3 Super resolution of LR images to 4 times its orginal size (magnification factor 4)

The LR image of size MxM is super resolved to an image of size 4Mx4M using the new directionlet method and block wavelet method.

The zooming factor 4 is obtained by iterating the critically sampled directionlet method two times, with output image of first iteration as the input low resolution image of the second iteration. The low resolution input image to the first stage is obtained using equation 3.1 as mentioned in section 3.2. SNR values obtained for the different methods for zooming factor 2 and 4 are shown in table 4.4. From section 4.3 it is seen that Yang et al's method gives the best result among the other existing SR methods. So the directionlet method is compared with Yang's method and block wavelet method only.

Table 4.4: SNR values for grey images with zooming factor 2 and 4

Method	SNR in DB			
	for zooming factor 2		for zooming factor 4	
	Barbara	Bush	Barbara	bush
Yang et al's sparse method	17.0718	12.6237	15.6918	0.6
Block wavelet method	18.5995	12.7631	17.4133	0.8083
Directionlet method	22.8995	12.8233	19.7996	1.9528

From table 4.4 SNR values obtained for the image Barbara for zooming factor 2 and 4 are 22.8995dB and 19.7996 for directionlet method,18.5995dB and 17.4133dB for block wavelet method, 17.0718dB and 15.6918dB for Yang's method respectively. So is the case with the image 'bush'. From SNR values it is clear that directionlet method gives better result.

In Figure 4.27 (a)low resolution image(64x64) (b)original image(128x128) (c)original image(256x256)(d),(e),(f) 2 times SR images (128x128) using Yang et al method, block wavelet method and directionlet method(g), (h), (i) 4 times SR images (256x256) using Yang et al method, block wavelet method and new

directionlet method. Cross lines can be noticed on the scarf for the yang et al and block wavelet method while those are almost removed in the directionlet method.

Figures 4.28 show the corresponding results obtained for the image 'bush' for the different methods. Blocking effect along the edges of the leaves are noticed in Yang et al and block wavelet method. Triangular protrusions along the leaf are missing in directionlet method and but it is sharper than the images by the other 2 methods.

In Figures 4.29 and 4.30, LR images of size 128x128 are obtained using equation 3.1, from the original image of size 256x256 and they are super resolved by a of 2 and 4 using Yang's method and directionlet method. For image butterfly the ringing effect on the edges can be seen in both cases in Yang method, while it is much less in directionlet method. So also is the effect on the moustache of the tiger in figure 4.30. SNR computation has not been done as original image of size 512x512 is not available.

Figure 4.27: (a)low resolution image(64x64) (b)original image(128x128)(f)original image(256x256) (c),(d),(e)2 times SR images (128x128) using Yang et al method, block wavelet method and directionlet method (g),(h),(i) 4 times SR images (256x256) using Yang et al method,block wavelet method and new directionlet method

(a)

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

(f) (g) (h)

(i)

Figure 4.28: (a)low resolution image(64x64) (b)original image(128x128)(f)original image(256x256)(c),(d),(e)2 times SR images (128x128) using Yang et al method, block wavelet method and directionlet method(g),(h),(i) 4 times SR images (256x256) using Yang et al method,block wavelet method and new directionlet method

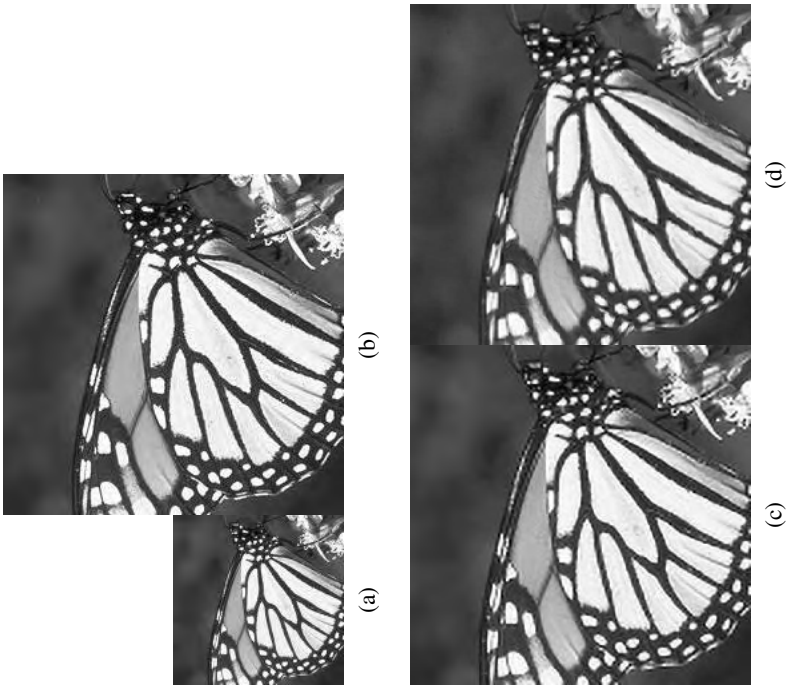


Figure 4.29: contd....

(e) (f)

Figure 4.29: (a)low resolution image(128x128) (b)original image (256x256)(c),(d)super resolved image with magnification factor 2 using Yang et al sparse method and new directionlet method, e),(f) super resolved image with magnification factor 4, using Yang et al sparse method and new directionlet method respectively.

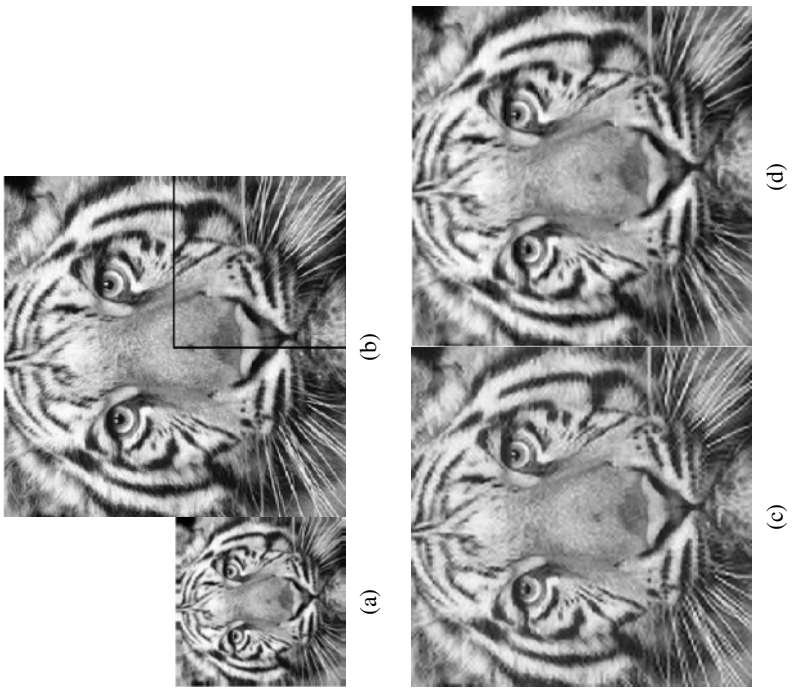
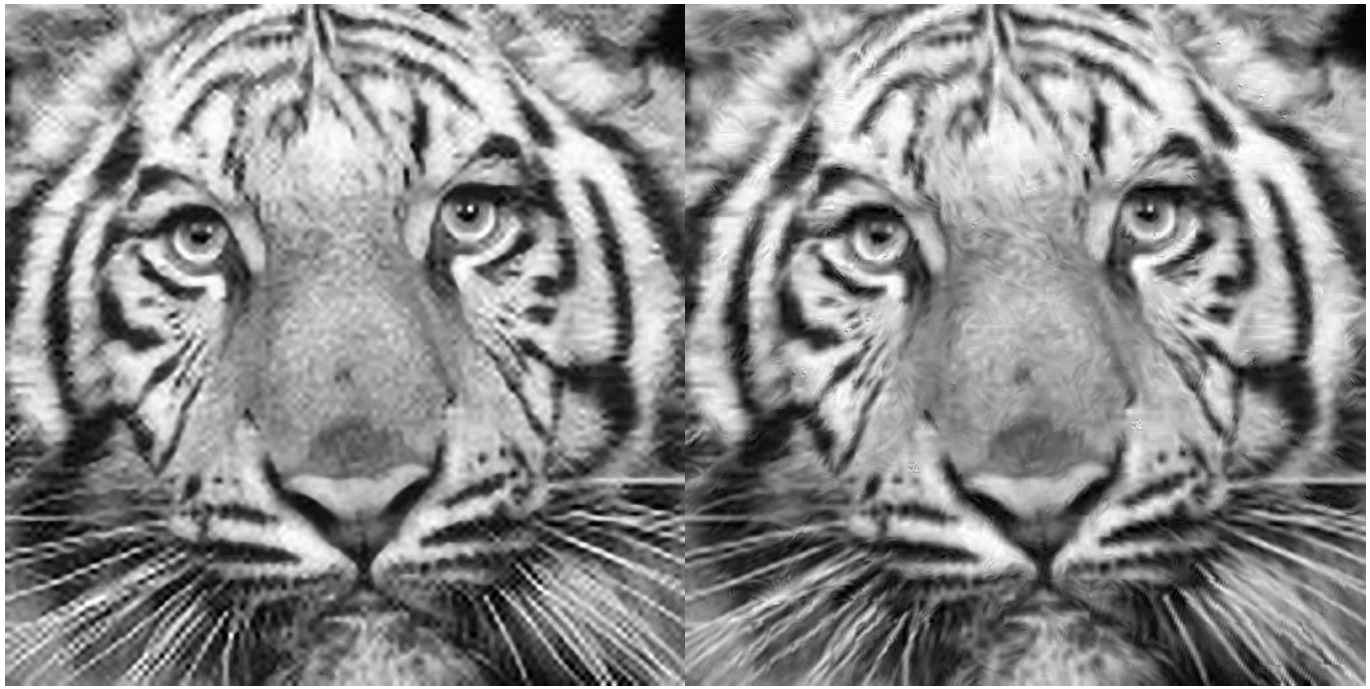


Figure 4.30: contd....



(e) (f)

Figure 4.30: (a)low resolution image(128x128) (b)original image (256x256)(c),(d)super resolved image with magnification factor 2 using Yang et al sparse method and new directionlet method, e),(f) super resolved image with magnification factor 4, using Yang et al sparse method and new directionlet method respectively.

4.4 Conclusion

In this chapter a novel single image super resolution method based on directionlet method is presented. The method is compared with Freeman method, wavelet 1 method, Sapan's method, Yang et al method and Block wavelet method. It is seen that directionlet method is better than the other methods. Super resolution for a factor of 4 is also implemented and results show that directionlet method is the best.

Chapter 5

Single Image Super Resolution Using Directionlets based on Lifting scheme

---

The last chapter explains a new method for single image super resolution using directionlet transform. Directionlet transform, also called skewed anisotropic wavelet transform (AWT(n_1,n_2)) can effectively capture anisotropic discontinuities present in an image unlike the isotropic wavelet transform. The advantage of the directionlet transform is that it retains the property of separability and one dimensional convolution simplicity of wavelet. Since convolution based directionlets are used, the new method presented in the last chapter needs more computations and hence more computation time. Lifting based wavelet transform is an alternative for conventional wave let transform which needs less computations, less time, and less memory. The lifting method is extended to the directionlet transform in this chapter and is used to super-resolve an LR image.

5.1 Introduction

The new single image super resolution using directionlet transform presented in this thesis produces high resolution images which are good both in terms of visual quality and SNR values. But like wavelet transform, the directionlet uses convolution operation and it is a computationally very intensive process, and hence faster implementation schemes are required. The Lifting based implementation of directionlet transform meets this requirement. Lifting scheme reduces the time required for this process by a factor of two or more with out any loss in quality. The lifting scheme is a simple and flexible tool for constructing pairs of forward and inverse wavelet transforms that can be adjusted to acquire various desired features while retaining the basic property to perfectly reconstruct the initial signal.

5.2 Background

5.2.1 Lifting scheme for wavelet transform

The lifting scheme for wavelet transform is explained in this section. This is extended to directionlet transform in section 5.2.4. The wavelet transform of a signal is a multi-resolution representation of a signal, which at each resolution level gives a highly uncorrelated representation. Discrete Wavelet Transform can be implemented using convolution, and several DWT architectures with filter convolution have been proposed. Filter banks are the fundamental tool to create discrete wavelet transforms. They are formed by the analysis and synthesis low and high pass filters and the intermediate stages composed by down and up sampling. However, such an implementation suffers many disadvantages. Due to convolution operation extra memory is needed to store the conventional DWT output. It takes more time to do the all operations. Conventional wavelet transform is called first generation wavelets.

Lifting scheme is a new approach to construct the so called second generation wavelets,that is wavelets which are not necessarily translations and dilations of one function[46]. This method was developed by Wim Sweldons in 1997 as a method to implement Discrete Wavelet Transform which overcomes above mentioned

disadvantages of first generation wavelet transform [31], [109]. Lifting enables constructing wavelets in spatial domain. DWT can be decomposed into a finite sequence of simple filtering steps, which are called lifting steps.

DWT can be viewed as prediction-error decomposition [82]. Using the wavelet transform a signal s_j with 2^j samples can be transformed into a coarser signal part s_{j-1} and a detail signal part d_{j-1}. The scaling coefficients at a given scale j are predictors for the data at the next higher resolution or scale (j-1). The wavelet coefficients are simply the "prediction errors" between the scaling coefficients and the higher resolution data. This interpretation has led to lifting scheme. It involves a polyphase decomposition or split of the original signal followed by Predict (P) and Update (U) steps.

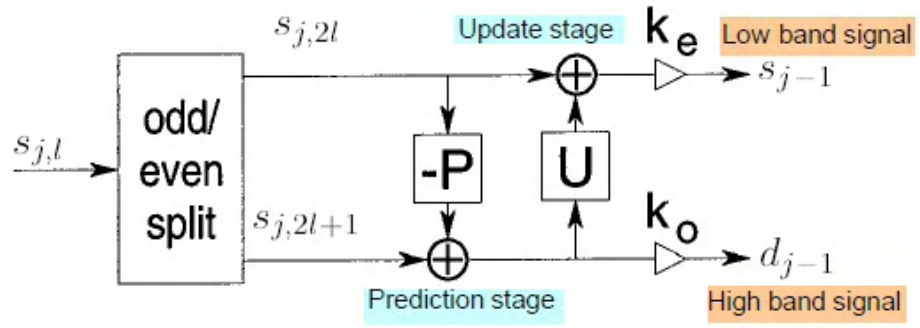


Figure 5.1: Lifting steps

The lifting based 1-D transform can be implemented as follows: There are three steps: split, predict and update which constitutes lifting stage as shown in Figures 5.1 and 5.2. The lifting steps are given below [83], [81] :

1. Split : The signal is split into two disjoint set of samples even indexed samples and odd indexed samples. Each set contains half samples of original

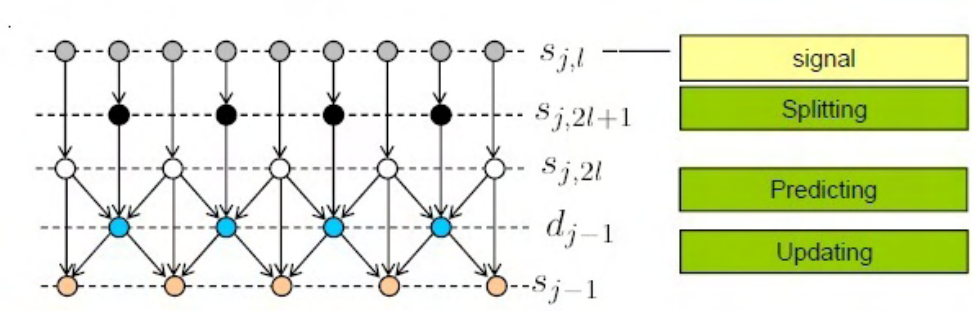


Figure 5.2: Lifting steps

signal. The process of splitting is called the lazy wavelet transform.

$$s_{j,2l} = s_{2l} \tag{5.1}$$

$$s_{j,2l+1} = s_{2l+1} \tag{5.2}$$

2. Predict(Dual lifting): The even and odd subsets are interspersed. If the signal has a local correlation structure, the even and odd subsets will be highly correlated with each other. In other words, given one of the two sets, it should be possible to predict the other one with reasonable accuracy. In this prediction step, which is also called dual lifting, the odd indexed samples $s_{j,2l+1}$ are predicted using the neighboring even indexed samples $s_{j,2l}$ and the prediction error. The difference between the odd sample and its prediction (detail or wavelet coefficients) replaces the original odd sample values, thus providing in-place calculations and are called detail (wavelet) coefficients.

$$d_{j-1,l} = s_{j,2l+1}(n) - P(s_{j,2l}) \tag{5.3}$$

3. Update : In the second lifting step, known as primal lifting (U), the even samples are replaced with smoothed values using the update operator U on previously computed details. The U operator is designed to maintain the

correct running average of the original sequence, in order to avoid aliasing.

$$s_{j-1,l} = s_{j,2l} + U(d_{j-1,l}) \tag{5.4}$$

4. Normalize: The outputs from the lifting steps are weighted by k_e and k_o. These values serve to normalize the energy of the underlying scaling and wavelet functions, respectively.

Iteration of the lifting stage, creates the complete set of DWT scaling and wavelet coefficients s and d. The output from the s channel after primal lifting steps provides a low pass sub band, whereas the output from d channel, after dual lifting steps provides a high pass sub band.

The lifting steps are easily inverted, even if P and U are nonlinear, space-varying, or non-invertible. The four steps in inverse lifting steps are: (i) Undo Normalize (ii) Undo Update (iii) Undo Predict (iv) Merge. The steps of inverse filtering is shown in Figures 5.3 and 5.4 [81].

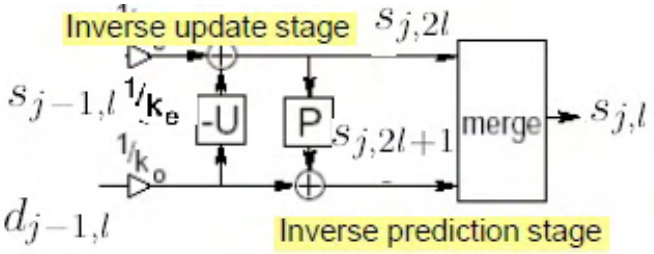


Figure 5.3: Inverse lifting steps

1. Undo split: The even samples are obtained by subtracting update information

$$s_{j,2l} = s_{j-1,l} - U(d_{j-1,l}) \tag{5.5}$$

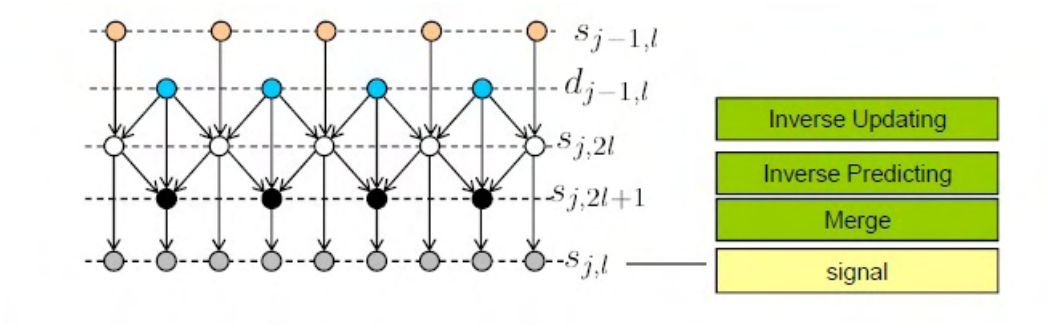


Figure 5.4: Inverse lifting steps

2. Undo Predict: The odd samples are obtained by adding the prediction information.

$$s_{j,2l+1} = d_{j-1,l} + P(s_{j-1,l}). \tag{5.6}$$

3. Merge(Inverse lazy wavelet): The even and odd samples are merged together to recover the original signal.

The inverse transform is found by reversing the order of the operations and flipping the signs.

Advantages of lifting scheme

Advantages of lifting based wavelet transform over conventional wavelet transform are as follows:

1. Lifting reduces the computational complexity of DWT involved with the convolution implementation.

2. The lifting scheme allows a fully in-place calculation of the wavelet transform. The extra memory required to store the results of the convolution can also be reduced by in place computation of the wavelet coefficient with the lifting scheme.

3. The inverse wavelet transform can be obtained by undoing the operations of the forward transform. In practise, this is done by changing each '+' operation into a '-' operation and vice versa. It has the same computational complexity as the forward transform.

4. Transforms signals with an arbitrary length (need not be of power of 2, that is 2^n).

5. All wavelet filters can be implemented using the lifting scheme.

6. Conventional wavelet transform convert floating point numbers to floating point numbers. Lifting scheme can be converted into a transform that maps integers to integers. At the same time it retains the perfect reconstruction property.

.

5.2.2 Wavelet transform in polyphase form

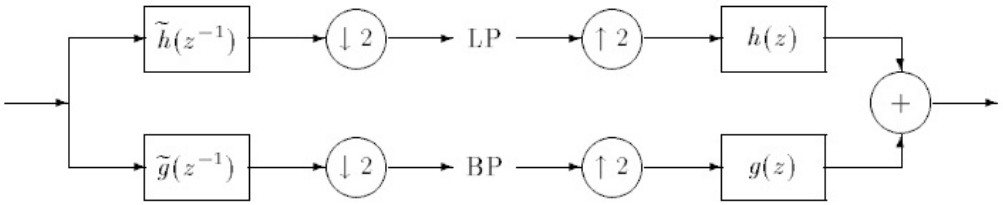


Figure 5.5: Filterbank representation of wavelet transform

Wavelet transform can be represented using multirate filter bank. Filter bank representation of DWT is shown in Figure5.5 [81].

The forward transform uses two analysis filters $\tilde{h}$ and $\tilde{g}$ followed by subsampling, while the inverse transform first upsamples and then uses two synthesis filters h

and g. The conditions for perfect reconstruction are given by equation 5.7.

$$h(z)\tilde{h}(z^{-1}) + g(z)\tilde{g}(z^{-1}) = 2$$

$$h(z)\tilde{h}(-z^{-1}) + g(z)\tilde{g}(-z^{-1}) = 0 \tag{5.7}$$

Each filter can be represented using polyphase representation. A filter h(z) can be represented by

$$h(z) = \sum_{k=-\infty}^{\infty} h(k)z^{-k} \tag{5.8}$$

Using polyphase representation, discrete time filter can be represented by the equation 5.9, where h_e and h_o are even and odd coefficients [73].

$$h(z) = h_e(z^2) + z^{-1}h_o(z^2) \tag{5.9}$$

$$h_e[z] = \sum h_{2k}z^{-k} \tag{5.10}$$

$$h_o[z] = \sum h_{2k+1}z^{-k} \tag{5.11}$$

The polyphase decomposition can be used to implement filter bank in efficient manner. The wavelet transform is represented by schematically in Figure 5.6. P(z) is called polyphase matrix for synthesis and is represented by;

$$P(z) = \begin{bmatrix} h_e(z) & g_e(z) \\ h_o(z) & g_o(z) \end{bmatrix} \tag{5.12}$$

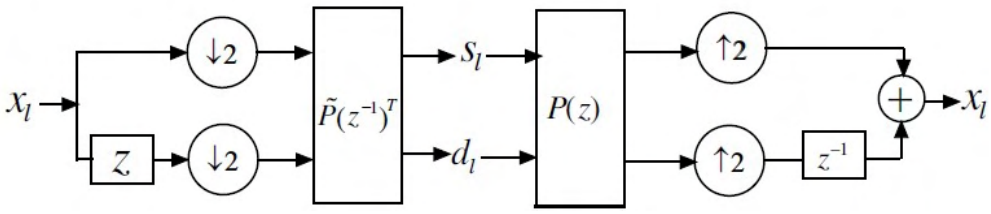


Figure 5.6: Polyphase representation of wavelet transform

In polyphase representation involves splitting the sequence into odd or even and then apply the poly phase matrix. In inverse transform polyphase matrix is followed by merging odd and even samples. The perfect reconstruction property is given by;

$$P(z)\tilde{P}(z^{-1})^t = I \tag{5.13}$$

where $\tilde{P}(z)$ is called polyphase matrix (dual poly phase matrix) for analysis bank.

Another concept used here is that filtering a signal with $H_e(z^2)$ and then down sampling by two is equivalent to down sampling the signal by two and then applying $H_e(z)$. The problem of finding an FIR wavelet transform thus amounts to finding a matrix P(z) with determinant 1. If matrix P(z) is known, $\tilde{P}(z)$ and the other four filters for the wavelet transform can be found out using following equations.

$$h_e(z) = g_o(z^-1) \tag{5.14}$$

$$h_o(z) = -g_e(z^-1) \tag{5.15}$$

$$g_e(z) = h_o(z^-1) \tag{5.16}$$

$$g_o(z) = -h_e(z^-1) \tag{5.17}$$

5.2.3 Polyphase representation of lifting method

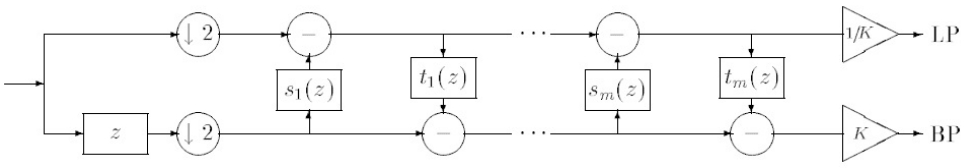


Figure 5.7: Lifting based forward wavelet transform

The polyphase representation provides an effective way of representing lifting method. As already explained, lifting consists of three steps: split, predict, and update. The odd/even split is another form of polyphase domain. For a given complementary filter pair (h, g), the polyphase matrix in equation 5.12 can be factored using Laurent polynomials $s_i(z)$, $t_i(z)$ and non zero constant K. Euclidian algorithm can be used to factor poly phase matrix. The equation 5.18 shows the factorisation of P(z).

$$P(z) = \prod_{1}^{m} \begin{bmatrix} 1 & s_i(z) \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ t_i(z) & 1 \end{bmatrix} * \begin{bmatrix} K & 0 \\ 1 & 1/K \end{bmatrix} \tag{5.18}$$

The dual polyphase matrix is given by

$$\tilde{P}(z) = \prod_{1}^{m} \begin{bmatrix} 1 & 0 \\ -s_i(z^{-1}) & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ -t_i(z^{-1}) & 1 \end{bmatrix} * \begin{bmatrix} 1/K & 0 \\ 1 & K \end{bmatrix} \tag{5.19}$$

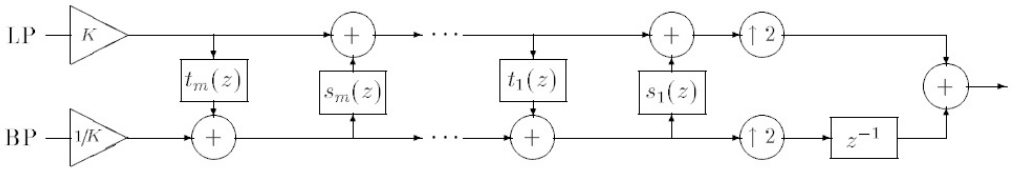


Figure 5.8: Inverse wavelet transform using lifting

The polynomial $s_i(z)$ represents primary lifting steps and $t_i(z)$ represents dual lifting steps. Figure5.7 shows forward wavelet transform using lifting and it includes lazy wavelet, alternating lifting and dual lifting steps and then scaling. Figure 5.8 shows inverse wavelet transform using lifting and it includes scaling, alternate dual lifting and lifting steps and finally inverse lazy transform.

5.2.4 Lifting scheme for directionlet transform

DWT is a special case of directionlet transform. So the lifting method applied to DWT can be extended to directionlet transform also. Lifting based directionlet also offers above mentioned advantages over conventional directionlet transform. Directionlet is based on lattice based filtering and subsampling. The lifting framework can be used to obtain the lattice based transform[40]. To apply a separable 2-D DWT on a 2-D signal like image using the lifting implementation, 1-D DWT with lifting implementation in the vertical and horizontal dimensions can be cascaded. To implement directionlet transform, lifting based wavelet transform is applied along selected pair of directions. I-D transform with directional prediction is performed by choosing the pixels from which a prediction (or update) is formed . These pixels are chosen along a direction which is not necessarily the horizontal or vertical direction. The lifting steps are as follows:

1. Split step : Split the pixels of lattice A located along the transform direction d_1, x(n) into two disjoint subsets: the odd and even polyphase samples $\mathrm{x}_o(n)$ and $\mathrm{x}_e(n)$. Although any disjoint split is possible, the original data set x(n) into $x_e(n) = x_(2n)$ and $x_o(n) = x(2n+1)$.

2. Prediction step: The wavelet coefficients or detail signal d(n) is generated as error in predicting $x_o(n)$ from $x_e(n)$ using prediction operator P, by keeping even samples changeless. The difference between the prediction value of $x_o(n)$ and the real value of $x_o(n)$ is defined as the high-frequency component or detail signal.

$$d(n) = x_o(n) - P(x_e(n)) \tag{5.20}$$

3. Update step:This detail coefficients are used to update even samples $x_e(n)$, to obtain the approximate signal which creates the low-frequency component. $x_e(n)$ and $d(n)$ are combined to obtain scaling coefficients c(n) which is a coarse approximation to the original signal x(n). This is done by applying an update operator U to the wavelet coefficients d(n) and adding the result to $x_e(n)$:

$$c(n) = x_e(n) + U(d(n)) \tag{5.21}$$

By performing the above processing along all co-lines in all cosets in both directions, the multi resolution lifting directionlet decomposition is obtained.

5.3 Single Image Super Resolution using directionlets based on lifting scheme

Directionlet transform based on lifting method is implemented and is used to super resolve a low resolution image to high resolution image. db4 wavelet is used for implementation.

5.3.1 Implementation

Implementation of db4 wavelets

Wavelet coefficients for db4 are shown in table5.1. The lifting coefficients are different for different wavelets. These coefficients are obtained from wavelet coefficients. Table5.2 shows the db4 coefficients used in the lifting scheme.

Euclid algorithm can be used to find DWT with a finite number of lifting steps starting from polyphase transform [31]. db4 wavelet transform can be expressed into lifting steps as follows,

1. db4 wavelet is an orthogonal base. Inverse transform uses two synthesis

Table 5.1: Wavelet coefficients of db4

Decomposition low pass filter,h		Decomposition high pass filter,g	
h_0	-0.0106	g_7	-0.2304
h_1	0.0329	g_6	0.7148
h_2	0.0308	g_5	0.6309
h_3	-0.1870	g_4	-0.0280
h_4	-0.0280	g_3	-0.1870
h_5	0.6309	g_2	0.0308
h_6	0.7148	g_1	0.0329
h_7	0.2304	g_0	-0.0106

filters h and g. Under the perfect reconstruction condition,

$$h(z) = h(0) + h_1 z^{-1} + h_2 z^{-2} + h_3 z^{-3} + h_4 z^{-4} + h_5 z^{-5} + h_6 z^{-7} + h_7 z^{-8} \tag{5.22}$$

$$g(z) = h_7 z^6 + h_6 z^5 + h_5 z^4 + h_4 z^3 + h_3 z^2 + h_2 z^1 + h_1 z + h_0 6 z^{-1} \tag{5.23}$$

Polyphase representation of synthesis filters is given by equations 5.24 and 5.25.

$$h(z) = h_e(z^2) + z^{-1} h_o z^2 \tag{5.24}$$

$$g(z) = g_e(z^2) + z^{-1} g_o z^2 \tag{5.25}$$

h_e contains the even coefficients, and h_o contains the odd coefficients.

2. Using polyphase representation of db4 synthesis filters h and g, synthesis polyphase matrix P(z) is obtained. Synthesis polyphase matrix is given by equation 5.26

$$P(z) = \begin{bmatrix} h_e(z) & g_e(z) \\ h_o(z) & g_o(z) \end{bmatrix} \tag{5.26}$$

$$h_e(z) = h_0 + h_2 z^{-1} + h_4 z^{-2} + h_6 z^{-3} \tag{5.27}$$

$$h_o(z) = h_1 + h_3 z^{-1} + h_5 z^{-2} + h_7 z^{-3} \tag{5.28}$$

$$g_e(z) = h_7 z^3 + h_5 z^2 + h_3 z + h_1 \tag{5.29}$$

$$g_o(z) = -h_6 z^3 - h_4 z^2 - h_2 z^1 - h_0 \tag{5.30}$$

Polyphase representation of wavelet transforms is shown in Figure5.6 .

3. Using Euclidean algorithm for Laurent polynomial, synthesis polyphase matrix P(z) can be factored into lifting steps.

$$\begin{aligned} P(z) = &\begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} * \begin{bmatrix} 1 & (\beta z^{-1} + \beta') \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ (\gamma z^{-1} + \gamma') & 1 \end{bmatrix} \\ &\begin{bmatrix} 1 & (\eta z^{-1} + \eta') \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\ &* \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & -(\lambda_3 z^3 + \lambda_2 z^2 + \lambda_1 z + \lambda_0) \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} -K^{-1} & 0 \\ 0 & K \end{bmatrix} \end{aligned} \tag{5.31}$$

The coefficients α, β etc in above equation can be obtained from Matlab 7.0 by the command function 'liftwave()', and are listed in table 5.2.

4. Obtain analysis polyphase matrix $\tilde{P}(z^{-1})^T$.

$$\tilde{P}(z^{-1})^T = P(Z^{-1})^T \tag{5.32}$$

The analysis polyphase matrix is factored as follows

$$\tilde{P}(z^{-1})^T = \begin{bmatrix} -k_1^{-1} & 0 \\ 0 & K_1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ -(\lambda_3 z^{-3} + \lambda_2 z^{-2} + \lambda_1 z^{-1} + \lambda_0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ (-1 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ (\eta z + \eta') & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ -(\gamma z + \gamma') & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ (\beta z + \beta') & 1 \end{bmatrix} * \begin{bmatrix} -K^{-1} & 0 \\ 0 & K \end{bmatrix} \quad (5.33)$$

This corresponds to the following implementation for the forward transform:

$$s(l)^1 = x(2l+1) + \alpha x(2l) \quad (5.34)$$
$$d(l)^1 = -\beta s_{l+1}^1 + \beta s_l^l + x_{2l+1} \quad (5.35)$$
$$s(l)^2 = s_l^1 - \gamma d_{l+1}^1 - \gamma^1 d_l^1 \quad (5.36)$$
$$d(l)^2 = -\eta s_{l+1}^2 - \eta^1 s_l^2 + d_l^1 \quad (5.37)$$
$$s(l)^3 = s(l)^2 + d(l)^2 \quad (5.38)$$
$$d(l)^3 = -s(l)^3 + d(l)^2 \quad (5.39)$$
$$d(l)^4 = -\lambda_3 s_{l-3}^3 - \lambda_2 s_{l-2}^3 - \lambda_1 s_{l-1} + s(l)^3 + d_l^3 \quad (5.40)$$
$$s(l) = -k_l^{-1} s(l)^3 \quad (5.41)$$
$$d(l) = -k_1 d(l)^4 \quad (5.42)$$

where s_l and d_l are smoothed values and details respectively.

5. Derive the inverse transform from the forward by running the scheme backward [110].

Table 5.2: Coefficients in lifting steps

k_1	1.362166720130752
λ_0	-1
λ_1	0.469083478901698
λ_2	0.14003923772683
λ_3	0.024791238156143
η	2.131816712755221
γ^1	0.117648086798478
γ	0.018808352726244
β^1	0.300142258748545
β	1.117123605160594
α	0.322275887997141

Figure 5.9: Training set image

Implementation of directiolet transform

Here the directionlet transform is implemented by applying lifting based 1-D wavelet transform along selected pair of directions. Here also lattice structure based filtering and subsampling are used. The same set of directions used in conventional directionlet method are used here.

Development of training set

The training set is obtained from high resolution images downloaded from Internet. One of the training set is shown in Figure5.9. This image is selected as training set image because it contains information in five pair of directions. It is of size 333x500 and in tiff format.

5.3.2 Results and discussion

Table 5.3: SNR values for different images with traditional directionlet method and lifting based directionlet method

Method	SNR in dB		
	Barbara	Butterfly	Tiger
directionlet method	20.3445	26.7953	25.1543
lifting based directionlet method	20.3513	26.8020	25.2685

Table 5.4: Time taken to generate training set with high resolution images of different size with traditional directionlet method and lifting based directionlet method

Input images		time in seconds	
name	size	Directionlet method	Lifting based directionlet method
image 1	256x256	625.551540	294.186
image2	500x233	1698.824	855.525

SNR values obtained for different standard images using conventional directionlet transform and lifting scheme based are shown in table 5.3. The time taken for training set formation and learning process is shown in tables 5.4 and 5.5

respectively. The SNR value obtained for super resolving image1 of size 256x256 to the size of 512x512 with conventional directionlet method is 20.3445dB while with lifting scheme based method it is 20.3513dB. SNR values show that both methods give almost same SNR values.

From the table 5.4, it can be seen that the time taken to build a training set with an image of size 500×333 using lifting scheme method is 855.525 seconds while it takes 1698.824 seconds using conventional directionlet method. It is 294.186 seconds with new lifting based method while with conventional directionlet method it is 625.551540 seconds for making a training set with an image of size 256×256.

It is clear from the table 5.5 that time taken to super resolve an image of size 256×256 to the size of 512x512, using conventional directionlet method is 1841.033seconds while that with lifting based directionlet it is 973.135 seconds. While the average time to convert an image of size 128×128 to an image of size 256×256 the traditional directionlet method needs 1142.806 seconds, lifting method requires only 288.523seconds.

Table 5.5: Time taken to super resolve low resolution images of different size with traditional directionlet method and lifting based directionlet method

Input images		time in seconds	
name	size	Directionlet method	lifting based directionlet method
image 1	500x233	3684.766759	2017.318
image2	128x128	1142.806190	288.523488
image3	256x256	1841.033	973.135

Figures 5.10, 5.11, 5.12, show super resolved images using conventional directionlet transform and lifting based directionlet transform. Figure 5.10 (a) is the low resolution image of size 256x256 (b) is the original image of size 512x512 (it is shown here for comparison), (c) super resolved image of size 512x512 using lifting method, (d) super resolved image using conventional directionlet method.

(a) (b)

(c) (d)

Figure 5.10: (a)low resolution image (b)original image(c)super resolved image using lifting based directionlet transform(d)super resolved image using conventional directionlet transform

Figure 5.11 (a) is the low resolution image of size 288x360 (b) is the original image of size 576x720 (it is shown here for comparison), (c) super resolved image

of size 576x360 using lifting method,(d) super resolved image using conventional directionlet method. From them, it is clear that both methods give almost same result.

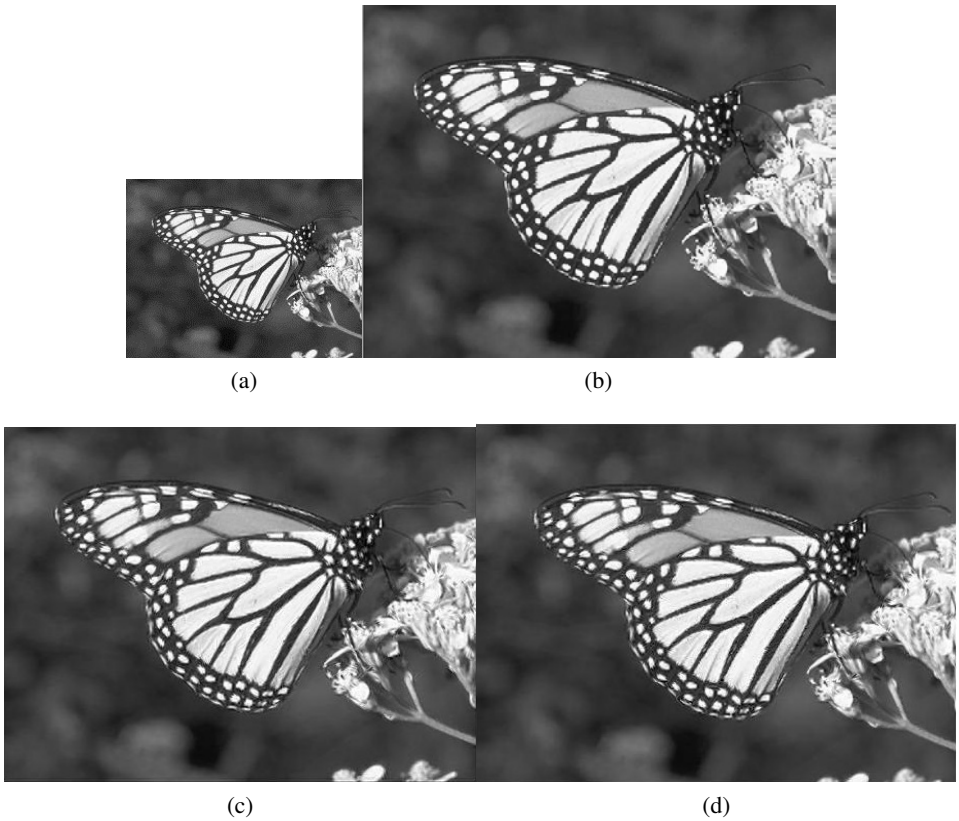
(a) (b)

(c) (d)

Figure 5.11: (a)low resolution image (b)original image(c)super resolved image using lifting based directionlet transform(d)super resolved image using conventional directionlet transform

Figure 5.12 (a) is the low resolution image of size 173x230 (b) is the original image of size 346x460 (it is shown here for comparison), (c) super resolved image

of size 346x460 using lifting method, (d) super resolved image using conventional directionlet method. From them, it is clear that both methods give almost same result. It is also obvious that using new directionlet method low resolution image of any size can be super resolved.



(a) (b)

(c) (d)

Figure 5.12: (a)low resolution image (b)original image(c)super resolved image using lifting based directionlet transform(d)super resolved image using conventional directionlet transform

5.4 Study of the effect of different wavelets in directionlet transform

The effectiveness of the directionlet transform in super resolution methods depends on the choice of the basis function or mother wavelet. This section proposes, a study of the different wavelet bases for a directionally adaptive, learning based, single image super resolution method using Directionlets. There are many available types of wavelet families, such as Daubechies, Meyer, Gaussian, Mexican Hat, Morlet and many more. Four types of wavelet families were examined: Haar Wavelet (HW), Daubechies Wavelet(db), Biorthogonal Wavelet(BW) and Symlet wavelet. Implementation using lifting scheme are done with the different wavelets. Patch size depends on length of wavelet coefficients.

The experiments are done on natural images of size 128x128 with wavelet basis har, db4 ,db5, bior1.3, bior3.3, bior4.4 and sym4 etc. SNR values with different wavelets are shown in table 5.6. SNR values and visual quality of resultant images show that rbior1.5 gave the best results followed by rbior1.3, bior 3.3, db5, db4 etc. SNR obtained for super resolved image boat with rbior1.5 is 29.488dB, while with bior3.3 is 28.8175dB. Figures 5.13(a), (b),(c),(d) and (e) shows main results obtained with wavelets db5, db4, b33, rb1.3, rb1.5, give good results. This figure also shows that rbior1.5 gives good visual effect compared with other wavelets.

But rbior1.5 has 10 coefficients and this necessitates that the block size of the high resolution patch should be 24×24 and the computation time is more compared with wavelets like db4, bior3.3. The time taken to super resolve an image of size 256×256 with wavelet rbior1.5 is 1800seconds and while it is 750seconds with bior3.3 or db4.

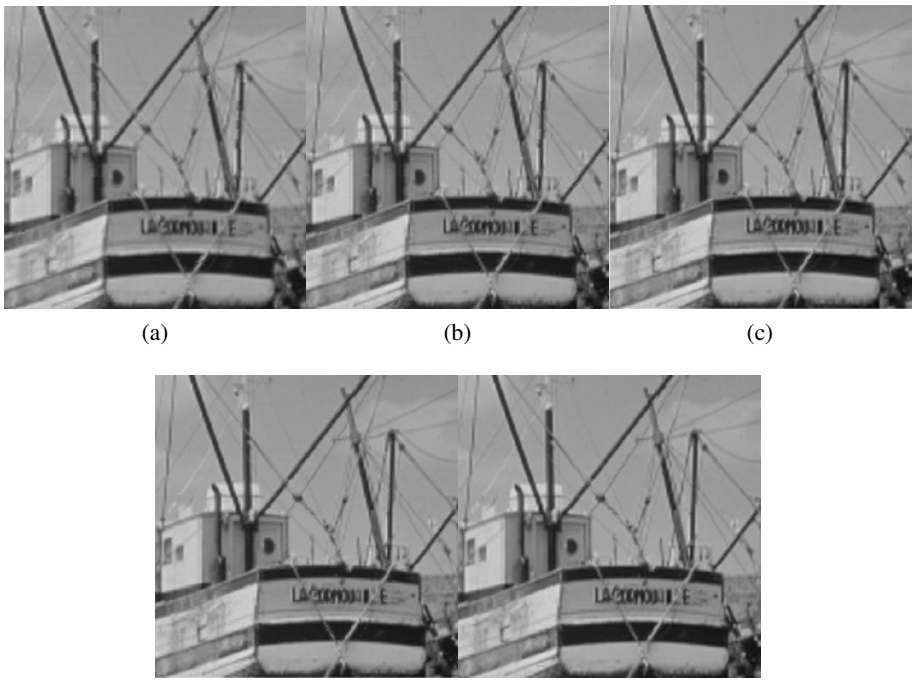
Table 5.6: Comparison of SNR of different images super resolved with different wavelets

method	SNR in dB		
	Boat	Butterfly	Tiger
db1	23.6123	22.7688	21.3765
db3	21.2480	20.5271	19.470
db4	**26.9486**	**25.5178**	**24.051**
db5	26.9486	24.1837	22.8421
bior1.3	**27.4063**	**21.7692**	**19.9327**
bior2.2	21.8980	21.61294	20.224
bior2.4	24.2918	24.1294	21.224
bior3.3	**28.8175**	**26.802**	**25.2686**
bior4.4	23.2918	22.8601	22.0323
sim3	20.9307	20.024	18.6481
sim4	25.5432	24.1286	23.0283
rbior1.3	**29.488**	**26.1995**	**24.732**
rbior1.5	**29.8164**	**27.9769**	**25.2685**
coif1	28.0486	21.9551	20.2354

(a) (b) (c)

(d) (e)

Figure 5.13: Super resolved images using (a) db5, (b)db4, (c)b3.3 (d)rb1.3 (e)rb1.5



(a) (b) (c)

(d) (e)

Figure 5.14: Super resolved images using (a) db5, (b)db4, (c)bior3.3 (d)rb1.3 (e)rb1.5

5.5 conclusion

In this chapter a new method for super resolving a low resolution image using learning based approach is presented. This is done using directionlet transform. Instead of the conventional filter bank implementation scheme, this chapter proposes a computationally less intensive method, namely lifting based directionlet transform. As a result, tremendous reduction in computation has been achieved. Lifting scheme is faster than the Filter Bank scheme; especially when the filter has more taps. The capability and time saving achieved by combining directionlet transform and lifting scheme will be very useful in real time super resolution problems. db4 wavelet has been chosen for the simulation. Effects of other wavelets on this super resolution method is also studied.

Chapter 6

Single image super resolution in color images

Single image super resolution for grey level image is extended to color images in this chapter. Two color image super resolution methods are implemented. In the first method color image in RGB format is converted to YC_bC_r format. The luminance component Y alone is super resolved and other two components are interpolated using standard methods. At the end the YC_bC_r format is converted back to RGB format. In the second method the three color components R, G, B are super resolved separately to obtain super resolved color image. It is found that the second method needs more computation time compared with the first method.

6.1 Introduction

The previous chapters presented single image super resolution methods in grey images. In this chapter the method is extended to color images.

6.2 Overview

6.2.1 Color Fundamentals

The colors that human beings observe in an object are determined by the nature of the light reflected from the object. The visible light is composed of frequencies that ranges from 400 to 700 nanometers in the electromagnetic spectrum. A body that reflects light that is balanced in all visible wavelengths appears white to the observer. However a body that favors reflectance in a limited range of the visible spectrum exhibits some shades of color. For example, green objects reflect light with wavelengths primarily in the 500 to 570 nm range, while absorbing most of the energy at other wavelengths [76]. There are two ways to characterize the color of an object: color of the pigments (the color of the light that is reflected) or the color of the light that is absorbed. (the complementary of the pigment color)[98].

Three basic quantities are needed to describe the quality of a chromatic light source: radiance, luminance, and brightness. Radiance is the total amount of energy that flows from the light source, and is usually measured in watts. Luminance, measured in lumens, gives a measure of the amount of energy an observer perceives from a light source [76].

Cones are sensors in the eye responsible for color vision. Due to the absorption characteristics of the human eye, colors are seen as variable combinations of the so called primary colors red, green, and blue. These primary colors are added to produce the secondary colors of light-magenta (red plus blue), cyan (green plus blue), and yellow (red plus green). Mixing the three primaries or secondary with its opposite primary color, in the right intensities produces white light.

The characteristics generally used to distinguish one color from another are brightness, hue, and saturation. As already explained brightness embodies the chromatic notion of intensity. Hue is an attribute associated with the dominant wavelength in a mixture of light waves. It represents the dominant color as perceived by an observer. Saturation refers to the relative purity or the amount of white light mixed with hue. The pure spectrum colors are fully saturated.

6.2.2 Color Model

A color model is a specification of a coordinate system within which each color is represented by a single point. Several popular color formats are used for image and video processing. A color model is an abstract mathematical system for representing color and has 3 dimensional abstractions for three primary colors along three dimensions. They can represent only limited number of colors and hence often can't represent all colors in the visible spectrum.

Gamut or Color Space

The range of colors that are covered by a color model is called Gamut or color space. Color models can be classified as either additive or subtractive.

Additive color describes the situation where color is created by mixing or adding the visible light emitted from differently colored light sources. Additive color uses transmitted light to display color. Human perception is additive since black is the absence of light and white the presence of all wavelengths of light. Subtractive colors are in contrast to additive colors. In this type color system light is removed from various part of the visible spectrum to create colors.

Figure 6.1 shows additive and subtractive models. The most common examples for additive light are computer monitors and televisions. Subtractive color is used in paints and pigments and color filters. The additive reproduction process usually uses red, green and blue light to produce the other colors. Combining one of these additive primary colors with another in equal amounts produces the additive secondary colors cyan, magenta, and yellow. The different color models are

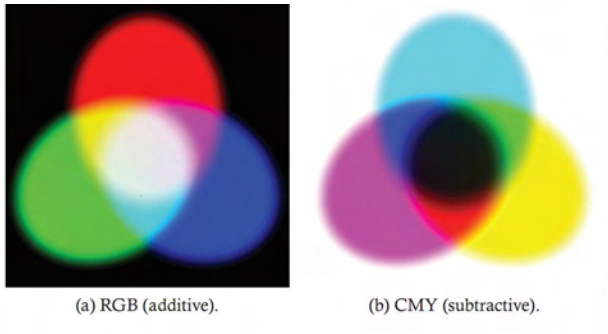


Figure 6.1: (a)Additive models(b)Subtractive models

- RGB Color Format
- CMY Color Format
- CMY(K) Color Format
- HSB Color Format
- YUV Color Format
- YIQ Color Format
- YCbCr Color Format

RGB Color Format

In the RGB model which is also called additive model, each color is represented as a combination of primary colors red, green, and blue. Figure 6.2 shows RGB color system. The primary colors can be added to produce the secondary colors of light - magenta (red plus blue), cyan (green plus blue), and yellow (red plus green). The combination of red, green, and blue at full intensities makes white. The importance of the RGB color model is that it relates very closely to the way that the human eye perceives color. This model is a basic color model for computer graphics because color displays use red, green, and blue to create the desired color. The number of bits used to represent each pixel in RGB space is called pixel depth.

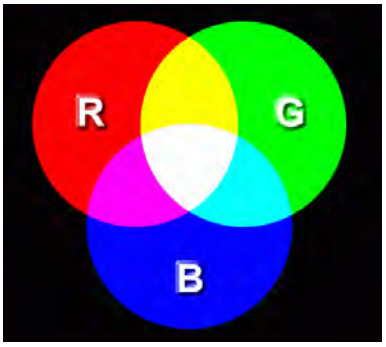


Figure 6.2: The RGB color system

A RGB color cube can be used to display smooth transitions between these colors. To generate any color within the RGB color cube, all three RGB components need to be of equal pixel depth and display resolution. This model is based on a Cartesian coordinate system. Schematic of RGB color cube is shown in Figure 6.3. Here RGB values are at three corners; cyan, magenta, and yellow are at three corners; black is at the origin; and white is at the corner farthest from the origin. The gray scale extends from black to white along the line joining these two points. Different colors are points on or inside the cube, and are defined by vectors extending from the origin [76]. Each color is in the range [0,1].

In RGB model, images consist of three component images, one for each primary color. Any modification of the image requires modification of all the three planes. If 8 bits are needed to represent each primary color image, then 24 pixels are needed to represent a RGB color pixel. The total number of colors in a 24 bit RGB image is 2^8.

CMY Color Format

Cyan, Magenta, and yellow are the secondary colors of light. When a surface coated with cyan pigment is illuminated with white light, no red light is reflected from the surface. That is Cyan subtracts (absorbs) red light from reflected white

light, which itself is composed of equal amounts of red, green, and blue light. The CMY color space is subtractive.

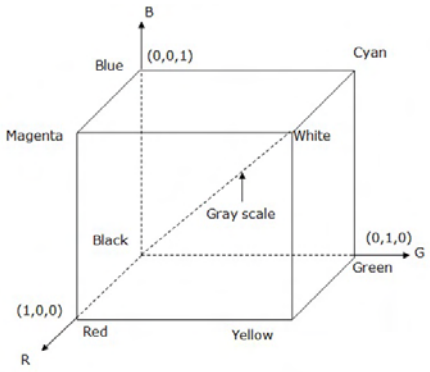


Figure 6.3: RGB color cube

Usually most devices like color printers and copiers convert RGB to CMY internally. Converting from RGB to CMY is done using the equation 6.1.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.1}$$

Here normalised color values are used. Equation 6.1 demonstrates that light reflected from a surface coated with pure cyan does not contain red (C=1-R), pure magenta does not contain green, yellow does not contain blue.

CMYK Color Format

In actual practice, mixing maximal amounts of cyan, magenta and yellow pigments creates a color that is not really black (dark muddy instead). So in order to produce true black a fourth color black is added giving rise to the CMYK color model.

Figure 6.4: HSB

HSB Format

Color models like RGB, CMY consider a color image as being composed of three primary images that combine to form single image. They are not suited for describing colors for practical human interpretation. HSB color model decomposes color according to perception rather than according to how it is physically sensed. A point within the HSB gamut is defined by Hue (the chromaticity or pure color), Saturation (the vividness or dullness of the color), Brightness(the intensity of the color).

According to this model, any color is represented by 3 numbers as in Figure6.4. The first number is the hue, and its value ranges from 0 to 360 degrees. Each degree represent a distinct color. First there is the red color (0 or 360 degrees,) and then there are all other colors (for example yellow at 120 degrees, green at 180 degrees and blue at 240 degrees), up to the violet color. All the VIBGYOR colors are represented here. The second number is the saturation. It represents the amount of color or, more exactly, its percentage. It is the purity of the color and is the amount of pure color mixed with white color. Its value ranges from 0 to 100, where 0 represents no color, while 100 represents the full color. Finally, the third number is the brightness. One can enhance the color brightness adding the white color, and can reduce it by adding the black color. In this case 0 represents the white color and 100 represents the black color. The more this value tends to 0, the brighter the color is. The more this value tends to 100 the darker the color is. The RGB values have been normalised to the range [0,1]. Given an RGB color format,

H component of each RGB pixel is obtained by

$$\begin{aligned} H &= \theta; if B \leq G \\ &= 360 - \theta; if B > G \\ where, \theta &= \cos^{-1}\{\frac{1/2[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{\frac{1}{2}}}\} \end{aligned}$$

The Saturation component is given by

$$S = 1 - \frac{3}{(R+G+B)}[min(R,G,B)] \tag{6.2}$$

The intensity is given by

$$I = \frac{1}{3}(R+G+B) \tag{6.3}$$

This is used by NTSC, PAL, and SECAM television standards. Like HSB, YIQ separates color into luminance Y and color channels(I and Q). Y is the luminance or gray scale component, I is the in phase component (amount of red-green) and Q is the quadrature (amount of blue-yellow). The YIQ color space is a rotation and distortion of RGB such that the Y axis lies along the RGB gray scale, the I axis is oriented roughly to red-green and the Q axis to blue-yellow. Converting RGB to YIQ is done by the equation6.4.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.321 \\ 0.211 & 0.523 & -0.312 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.4}$$

YUV Color Model

The YUV color model is the basic color model used in analogue color TV broadcasting. Initially YUV is the recoding of RGB for transmission efficiency (minimizing bandwidth) and for downward compatibility with black and white television. The YUV color space is derived from the RGB space. It comprises the luminance (Y) and two color difference U, V components. The luminance

can be computed as a weighted sum of red, green and blue components; the color difference, or chrominance, components are formed by subtracting luminance from blue and from red.

YC_bC_r Color Model

Figure 6.5: A color image and its Y, C_b and C_r components.

YC_bC_r color space used for component digital video is a scaled and offset version of the YUV color space. Y is the luminance component and C_b and C_r are the blue difference and red-difference chroma components. Figure 6.5 shows a color image and its Y, C_b and C_r components. It can be noted that the Y image is essentially a grey scale copy of the original image. The principal advantage of the YC_bC_r model in image processing is decoupling of luminance and color information. The importance of this decoupling is that the luminance component of an image can be processed without affecting its color component. For example, the histogram equalization of the color image in the YC_bC_r format may be performed simply by applying histogram equalization to its Y component. There are many combinations

of YC_bC_r values from nominal ranges that result in invalid RGB values, because the possible RGB colors occupy only part of the YC_bC_r space limited by these ranges.

Conversion from RGB to YCbCr using equation6.5

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112 & -93.786 & -18.214 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.5}$$

6.3 Single image super resolution using directionlets for color images

There are different ways to extend the gray image super-resolution algorithm for processing color images in RGB format. The typical solution involves applying monochromatic SR algorithms to each of the color channels independently, while using the color information to improve the accuracy of motion estimation [90]. This process can result in better performance than super-resolving only the luminance components, both in terms of SNR values and in visual plausibility. But it is at the expense of three times larger run-time complexity. Humans are more sensitive to the brightness information(luminance) than color information(chrominance components). Another approach is transforming the problem to a different color space, where chrominance layers are separated from luminance, and super resolution method is applied only to the luminance channel. The chrominance or color channels, are then upsampled using interpolation methods (eg. bilinear, bicubic) and the final RGB is computed by recombining the new SR luminance image with the interpolated chrominance. Processing the luminance information does not reduce the quality of resultant image, but this method reduces the computation time[88].

Above mentioned two methods are implemented here. The simplest way is the second method which super resolves only the luminance component of a given

color image.

6.3.1 Color image Super resolution with super resolution on luminance component only

Methodology

In this method the RGB color image is converted into YC_bC_r model and the super-resolution is performed only on the Y component. The main advantage of the YC_bC_r model in image processing is that the luminance and the color information are independent. Thus, the luminance component can be processed without affecting the color contents. The details information in a digital image is mainly present in the image luminance component. There fore, one can take advantage of the high sensibility of the human visual system to the brightness variation than to the chrominance variation. Consequently, more computational resources can be allocated to enlarge the brightness values while color components can be enlarged using a simpler approach. The final result is then obtained by combining the super-resolved Y component with the interpolated C_b and C_r components. As mentioned earlier, this scheme has the important advantages of enabling the gray level model being applied directly to color images, resulting in the run-time complexity to be the same as that of gray level case. Finally, the YC_bC_r model is converted to the RGB model to generate the synthetic image.

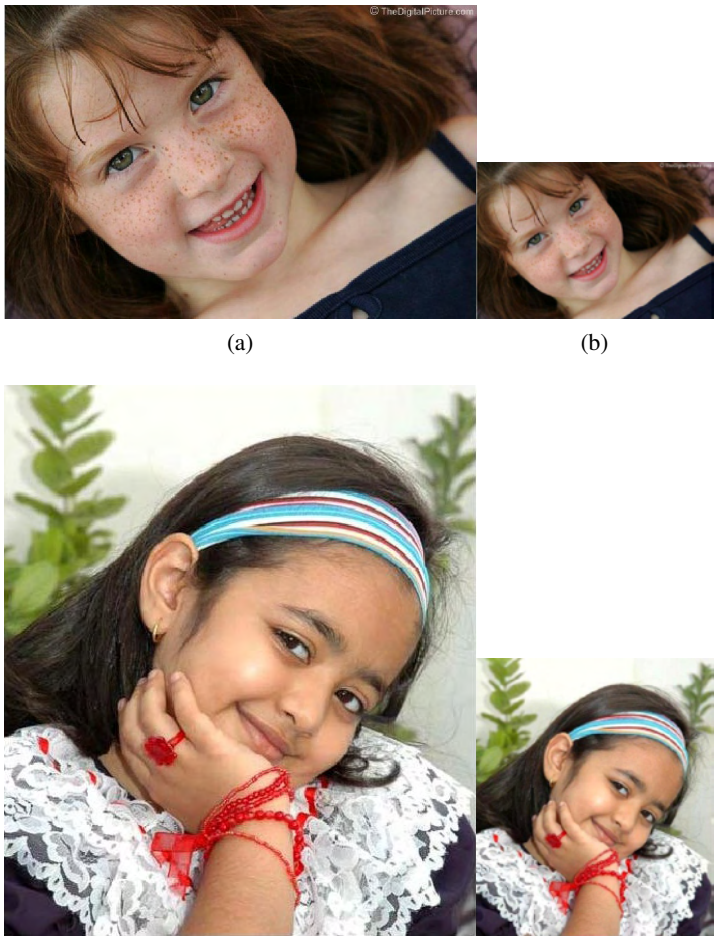
Implementation

High resolution color images are obtained from the Internet and are used to form training set. The images are obtained from Canon digital Gallery [1]. One of the training set image is shown in Figure6.6. It is of size 333x500 and in tiff format. Some of high and corresponding low resolution images are shown in Figure 6.7. The low resolution images are obtained by subsampling the each R, G, B components of their high resolution images separately.

In this method the RGB formatted images are converted into YC_bC_r format. The luminance component Y alone is super-resolved using the new directionlet method. For this the training set is formed using the luminance components of high and low

Figure 6.6: Training set color image

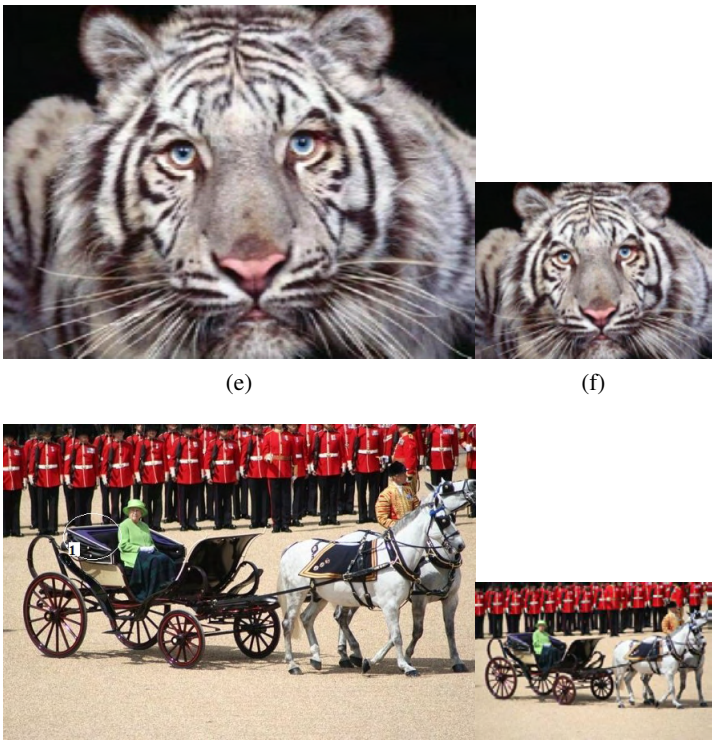
resolution color images. The low and high resolution luminance components are divided into patches and directionlet transform is applied on these patches. These coefficients form the training set. The missing high resolution Y component is learned from this training set. The other two components C_b and C_r are upsampled using cubic spline interpolation. The super resolved Y component and cubic spline interpolated C_b, C_r components are converted back to RGB format to obtain the high resolution image.



(a) (b)

(c) (d)

Figure 6.7: (a),(c)Original images of girl1 and girl2(b),(d)Low resolution images, contd.....



(e) (f)

(g) (h)

Figure 6.7: (e),(g) Original images of tiger and queen (f),(h)Low resolution images

Results

For color images, results obtained using new directionlet method is compared with Yang et al sparse method [36] and Shan et al method [75]. The paper[75], is based on an image formation process that models how a coarselevel image is produced from its clear finer-resolution version.

SNR values are calculated and shown in table 6.1 for new directionlet method, Yang et al and Shan et al super resolution methods. Table shows that the super resolved image girl1 using directionlet method has SNR of 43.5548dB while super resolved images using Yang et al method and Shan et al method have SNR of

Table 6.1: SNR values of Super resolved Color images

method	SNR in DB			
	girl1	girl2	tiger	queen
super resolved using Shan et al method	39.2839	32.1571	23.7560	25.5211
super resolved using Yang et al method	39.3904	33.8038	26.6926	27.6746
Directionlet method	43.5548	37.3297	29.2922	31.5619

39.3904dB and 39.2839dB only. Same is the case with the images of girl2, tiger and queen.

Figure 6.8 (a), (c), (e), (g) show original image, super resolved images of low resolution image Figure6.7(b) of girl1 using existing methods (Yang et al method and Shan et al method) and new directionlet method. Here the low resolution image of size 166×250 is super resolved to an image of size 332×500. Figures 6.8(b), (d), (f), (h) represent the zoomed portions of forehead of girl in Figure (a), (c), (e), (g) respectively. It is clear that the image obtained by directionlet super-resolution method, consists of sharper details than the images from other methods. Blocking effect is present in hair strands of image using existing methods but it is almost removed in the new method.

Figure 6.9 shows another color image of girl2 in which new super resolution method is applied. The low resolution image is of size 215×251 and is shown in Figure 6.7(d). Figure 6.9 (b), (d), (f), (h) show zoomed portions of original image(a), super resolved using Yang et al method (c), Shan et al method (e) and super resolved method using directionlet (g) respectively. The artifacts in the hair band present in super resolved image using existing method is almost removed in directionlet based super resolution method. Also the eyebrows and hair strands are sharper and close to original in the directionlet method. Table 6.1 shows that the directionlet based super resolved image has SNR of 37.3295dB while super resolved images with Yang et al method and Shan et al method have SNR value 33.8038dB and 32.1571 respectively.

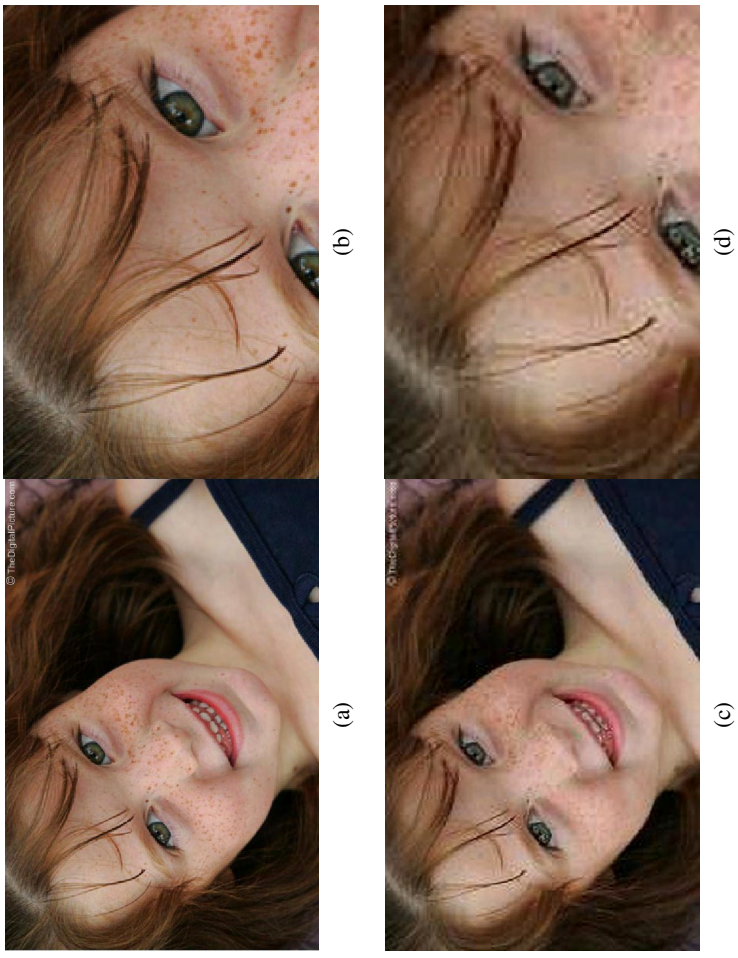
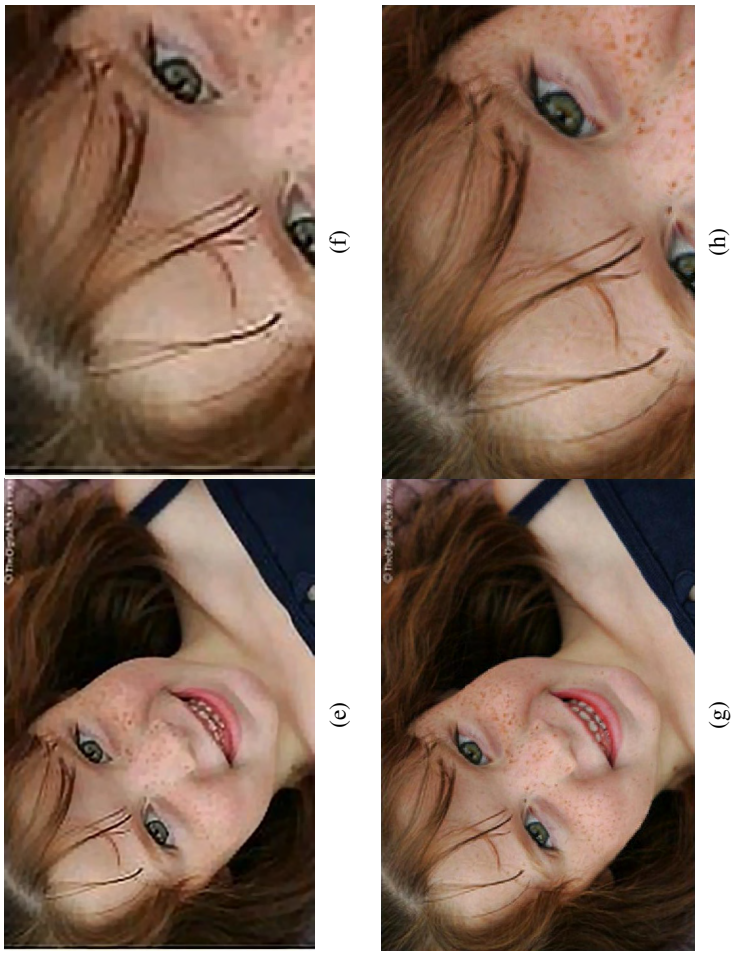


Figure 6.8: (a)original image(c)super resolved using Yang et al method (b),(d)zoomed portion of the face of (a),(c)respectively. contd.....



(e) (f) (g) (h)

Figure 6.8: (e)super resolved using Shan et al method(g)super resolved image using directionlets (f),(h)zoomed portion of the face of (e),(g)respectively.

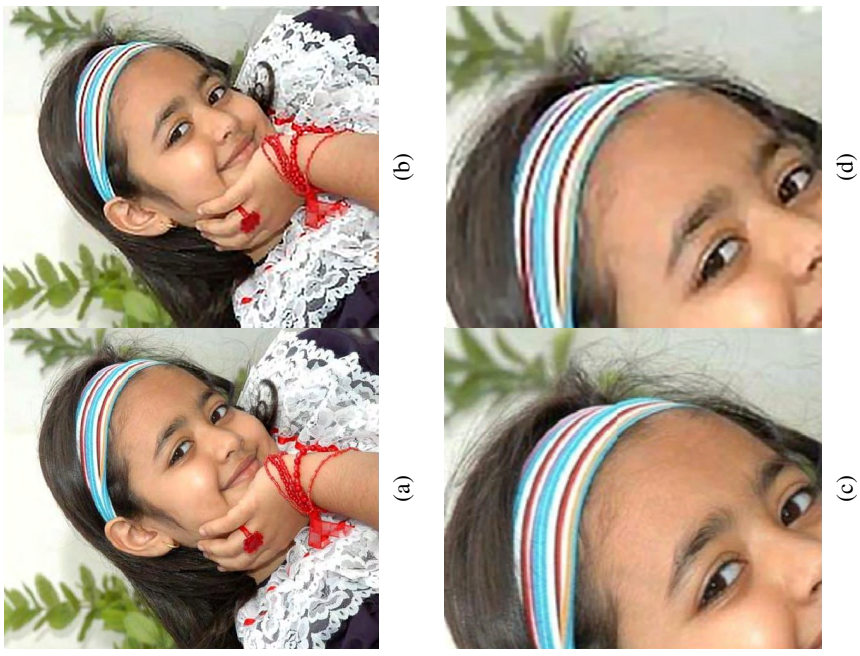


Figure 6.9: (a)original image(c)super resolved using Yang et al method(b),(d)zoomed portion of the face of (a),(c)respectively . contd.....

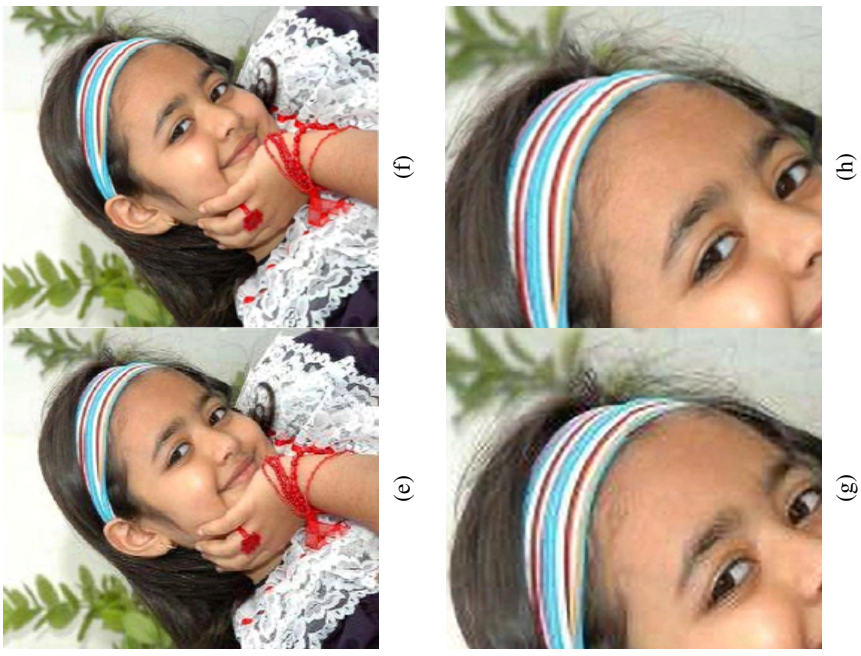


Figure 6.9: (e)super resolved using Shan et al method(g)super resolved image using directionlets (f),(h)zoomed portion of the face of (e),(g)respectively.

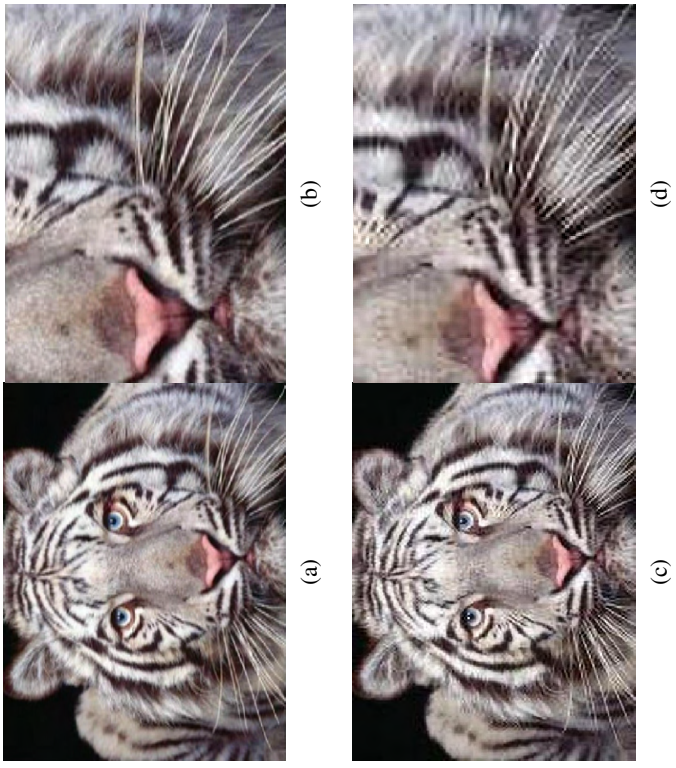
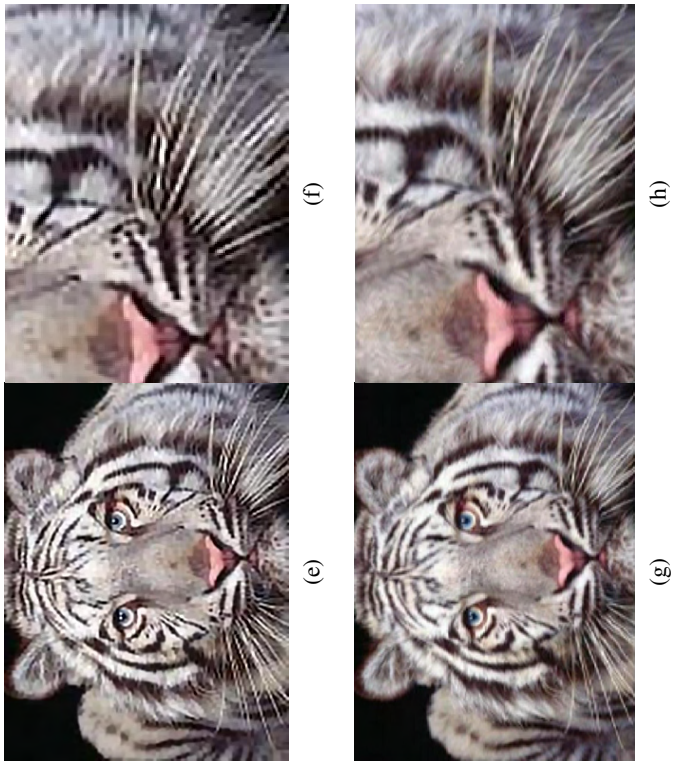


Figure 6.10: (a)original image(c)super resolved using Yang et al method(b),(d)zoomed portion of the face of (a),(c)respectively. contd.....



(e) (f) (g) (h)

Figure 6.10: (e)super resolved using Shan et al method(g)super resolved image using directionlets (f),(h)zoomed portion of the face of (e),(g)respectively

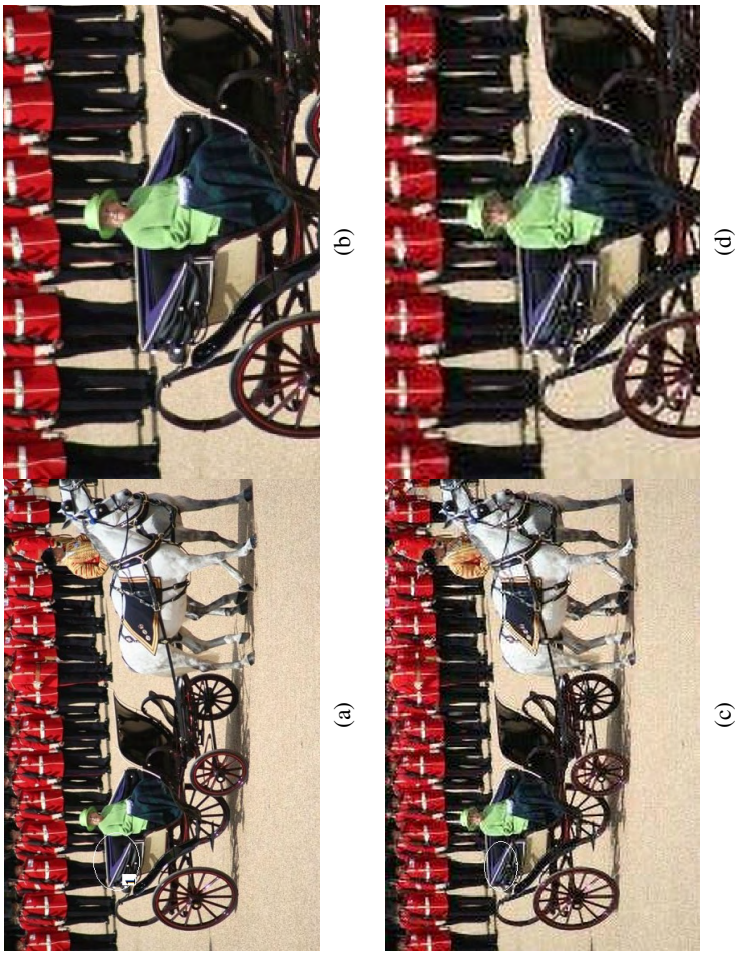


Figure 6.11: (a)original image(c)super resolved using Yang et al method(b),(d)zoomed portion of the face of (a),(c)respectively. contd.....

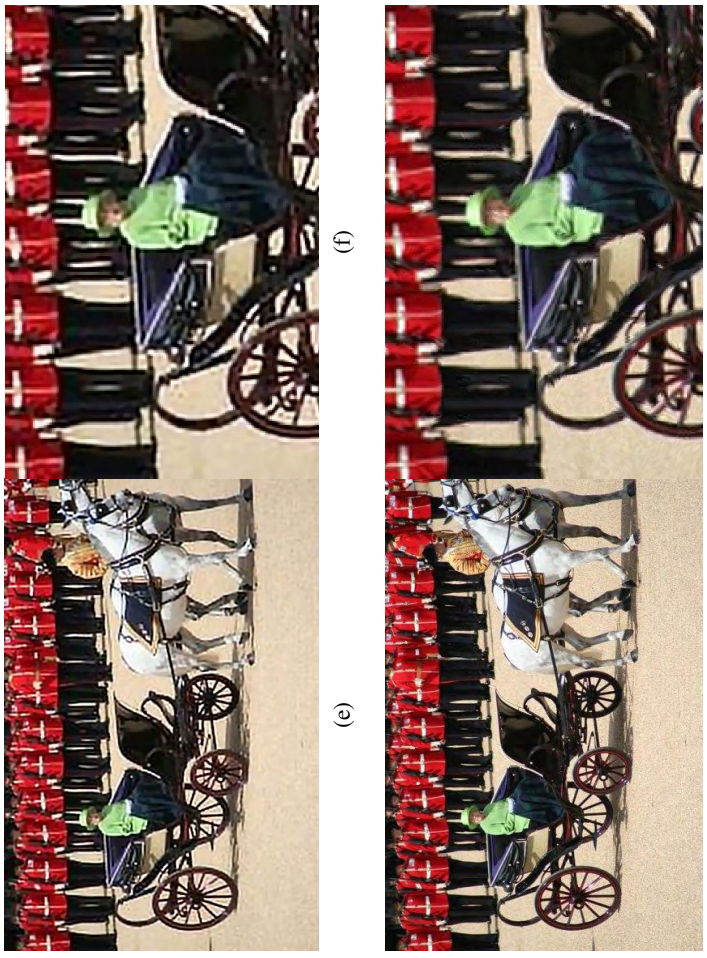


Figure 6.11: (e)super resolved using Shan et al method(g)super resolved image using directionlets (f),(h)zoomed portion of the face of (e),(g)respectively

The third low resolution image used is shown in 6.7(f). It is of size 200×150. Figures 6.10(a), (c), (e), (g) show original image, super resolved image using Yang et al method, super resolved using Shan et al method and directionlet based super resolved image (400×300). Figures 6.10(b), (d), (f), (h) show zoomed portions of (a), (c), (e), (g) respectively. The ringing effects present in the mustache of tiger in the super resolved images using Yang et al method and Shan et al method is almost removed in directionlet based method.

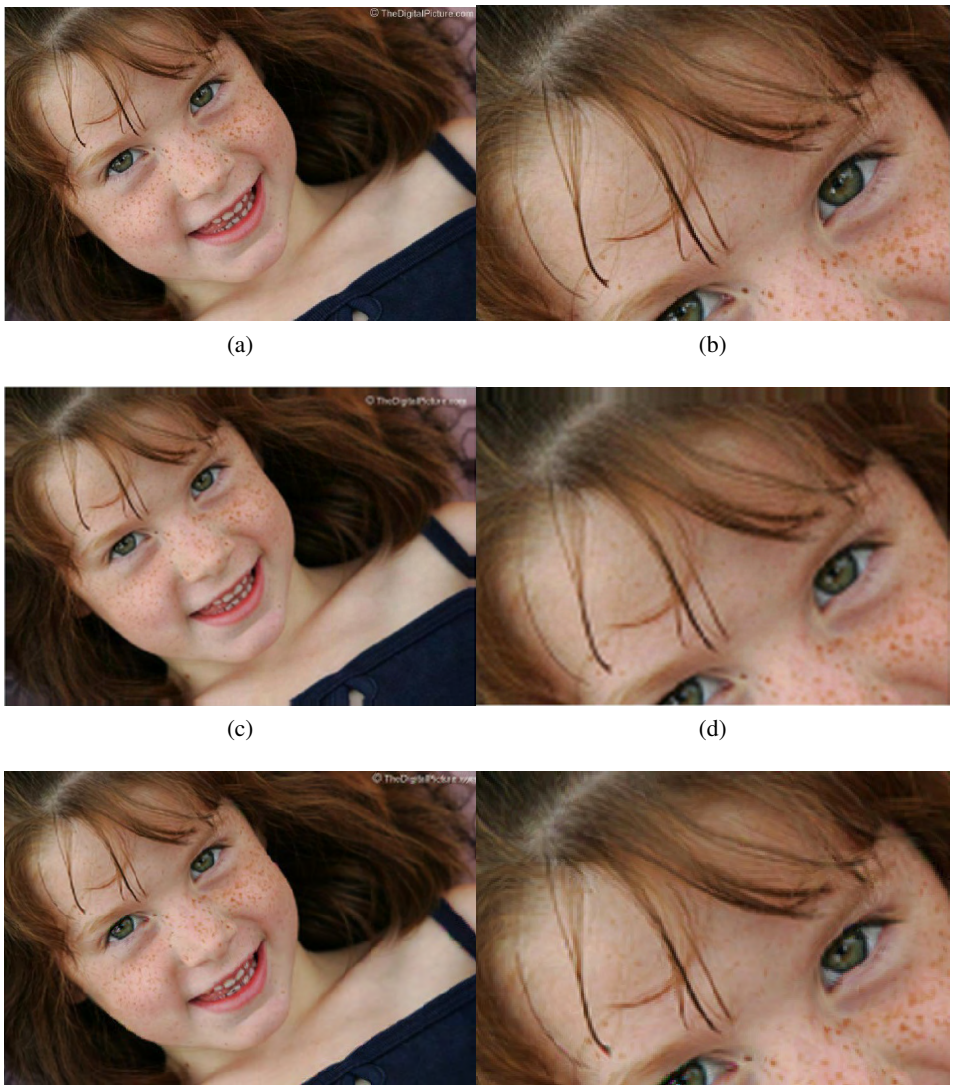
In Figures 6.11(a), (c), (e), (g) show original image, super resolved image using Yang et al sparse method, Shan et al method and new super resolved image of low resolution image in Figure 6.7(h). Low resolution image is of size 196x292 and it is super resolved to the size of 392x584. The block effect in the marked area of super resolved images using Yang et al method and Shan et al method is almost removed in the new super resolved image.

6.3.2 Super resolution on R, G, B components

Methodology

To form the training set, each high and low resolution color component is divided into patches and directionlet transform is applied on these patches. Thus there are three groups of directionlet coefficients corresponding to three color components. The training set contains three groups of directionlet coefficients. Each group contains directionlet coefficients of low and high resolution color component patches. In MATLAB, the format of input color image is in RGB. The high resolution color components R, G, B are learned from the training set individually.

To super resolve an input low resolution color image, it's three high resolution components corresponding to different colors are learned separately from the training set. Each color component is super resolved separately.



(a) (b)

(c) (d)

(e) (f)

Figure 6.12: (a)Original image(c)image super resolved by super resolving luminance components only(e)image super resolved by super resolving R, G, B components (b),(d),(f)zoomed portion of the face of (a),(c),(e)respectively

Implementation

The same high resolution images which are used in the first method are used as the training set images. The low resolution image of girl1 is super resolved to its double size.

Results and discussions

Figure 6.12(a), (c), (e) show the original image, the SR image by super resolving luminance the components only and SR image by super resolving all the three color components separately and (b), (d), (f) are the zoomed portions of (a), (c), (e) respectively. The super resolved image in Figure 6.12(e) has almost same subjective quality compared to the image in Figure 6.12(c) , which is obtained by super resolving the luminance component only. Also SNR of Figure 6.12(e) is 43.9447dB while that of Figure 6.12(c) is 43.5548dB. But the time taken for the second method is very large (three times) compared with the first method.

6.4 Super resolution of LR color images to 4 times its orginal size (magnification factor 4)

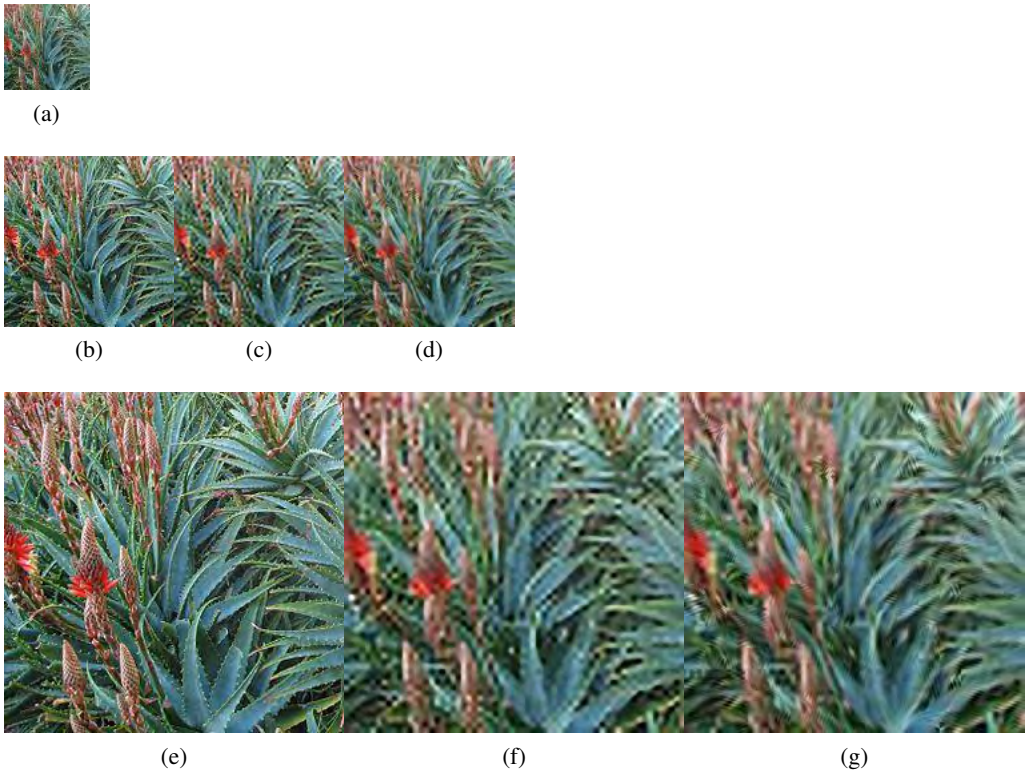
The LR color image of size MxM is super resolved to an image of size 4Mx4M using new directionlet method. It is compared with the super resolved image obtained using Yang et al's sparse method.

As in the case of grey images, the zooming factor 4 is obtained by iterating the critically sampled directionlet method two times with output image of first iteration as the input low resolution image of the second iteration. SNR values obtained for different methods are shown in table 6.2.

From table 6.2, SNR values obtained for the image Lena for zooming factor 2 and 4 are 22.5731dB and 13.8592dB for directionlet method, 22.0944dB and 13.8316dB for Yang's method respectively. So is the case with image bush. From SNR values, it is clear that directionlet method gives better result.

Table 6.2: SNR values for color images with zooming factor 2 and 4

Method	SNR in DB			
	for zooming factor 2		for zooming factor 4	
	Lena	Bush	Lena	bush
Yang et al's sparse method	22.0944	9.6565	13.8316	6.4616
Directionlet method	22.5731	10.6168	13.8592	6.7398



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

Figure 6.13: (a)low resolution image(64x64) (b)original image(128x128) (e)original image(256x256) (c),(d)2 times SR images (128x128) using Yang et al method and directionlet method(f),(g) 4 times SR images (256x256) using Yang et al method and new directionlet method respectively.

(a)

(b) (c) (d)

(e) (f) (g)

Figure 6.14: (a)low resolution image(64x64)(b)original image(128x128)(e)original image(256x256) (c),(d)2 times SR images (128x128) using Yang et al method and directionlet method (f),(g) 4 times SR images (256x256) using Yang et al method and new directionlet method respectively.

Figures6.13(a)low resolution image(64x64)(b)original image (128x128)(e)original image(256x256) (c), (d) 2 times SR images (128x128) using Yang et al method and directionlet method(f), (g) 4 times SR images (256x256) using Yang et al method and new directionlet method. Blocking effect along the edges of the leaves are noticed in Yang et al method. Triangular protrusions along the leaf are missing in directionlet method and but it is sharper than the image by the other method.

Figures 6.14(a) low resolution image(64x64) (b)original image(128x128) (e)original image(256x256) (c),(d)2 times SR images (128x128) using Yang et al method and directionlet method(f),(g) 4 times SR images (256x256) using Yang et al method and new directionlet method. Here ringing effect present in the edge of hat of Yang et al method is less in new directionlet method.

6.5 conclusion

In this chapter directionlet based super resolution method is extended to color images. Two super resolution methods are implemented here. In the first method, RGB format is converted to YC_bC_r format and super resolution method is applied to the luminance component, Y alone. In the second method each R, G, B components are super resolved separately. The two methods (YCbCr method) give almost same result both in numerical and visual quality, but the second method needs more time compared with first method. They are compared with the Yang's method which is proved to be the best among existing methods. It is seen that directionlet based super resolved method outperforms the existing super resolution methods, in the case of color images also.

Chapter 7

Single Image Super-Resolution using directionlets on Noisy images

---

As in the case of color images, the single image super resolution method using directionlets can be extended to noisy low resolution images also. Since the learning based method involves mapping between the low resolution image and the training set, pre-processing using bilateral filter is done here. Advantage is that the same training set which has been used in grey image super-resolution can be used here. The new super resolution method is found to be effective for all types noises used.

7.1 Introduction

The super-resolution methods described in the previous chapters considered low resolution images free of noise. But this is not the actual case. An image is often corrupted by noise in its acquisition, recording and transmission. The performance of sensors is affected by a variety of factors, such as environmental conditions during image acquisition and by the quality of the sensing elements themselves. For instance, while acquiring images with a CCD camera, light levels and sensor temperatures are major factors affecting the amount of noise in the resulting image. Images are corrupted during transmission principally due to interference in the channel used for transmission. For example, an image transmitted using a wireless network might be corrupted as a result of lightening or other atmospheric disturbances [76]. These random distortions make it difficult to perform the required picture processing.

With the exception of spatially periodic noise, noise is independent of spatial coordinates and is uncorrelated with respect to the image itself. That is, there is no correlation between pixel values and values of noise components. Image denoising methods are used to remove the additive noise while retaining as much as possible the important signal features.

The effects of directionlet-based single image super resolution method on noisy images is presented here. The training set contains directionlet transform coefficients of high resolution images and their low resolution images. The idea used is that the noisy input low resolution image is decomposed into different frequency bands using directionlet transform. These coefficients are compared with the corresponding training set coefficients to select the most similar ones. The higher bands of high resolution images are learned from the training set. The low bands are obtained from the low resolution noisy image. The inverse directionlet transform of all these bands gives the super resolved image of the noisy image.

7.2 Overview: Noises

Unlike analog signals which are more prone to noises, discreteness nature of digital signals offers some built-in tolerance to noise . The three most common types of random noise likely to be encountered in images are: Gaussian noise, Salt and Pepper noise, and Speckle noise.

Gaussian noise

Because of its mathematical traceability in both the spatial and frequency domains, Gaussian noise models are used frequently in practice. In fact, this traceability is so convenient that it often results in Gaussian models being used in situations where they are marginally applicable at best. The Probability Distribution Function (PDF) of a Gaussian random variable z, is given by

$$P(z) = \frac{1}{\sqrt{2\Pi\sigma}} e^{\frac{-z-\mu^2}{2\sigma^2}} \tag{7.1}$$

where z represents gray level, μ is the mean or average value of z, and σ is its standard deviation. The standard deviation squared, σ^2, is called the variance of z.

Salt and Pepper Noise(Impulse noise)

The PDF of impulse noise is given by equation

$$\begin{aligned} P(z) &= P_a; for, z = a \\ &= P_b; for, z = b \\ &= 0; otherwise \end{aligned} \tag{7.2}$$

If b>a, gray level b will appear as a light dot in the image and level a will appear like a dark dot. Impulse noise will resemble salt and pepper granules randomly distributed over the image. For this reason bipolar noise is also called Salt and Pepper noise .

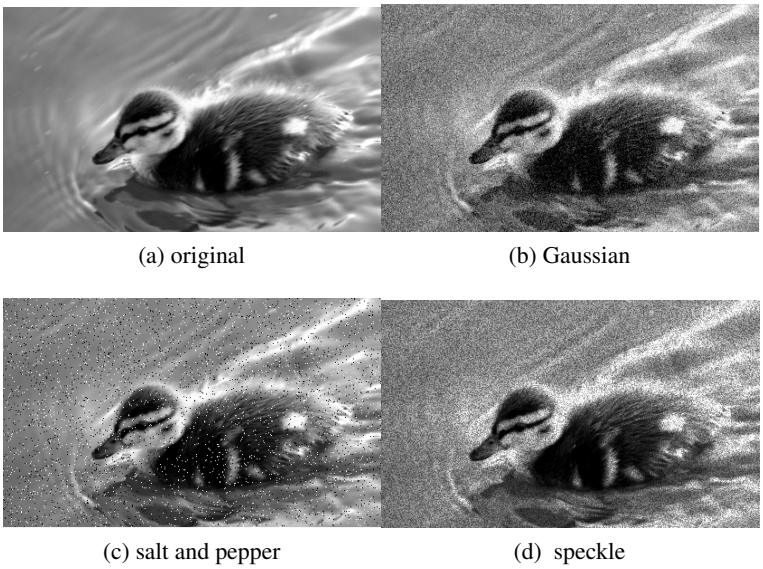
"Spike" or impulse noise that drives the intensity values of random pixels to either their maximum or minimum values. It is simple to handle. They occur when

pixel's intensities are either driven to minimum or maximum values. An effective noise reduction method for this type of noise involves the usage of a median filter, morphological filter or harmonic mean filter [3]. Salt and pepper noise creeps into images in situations where quick transients, such as faulty switching, take place.

Speckle Noise

It is a form of multiplicative noise in which the intensity values of the pixels in the image are multiplied by random values.

Figure 7.1 shows images with these noises.



(a) original (b) Gaussian

(c) salt and pepper (d) speckle

Figure 7.1: Images with different noises

7.3 Denoising Methods

Many denoising methods have been developed over the years. Multi resolution analysis has been proven to be an important tool for eliminating noise in signals. Using multi resolution analysis, it is possible to distinguish between noise and

image information better at one resolution level than another. The problem of image de-noising can be summarized as follows. Let A(i,j) be the noise-free image and B(i,j) the image corrupted with independent Gaussian noise Z(i,j), then

$$B(i,j) = A(i,j) + Z(i,j) \tag{7.3}$$

The problem is to estimate the desired signal as accurately as possible according to some criteria. There exist many denoising algorithms which uses averaging filter, median filter etc. Among many denoising methods, wavelet thresholding is one of the most popular approaches.

7.3.1 Wavelet Methods

In wavelet thresholding, a signal is decomposed into its approximation (low-frequency) and detail (high-frequency) sub-bands. Since most of the image information is concentrated in a few large coefficients, the detail sub-bands are processed with hard or soft thresholding operations. In the wavelet domain, if an orthogonal wavelet transform is used, the problem can be formulated as

$$Y(i,j) = W(i,j) + N(i,j) \tag{7.4}$$

where Y(i,j) is noisy wavelet coefficient, W(i,j) is true coefficient and N(i,j) noise [93]. De-noising of natural images corrupted by Gaussian noise using wavelet techniques is very effective because of its ability to capture the energy of a signal in few energy transform values. The wavelet transform yields a large number of small coefficients and a small number of large coefficients. Simple de-noising algorithms that use the wavelet transform consist of three steps.

1. Calculate the wavelet transform of the noisy signal.

2. Modify the noisy wavelet transform coefficients according to some rule. Coefficients that are supposed to be affected by noise are replaced by zero or by another suitable value, and also the other coefficients may be modified. This process is called thresholding.

3. Compute the inverse transform using the modified coefficients to obtain the noise free image.

.

7.3.2 Bilateral filter

The bilateral filter takes a weighted sum of pixels in a local neighborhood; the weights depend on both the spatial distance and the intensity distance. In this way, edges are preserved well while noise is averaged out. The bilateral filter is a nonlinear filter that does spatial averaging without smoothing edges; it has shown to be an effective image denoising technique [106], [56].

7.4 Single image super resolution in noisy images

Single image super resolution method described in previous chapters can easily be extended to noisy images. The noisy low resolution image is super resolved to a noiseless high resolution image. The method explained here is based on the paper [19] and single image super resolution method using directionlet transform.

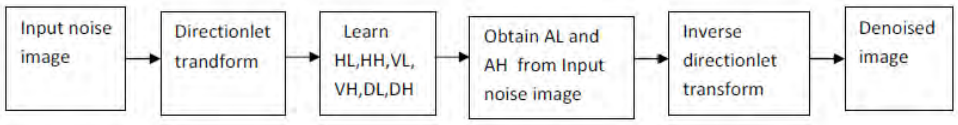


Figure 7.2: Block diagram of noisy image super resolution using directionlet transform

To get the noise free high resolution image of an input noisy low resolution image, it's high resolution frequency bands are learned from a training set. The training set contains directionlet coefficients of patches of noise free high resolution images and their low resolution images. To obtain the low frequency bands, up sample the low resolution noisy image patches and apply the directionlet transform on it. The bands AL and AH obtained by this process are used as the AL and AH

of the high resolution image. The Figure 7.2 shows block diagram of noisy image super resolution using directionlet transform.

Implementation and Discussions

Implementation

The same training set used in the earlier implementations with grey level images is used here. The training set contains directionlet transform coefficients corresponding to noiseless low resolution patches and high resolution patches. Experiments are conducted with different types of noises. Three types of noises are used : Gaussian, Speckle, Salt and Pepper. The low resolution noisy images are obtained by adding noises in low resolution images. A magnification factor of q=2 is used.

Discussions

The results are shown in Figure7.3. Different types of noises with standard deviation 0.1 are added to the low resolution image to simulate different noisy images. The noisy input image itself is used for obtaining AL and AH components and matching during the learning process. Here the new method is compared with Sapan et al method because this method uses wavelet denoising method. Figures7.3(a), (d), (g) are low resolution images with noises Gaussian, Speckle and Salt and Pepper (with standard deviation σ=0.1). Figures7.3(b), (e), (h) are super resolved images using Sapan et al method and Figures7.3(c), (f), (i) are super resolved images of Figures7.3 (a),(d),(g) using new directionlet method, without using any preprocessing for noise removal. From results it is observed that smooth regions of super resolved image is still noisy and when compared with the existing method, quality of super resolved image is not good.

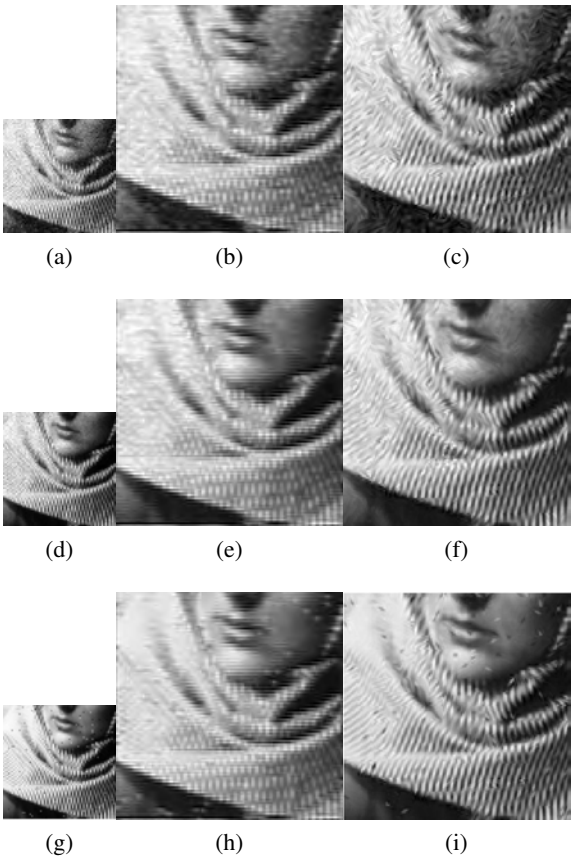


Figure 7.3: (a),(d),(g)Low resolution images with gaussian,speckle and salt and pepper noises σ=0.1 (b),(e), (h)Super resolved images using Sapan et al method(c), (f), (i)super resolved images of (a), (d), (g) respectively.

7.5 Single image super resolution in noisy images using Bilateral filter

Since super resolution involves mapping between input image patches and training set patches, the noisy input image is preprocessed before the process of super resolution. The above mentioned method is modified by applying a bilateral filter on low resolution input image. It's block diagram is shown in Figure7.4. The

input noisy image is pre processed using bilateral filter. The preprocessed noisy image is divided into patches and the Directionlet transform is applied on input noisy image patches and the high frequency bands HL, HH, VL, VH, DL, DH are learned from the training set. The low frequency bands are obtained from the input image. The inverse directionlet transform gives the almost noise free high resolution image.

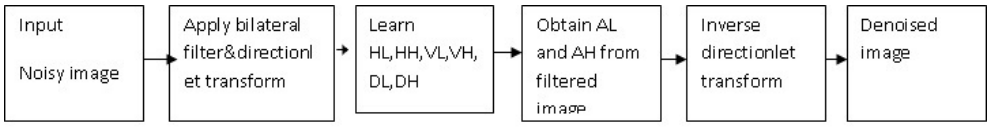


Figure 7.4: Block diagram of noisy image super resolution using directionlet transform with Bilateral filter for preprocessing

Implementation and Discussions

The experiments are done on images with Gaussian noise, Speckle noise and Salt and Pepper noise of different values of standard deviation(variance). The same training set used with grey images is used here.

7.5.1 Single image super resolution on images with Gaussian noise

Table 7.1 shows the SNR values obtained for super resolving different low resolution images with Gaussian noise of standard deviation σ=0.1, 0.2, 0.3. From the table it is clear that the new directionlet method gives better result than Sapan et al wavelet method. For example SNR value for Butterfly is 23.92dB while it is only 19.55dB for Sapan et al method. It is also clear that as noise increases the quality of super resolved image decreases.

Figure7.5, Figure7.6 show the results obtained with Gaussian noise. The results are compared with super resolved images using Sapan et al method. In the new super resolution method the missing high frequency bands are learned from the training set which are entirely free of noise. The low frequencies are obtained from

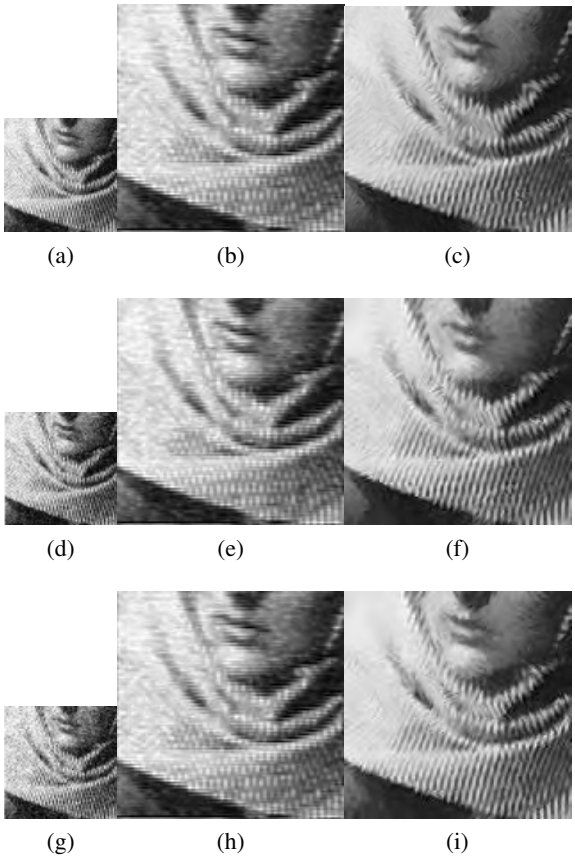
Table 7.1: SNR values for different low resolution images with Gaussian noise

images	Methods	SNR in dB		
		σ=0.1	$\sigma = 0.2$	σ=0.3
Butterfly	using Sapan et al method	19.55	19.5761	19.5683
	Directionlet Method	23.9251	22.549	21.031
Barbara	using Sapan et al method	14.5125	14.4897	14.4897
	Directionlet Method	19.401	17.9563	16.76

the directionlet transform of interpolated version of low resolution image. Thus the noise present in the low resolution image is almost removed by the averaging process of low pass filters in low frequency bands.

Figures7.5(a), (d),(g) show low resolution images of 'Barbara' with Gaussian noise of standard deviation σ=0.1, 0.2, 0.3. Figures7.5(b), (e), (h) show super resolved images using Sapan et al method and Figures7.5(c), (f), (i) show super resolved images using new directionlet method.

Figures7.6(a), (d),(g) show low resolution images of 'butterfly' with Gaussian noise σ=0.1, 0.2, 0.3. Figures7.6(b), (e), (h) show super resolved images using Sapan et al method and Figures7.6(c), (f), (i) show super resolved images using new method. It is clear that directionlet super resolved images are better than the images super resolved with existing wavelet transform based method.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 7.5: (a),(d),(g)Low resolution images with Gaussian noise (σ=0.1, 0.2, 0.3)(b),(e),(h)super resolved images using Sapan et al wavelet method (c),(f),(i) super resolved images using directionlet method

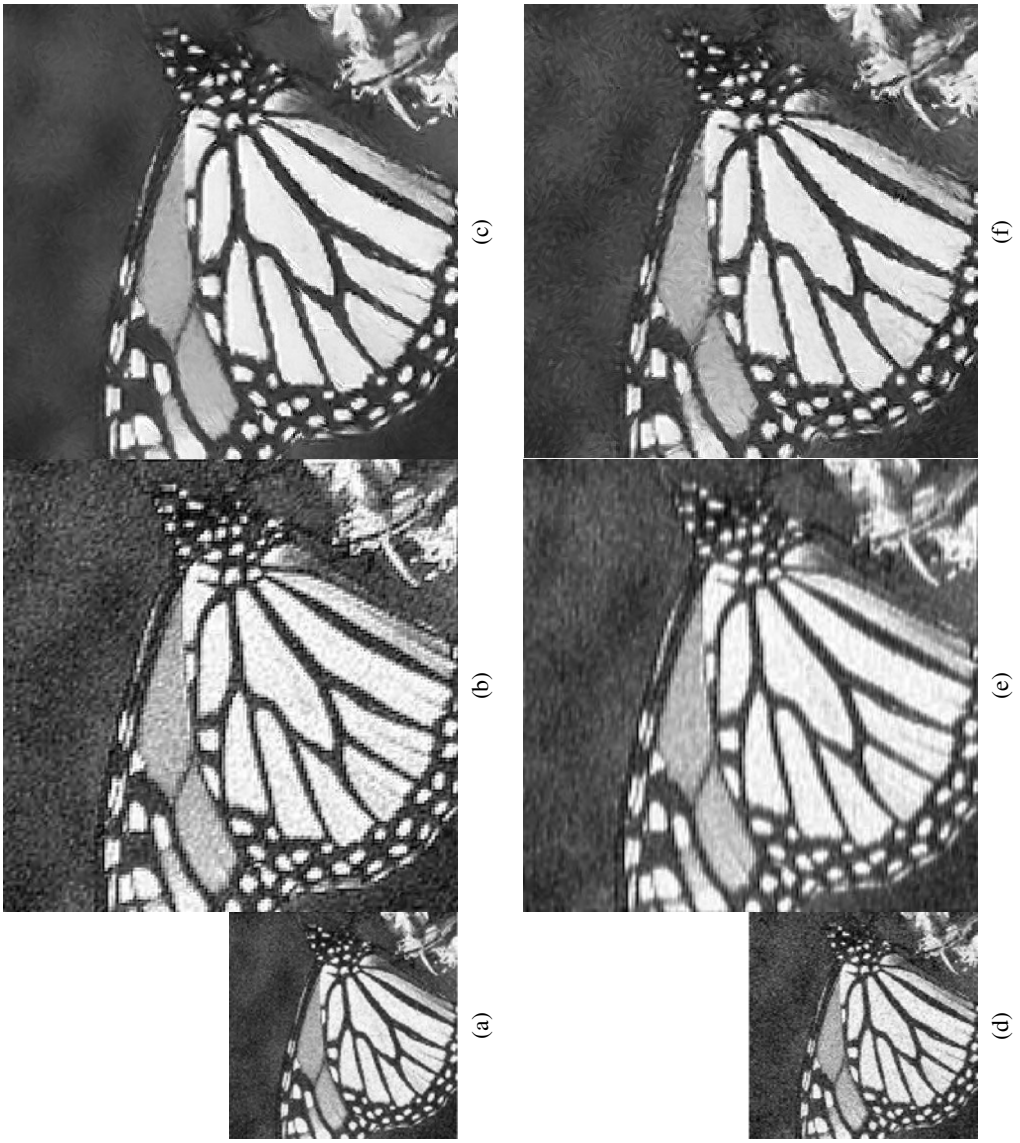
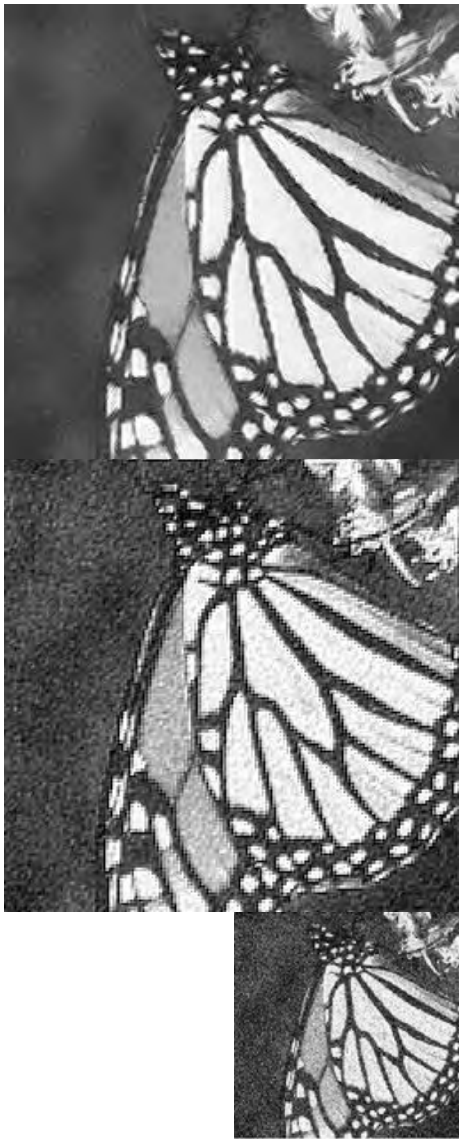


Figure 7.6: contd.....



(g) (h) (i)

Figure 7.6: (a),(d),(g)Low resolution images with Gaussian noise (σ=0.1, 0.2, 0.3)(b),(e),(h)super resolved images using Sapan et al wavelet method (c),(f),(i) super resolved images using directionlet method

7.5.2 Single image super resolution on images with Speckle noise

Table7.2 shows the SNR values obtained for super resolving different low resolution images with Speckle noise of standard deviation σ=0.1, 0.2, 0.3. The SNR value of Butterfly for σ=0.1 is 24.59dB for new directionlet method while it is 20.4610dB for existing Sapan et al method.

Table 7.2: SNR values for different low resolution images with speckle noise

Images	Methods	SNR in dB		
		σ=0.1	$\sigma = 0.2$	σ=0.3
Butterfly	using Sapan et al method	20.4610	20.05	19.6978
	Directionlet Method	24.59	24.0533	23.46
barbara	using Sapan et al method	14.8994	14.6246	14.4224
	Directionlet Method	20.31	19.75	19.37

Figures7.7(a), (d), (g) show low resolution images of 'barbara' with Speckle noise σ=0.1, 0.2, 0.3. Figures7.7(b), (e), (h) show images with Sapan et al method and Figures7.7(c), (f), (i) show super resolved images with new directionlet method.

Figures7.8(a), (d), (g) show low resolution images of butterfly with Speckle noise σ=0.1, 0.2, 0.3. Figures7.8(b), (e), (h) show images with Sapan et al method and Figures7.8(c), (f), (i) show super resolved images with new directionlet method. It is clear that super resolved images using new method are better than images using existing super resolution method. The directionlet method is better for removing speckle noise also.

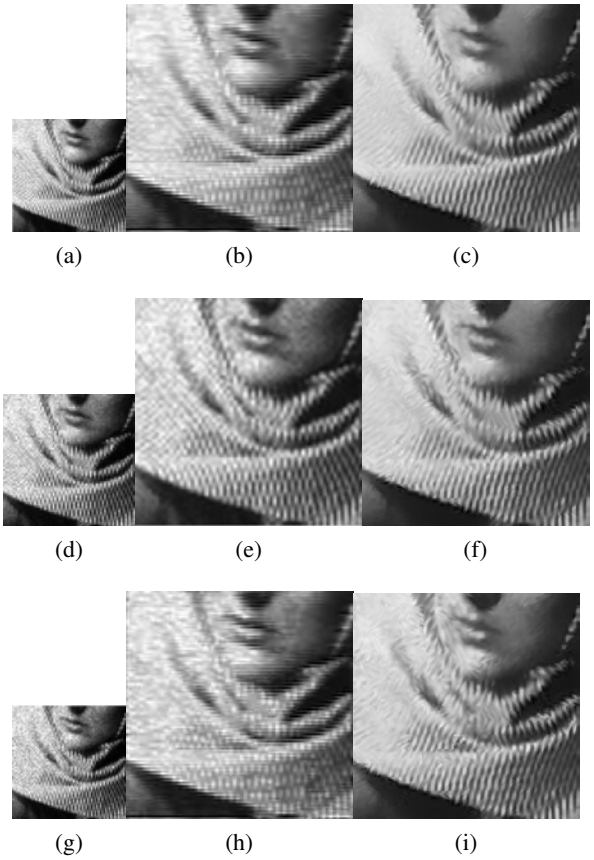


Figure 7.7: (a),(d),(g)Low resolution images with Speckle noise (σ=0.1, 0.2, 0.3)(b),(e),(h)super resolved images using Sapan et al wavelet method (c),(f),(i) super resolved images using directionlet method

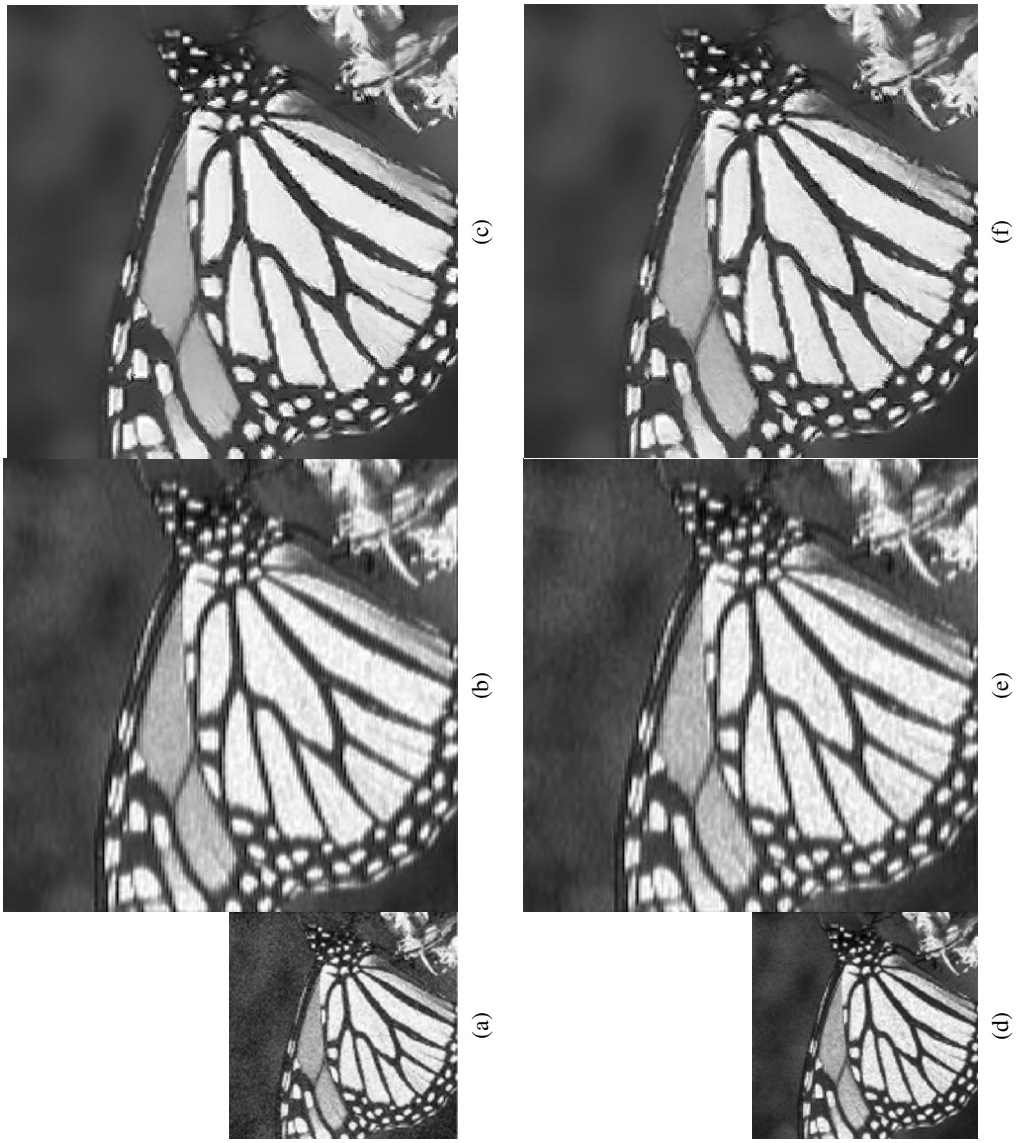


Figure 7.8: contd.....

(g) (h) (i)

Figure 7.8: (a),(d),(g)Low resolution images with Gaussian noise (σ=0.1, 0.2, 0.3)(b),(e),(h)super resolved images using Sapan et al wavelet method (c),(f),(i) super resolved images using directionlet method

7.5.3 Single image super resolution on images with Salt and Pepper noise

Table 7.3 shows the SNR values obtained for super resolving different low resolution images with Salt and Pepper noise of standard deviation σ=0.1, 0.2, 0.3. The SNR values for Butterfly is 24.39dB while it is 20.3163dB for Sapan et al method for σ=0.1.

Table 7.3: SNR values for different low resolution images with Salt and Pepper noise

Images	Methods	SNR in DB		
		σ=0.1	$\sigma = 0.2$	σ=0.3
Butterfly	using Sapan et al method	20.3163	19.7292	19.17
	Directionlet Method	24.39	23.66	22.92
barbara	using Sapan et al method	14.8965	14.654	14.43
	Directionlet Method	20.03	19.51	18.82

Figures7.9(a), (d), (g) show low resolution images of 'Barbara' with Salt and Pepper noise σ=0.1, 0.2, 0.3. Figures7.9(b), (e), (h) show images with Sapan et al method and Figures7.9(c), (f), (i) show super resolved images with new directionlet method. It is clear that directionlet based super resolved image is better than images using Sapan et al super resolution method.

Figures7.10(a), (d), (g) show low resolution images of 'butterfly' with Salt and Pepper noise σ=0.1, 0.2, 0.3 etc. Figures7.10(b), (e), (h) show images with Sapan et al method and Figures7.10(c), (f), (i) show super resolved images with new method. It is clear that directionlet based super resolved image is better than images using Sapan et al super resolution method.

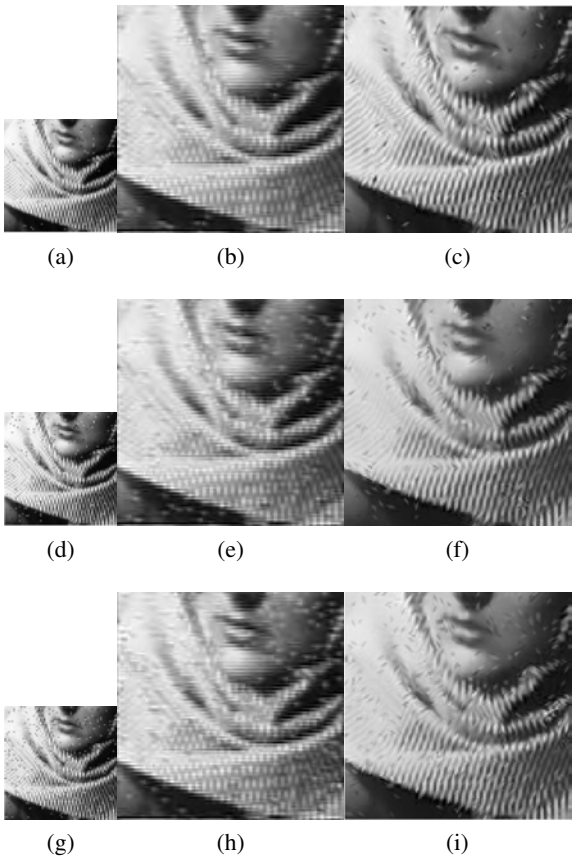


Figure 7.9: (a),(d),(g)Low resolution images with Salt and Pepper noise (σ=0.1, 0.2, 0.3)(b),(e),(h)super resolved images using Sapan et al wavelet method (c),(f),(i) super resolved images using directionlet method

Figure 7.10: contd....

(g) (h) (i)

Figure 7.10: (a),(d),(g)Low resolution images with Gaussian noise (σ=0.1, 0.2, 0.3)(b),(e),(h)super resolved images using Sapan et al wavelet method (c),(f),(i) super resolved images using directionlet method

7.6 Conclusion

The single image super resolution method with directionlet transform is extended to noisy images. The low resolution noisy images are filtered using bilateral filter before undergoing super resolution process. The same training set used with grey image super resolution is used here. Experiments are conducted on different images with different types of noises like Gaussian, Speckle, Salt and Pepper with different variances. Results are compared with the existing wavelet transform method and it is observed that the new directionlet method is able to remove noise contents to a greater extent.

Chapter 8

Conclusion

A brief summary of the research work conducted and the important conclusions thereon are highlighted in this chapter. The scope for further work in this field as an extension of the present study has also been discussed.

8.1 Thesis summary and Conclusions

In this thesis, problem of single image super resolution method is addressed. Resolution beyond the limit of image capturing device is achieved by using directionlet transform to extract high frequency features from the high resolution images in the training set. The presented method is different from other conventional super resolution methods in the way that it is adaptive to local directional variations present in images.

Different methods presented in this thesis are summarized here. Initially a learning based super resolution method using learned wavelets is presented. It is obtained by modifying the method proposed by Jiji et al using patch based approach. Advantage of the new patch based wavelet method is that low resolution images of any size can be super resolved using a single training set. Artifacts are also reduced in the super resolved images using this wavelet method. But this method including traditional methods fail to remove artifacts like ringing effects and aliasing.

Next, a novel learning based super resolution method using directionlet transform is introduced. Advantage of this method is that for each patch, direction of transform is selected according to the information present in it. This method out performs the standard interpolation and wavelet methods. Artifacts like aliasing and ringing effect are also reduced by this method. This new method needs more computations and hence need more computation time. To speed up this process lifting based directionlet transform is used instead of conventional convolution based directionlet transform and this lifting based method is faster than the convolution based directionlet method.

A study is done to analyze the effect of different wavelets on the directionlet based super resolution method. Results show that though the wavelet basis 'rbior1.5' gives better super resolved image compared to other wavelets, it takes three or four times more time compared to 'db4' or 'bior3.3'

The new single image super resolution method using directionlet transform is extended to color and noisy image also. It is found that the new method based on directionlet is effective in the case of colour and noisy images.

Thus a new directionally adaptive single image super resolution method is developed to super resolve grey ,colour and noisy images. It super resolves images to their double size. It can be also used for super resolving to 4 times the size of the original image.

8.2 Suggestions for Future Work

These methods are computationally too expensive for real time applications. One should think of ways to speed up these algorithms. Training set needs large amount of memory and compression techniques can be used to reduce the size of memory. Here low resolution images which are free from blurring are used. The new super resolution method presented here can easily be extended to blurred observations also. Here five sets of directions are selected. By extending to more directions the quality of super resolved image can be increased further. Magnification factor two and four are used in the presented work. This can be extended to higher values.

List of Publications

i. A.P Reji , Dr.Tessamma Thomas, *A Learning Based Single Image Super Resolution Method Using Directionlets* , International Conference on Advances in Computer Engineering,ACE, Bangalore,June 21-22,(2010), 1048–1053.

ii. A.P Reji , Dr.Tessamma Thomas, *Single Frame Image Super Resolution Using Learned Directionlets*, International Journal of Artificial Intelligence and Applications, AIRCC Publishers, Volume1,Number 4,(2010),pp 17-27.

iii. A.P Reji , Dr.Tessamma Thomas, *A Study Using Different Wavelet Basis on Single Image Super resolution with Learned Directionlets*, International Journal on Recent Trends in Engineering and Technology, ACEEE, USA (a division of IDES), volume 4 (2010) 82-85.

iv. A.P Reji , Dr.Tessamma Thomas, *Directionally Adaptive Single Image Super Resolution* , International journal of Innovative Computing and Applications,Inderscience Publishers,3(3)(2010), 117-125.

v. A.P Reji , Dr.Tessamma Thomas, *Single Image Super Resolution for Grey and Color Images using Lifting based Directionlets*, SIAM Journal on Imaging Sciences.(Communicated)

vi. A.P Reji , Dr.Tessamma Thomas, *Single Image Super Resolution for noisy images*, Computer Vision,IET.(Communicated)

Bibliography

[1] http://www.the-digital-picture.com/pictures.

[2] Aaron.Hertzmann, Charles.E. Jacobs, Nuria Oliver, Brian Curless, and David .H. Salesin. "Image analologies". *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 233–242, 2001.

[3] A.Bovik. *Handbook of Image and Video Processing*. New York, Academic, 2000.

[4] A.G.Haus and J.E Cullinan. "Screen film processing systems for medical radiography: a historical review". *RadioGraphics,The Journal of Continuing medcal education in radiology*, 9:1203–1224, 1989.

[5] Andrew.J.Patti, M.Ibrahim.Sezan, and A.Murat.Tekalp. "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time". *IEEE Transactions on Image Processing*, 6:1064–76, 1997.

[6] Antoni.Buades, Bartomeu.Coll, and Jean.Michel.More. "Self-similarity-based image denoising". *Communications of the ACM*, 68:157–168, 2011.

[7] Assaf.Zomet, Alex.Rav-Acha, and Shmuel.Peleg. "Robust super-resolution". *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,*, 1:645 –650, 2001.

[8] Assaf.Zomet and Shmuel.Peleg. "Efficient super-resolution and applications to mosaics". *Proceedings. 15th International Conference on Pattern Recognition*, 1:579 – 583, 2000.

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

[10] Bai.J, Wu.J, Lu.S, and Jiao.L. "Lifting-based directionlet transform for image coding". *International Conference on Signal Processing, 26–29 October, ICSP 2008,Beijing, China.*

[11] Baker.S and Kanade.T. "Limits on super-resolution and how to break them". *Proceedings. IEEE Conference on Computer Vision and Pattern Recognition,*, pages 372 – 379, 2000.

[12] Begin.I and Ferrie F.P. "PSF recovery from examples for blind super-resolution". *IEEE International Conference on Image Processing,*, 5:421–424, 2007.

[13] Burt.P.J and Adelson.E.H. "The lapalacian pyramids as a compact image code". *IEEE Tranctions Communications*, 31:532–540, 1983.

[14] C.V.Jiji. "Single-frame image super-resolution",ph.d thesis,department of electrical engineering,iit,bombay. 2007.

[15] C.V.Jiji, M.V.Joshi, and Subhasis Chaudhuri. "Single-frame image super-resolution using learned wavelet coefficients". *International Journal of Imaging Systems and Technology*, 14:105–112, 2004.

[16] C.V.Jiji and Subhasis.Chaudhuri. "Single-frame image super-resolution through contourlet learning". *EURASIP Journal on Applied Signal Processing*, 104, 2006.

[17] C.V.Jiji, Subhasis.Chaudhuri, and Priyam.Chatterjee. "Single frame image super-resolution: should we process locally or globally? ". *Multidimensional Systems and Signal Processing*, 18:123–152, 2007.

[18] Daniel.Glasner, Shai.Bagon, and Michal.Irani. "super-resolution from a single image",. *International conference on computer vision (ICCV),October 2009.*

[19] D. L. Donoho and I. M. Johnstone. "Adapting to unknown smoothness via wavelet shrinkage". *Journal of the Statistical Association*, page 1200–1224, 1995.

[20] Elad.M and Feuer.A. "Restoration of a single super resolution image from several blurred, noisy, and under sampled measured image". *IEEE Transactions on Image Processing*, 6(12):1646–1658, 1997.

[21] Fang.Qiu, Yi.Xu, Ci.Wang, and Yuhong.Yang. " Noisy image super-resolution with sparse mixing estimators". *Image and Signal Processing (CISP), 4th International Congress on,vol.2,pp.1081 - 1085,Oct. 2011*.

[22] Farsiu, Robinson.M.D S Santa.Cruz, Elad.M, and Milanfar.P. "Fast and robust multiframe super resolution". *IEEE Transactions on Image Processing*, 10(13):1327 – 1344, 2004.

[23] Gajjar.P.P and Joshi.M.V. "Single frame super-resolution: A new learning based approach and use of igmrf prior". *ICVGIP '08. Sixth Indian Conference on Computer Vision, Graphics and Image Processing*, 17:636–643, 2008.

[24] Gao.X, Zhang.K, Tao.D, and Li.X. "Image super-resolution with sparse neighbor embedding". *IEEE Transactions on Image Processing*, 21(6):1–5, March,2012.

[25] Giannis.Chantas, Nikolaos.Galatsanos, and Nathan.Woods. "Super resolution based on fast registration and maximum a posteriori reconstruction". *IEEE Transactions on Image Processing*, pages 1821 – 1830, July 2007.

[26] Gilbert.strang and Truong.Nguyen. *Wavelets and Filter banks, Wellesley-Cambridge Press, 2nd Ed. 1997.*

[27] Gotoh.T and Okutomi.M. "Direct super-resolution and registration using raw cfa images". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 600–607, 2004.

[28] Gupta.M.D, Rajaram.S, Petrovic. N., and T.S Huang. "Non-parametric image super-resolution using multiple images". *IEEE International Conference on Image Processing,Genova, Italy*, 2:89–92, 2005.

[29] H.M.Shapiro. "embedded image coding". *IEEE Tranctions of signal processing*, 41(12):3435–3462, 1993.

[30] Hyunwoo.Kim and Ki-Sang.Hong. "Variational approaches to super-resolution with contrast enhancement and anisotropic diffusion". *Journal of Electronic Imaging,*, 12:244–251, 2003.

[31] Ingrid.Daubechies and Wim.Swelden. "factoring wavelet transforms into lifting steps". *Journal of Fourier analysis and applications*, 4(3), 1998.

[32] Ioana.Adam. *"Complex Wavelet Transform: application to denoising"*. PhD thesis, Politehnica University Of Timisoara.

[33] Isabelle.B and Frank.P.Ferrie. "Comparison of super-resolution algorithms using image quality measures". *The 3rd Canadian Conference on Computer and Robot Vision*, 2006.

[34] Jain.V, Murray.J.F, Roth.F.and Turaga.S, Zhigulin.V, Briggman, K.L, Helmstaedter M.N., Denk. W, and Seung . H.S. "Supervised learning of image restoration with convolutional networks". *IEEE 11th International Conference on Computer Vision*, 21:1–8, 2007.

[35] J.H.Conway and N.J.A.Sloane. *Sphere packings, lattices and groups*. pringer-Verlag, 1998.

[36] Jianchao.Yang, John.Wright, Yi.Ma, and Thomas.Huang. "Image super-resolution as sparse representation of raw image patches". *IEEE Conference on Computer Vision and Pattern Recognition,vol.15,no.7*, 2008.

[37] Jianchao.Yang, Wright.J, Huang.T.S, and Yi.Ma. "image super-resolution via sparse representation". *Comm. ACM,vol 19,*, pages 2861–2873, 2010.

[38] Jianping.Qiao and Ju.Liu. "A novel log-wt based super-resolution algorithm ". *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 151–154, 2006.

[39] Jian.Sun, Jian.Sun, Zongben.Xu, and Heung-Yeung Shum. "Image super-resolution using gradient profile prior". *IEEE Conference on Computer Vision and Pattern Recognition,*, pages 1–8, 2008.

[40] Jing.Bai, Jiaji.Wu, Shan.Lu, and Licheng.Jiao. "Lifting-based directionlet transform for image coding". *International Conference on Signal Processing,Beijing, China*, 26–29 October, ICSP 2008, Beijing, China, 2008.

[41] J.M.Combes. Wavelets,time-frequency methods,and phase space (lecture notes on ipti). *New York: Springer-Verlag,1989*, 1989.

[42] Joshi.M.V, Chaudhuri.S, and Panuganti.R. "Learning based super-resolution imaging: Use of zoom as a cue". *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, pages 527–537, 2003.

[43] Kaibing.Zhang, Xinbo.Gao, Xuelong.Li, and Dacheng.Tao. "Partially supervised neighbor embedding for example-based image super-resolution". *IEEE Journal of Selected Topics in Signal Processing,vol.5,issue.2*, April.2011.

[44] Kim.S.P. "Recursive high-resolution reconstruction of blurred multiframe images". *IEEE transactions on image processing*, 2:534–539, 1993.

[45] Kim.S.P, Bose.N.K, and Valenzuala.H.M. "Recursive reconstruction of high resolution image from noisy undersampled multiframes". *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(6):1013–1027, 1990.

[46] K.P.Soman and K.I.Ramachandran. *"Insight into Wavelets:From theory to Practice"*. Prentice hall of India, 2006.

[47] Kwang.In.Kim and Younghee.Kwon. "single-image super-resolution using sparse regression and natural image prior",. *IEEE Trans. Pattern Anal. Mach.Intell., vol. 32, no. 6, pp. 1127–1133, Jun. 2010*.

[48] Kwang.In.Kim and Younghee.Kwon. "Example-based learning for single-image super-resolution". *Pattern Recognition Lecture Notes in Computer Science*, 5096:456–465, 2008,.

[49] L.Guan, S.W. Perry, , and H. Wong. "adaptive image processing: A computational intelligence perspective". 2002.

[50] Liyakathunisa and V.K.Ananthashayana. "Super resolution blind reconstruction of low resolution images using wavelets based fusion". *International Journal of Computer, Information Systems, Science and Engineering*, pages 105–109, 2008.

[51] Lyndsey.Pickup, Stephen.Roberts, and Andrew.Zisserman. "Optimizing and learning for super-resolution". *In Chantler,M.J, Trucco, E. and Fisher, R.B (eds) Proc. British Machine Vision Conf.*, 3:439–448, 2006.

[52] Mahesh.B.Chappalli and N. K. Bose. "Simultaneous noise filtering and super-resolution with second-generation wavelets". *IEEE Signal Processing Letters,vol.12*, 12:772 – 775, 2005.

[53] M.C.Chiang and T.E.Boult. "Imaging-consistent super-resolution". *DARPA Image Understanding Workshop*, 1997.

[54] M.C.Chianga and T.E.Boult. "Efficient super-resolution via image warping". *Image and Vision Computing*, 18(10):761–771, 2000.

[55] M.C.Hong, M.G.Kang, and A.K.Katsaggelos. "An iterative weighted regularized algorithm for improving the resolution of video sequences".

Proceedings International Conference in Image Processing, 2:474–477, 1997.

[56] Michael.Elad. "on the origin of the bilateral filter and ways to improve it". *IEEE transactions on image processing*, 11(10):1141–1151, 2002.

[57] Michael.Elad and Arie.Feuer. "Super-resolution reconstruction of image sequences". *Ieee transactions on pattern analysis and machine intelligence*, 21:817–834, 1999.

[58] Michael.Elad and Arie.Feuer. "Superresolution restoration of an image sequence: Adaptive filtering approach". *IEEE transactions on image processing*, 8:387–395, 1999.

[59] Michael.Elad and Dmitry.Datsenko. "example-based regularization deployed to super-resolution reconstruction of a single image ",. *The Computer Journal Advance Access,April 20,2007*.

[60] Michael.Elad and Yacov.Hel-Or. "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur". *IEEE transactions on image processing*, 10:1187–1193, 2001.

[61] Michael.K.Ng, Jaehoon.Koo, and N.K.Bose. "Constrained total least-squares computations for high- resolution image reconstruction with multisensors". *nternational Journal of Imaging Systems and Technology*, 12(1):35–42, 2002.

[62] Michal.Irani and Shmuel.Peleg. "Super resolution from image sequences ". In *Proc. 10th International Conference on Pattern Recognition*, volume 2, pages 115 – 120, 1990.

[63] Moon.Gi.Kang Min-Cheol.Hong and Aggelos.K. Katsaggelos. "A regularized multichannel restoration approach for globally optimal high resolution video sequence". *SPIE VCIP*, 3024(7):1306–1317, 1997.

[64] Nathan.A.Woods, Nikolas.P.Galatsanos, and Aggelos.K.Katsaggelos. "Stochastic methods for joint registration,restoration, and interpolation of multiple undersampled images". *IEEE transactions on image processing*, 15:201–213, 2006.

[65] Nhat.Nguyen. "An efficient wavelet-based algorithm for image superresolution". *IEEE Trans.On Parallel and Distributed Systems*, 16(4), 2000.

[66] Nhat.Nguyen and Peyman.Milanfar. "A computationally efficient superresolution image reconstruction algorithm". In *IEEE Transactions on Image Processing*, volume 10, pages 573–583, 10-13 Sept. 2001.

[67] Nigel.D.Browning. *"4*D *Electron Microscopy: Imaging in Space and Time "*. Journal of the American Chemical Society, 2010.

[68] N.K.Bose, H.C.Kim, and H.M.Valenzuela. "Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframes". *Proc. IEEE Conf. Acoustics, Speech and Signal Processing, Minneapolis*, 5:269–272, 1993.

[69] Pelletier.S and Cooperstock.J.R. "Preconditioning for edge-preserving image super resolution". *IEEE Transactions on Image Processing*, 41:159–170, Jan.12.

[70] Peyman.Milanfar. "Super-resolution imaging". *CRC Press ,Taylor and Francis Group*.

[71] Peyman.Milanfar, Nhat.Nguyen, and Gene.golub. "Pre conditioners for regularised image super resolution". *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, 6:3249 – 3252, 1999.

[72] Pickup.L.C, Roberts.S.J, and Zisserman.A. "A sample texture prior for image super-resolution advances". *Neural Information Processing Systems*, page 1587–1594, 2003.

[73] P.P.Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[74] Qiang.Wang, Xiaoou.Tang, and Shum.H. "Patch based blind image super resolution". *Tenth IEEE International Conference on Computer Vision,*, 1:709–716, 2005.

[75] Qi.Shan, Zhaorong.Li, Jiaya.Jia, and Chi-Keung.Tang. Fast image/video upsampling. *ACM Transactions on Graphics*, 2008.

[76] Rafael.C.Gonzalez and Richard.E.Woods. *Digital Image Processing*. Pearson Education, Asia, 2002.

[77] Raymond.H.Chan, Tony.F. Chan, Lixin.Shen, and D Zuowei.Shen. "Wavelet algorithms for high-resolution image reconstruction". *SIAM J. SCI. COMPUT*, 24(4).

[78] Richard.R.Schultz and Robert.L. Stevenson. "Extraction of high-resolution frames from video sequences". *IEEE Transactions on Image Processing*, 54(79), 1996.

[79] Richard.R.Schultz, Studerit, and Robert.L.Stevenson. "A bayesian approach to image expansion for improved definition". *IEEE transactions on image processing*, 3:233–242, 1994.

[80] Rob.Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. "Removing camera shake from a single photograph". *ACM Transactions on Graphics (TOG)*, 25.

[81] Roger.L.Claypoole. *"Adaptive Wavelet Transforms via Lifting"*. PhD thesis, Electrical and Computer Engineering Rice University, Houston, Texas, October 1998.

[82] Roshen.Jacob. *"Development of time - frequency techniques for sonar applications"*. PhD thesis, Ph.D thesis,Cochin University of Science and Technology, 2010.

[83] R.Satyabama and S. Annadurai. Image compression using adaptive primal and dual lifting scheme based on mean square error minimization. *European Journal of Scientific Research*, 63(1):99–109, 2011.

[84] Russell.C.Hardie, Kenneth.J.Barnard, and Ernest.E.Armstrong. "Joint map registration and high-resolution image estimation using a sequence of undersampled images". *IEEE transactions on image processing*, 6(12):1624 – 1633, 1997.

[85] Sanjeev.K and Nguyen.T. "Learning the kernel matrix for superresolution". *IEEE 8th Workshop on Multimedia Signal Processing*, pages 441–446, 2006.

[86] Sapan.Naik and Nikunj.Patel. "single image super resolution in spatial and wavelet domain". *The International Journal of Multimedia andIts Applications (IJMA)*, 5:23–32, 2013.

[87] S.H.Rhee and M.G.Kang. "Discrete cosine transform based regularized high-resolution image reconstruction algorithm". *Proceedings 1999 International Conference on Image Processing*, 38:1348–1356, 1999.

[88] Shuaicheng.Liu, Michael.S.Brown, Seon.Joo Kim, and Yu-Wing.Tai. "Colorization for single image super resolution". *Computer Vision – ECCV 2010 Lecture Notes in Computer Science, Volume 6316/2010, 323-336*, 2010.

[89] Shyamsundar.Rajaram, Mithun.Das.Gupta, Nemanja.Petrovic, , and Thomas.S. Huang. "Learning-based nonparametric image super-resolution. *EURASIP Journal on Applied Signal Processing*, pages 1–11, 2006.

[90] Sina.Farsiu, Dirk.Robinson, Michael.Elad, and Peyman.Milanfar. "dynamic demosaicing and color super-resolution of video sequences". *Proc. of SPIE*, 5562:169–178, 2004.

[91] Sina.Farsiu, Dirk.Robinsona, Michael.Elad, and Peyman.Milanfar. "Robust shift and add approach to super-resolution". *Applications of Digital Image Processing XXVI*, Proceedings Vol. 5203:121–130, 2003.

[92] Sina.Farsiu, Michael.Elad, and Peyman.Milanfar. "Multi-frame demosaicing and super-resolution from under-sampled color images". *IS and T/SPIE 16th Annual Sympo on Electronic Imaging, January*, pages 222–233, 2004.

[93] S.Kother.Mohideen, Dr. S. Arumuga Perumal, and Dr. M.Mohamed Sathik. "Image de-noising using discrete wavelet transform". *IJCSNS International Journal of Computer Science and Network Security*, 8:213–216, 2008.

[94] S.Mallat. "multiresolution approximation and wavelets". *Technical Report, Department Computer Inform. Sci., Univ. Pennsylvania, Philadelphia*, 1987.

[95] S.Mallat. "a theory for multiresolution signal decomposition:the wavelet representation". *IEEE Transactions on Patterns Anallysis and Machine Intelligence*, 11(7):674–693, 1989.

[96] Stark, Henry, Oskoui, and Peyman. "High-resolution image recovery from image-plane arrays, using convex projections". *Optical Society of America, Journal, A: Optics and Image Science*, 6:1715–1726, 1989.

[97] Shikha.Tripathi Subrahmanyam.Ravishankar, Challapalle.Nagadastagiri.Reddy and K.V.V.Murthy. "Image super resolution using sparse image and singular values as priors". *Computer Analysis of Images and Patterns,Lecture Notes in Computer Science,Volume 6855/2011,pp: 380-388,.*

[98] Tamal.Bose. *Digital Signal and Image processing*. John Wiley and Sons.

[99] Todd.A.Stephenson and Tsuhan.Chen. "Adaptive markov random fields for example-based super-resolution of faces". *EURASIP Journal on Applied Signal Processing*, pages 1–11, 2006.

[100] Tsai.R.Y and Huang.T.S. "Multiframe image restoration and registration". *Advances in Computer Vision and Image Processing*, 22:317–339, 1984.

[101] Vagefi.M.R, Lin C.C, and McCann J.D. "Color theory and color imaging systems : Past, present and future". *The Journal of imaging science and technology*, 42:70–78, 1998.

[102] Velisavljevic.V, Beferull-Lozano.B, Vetterly.M., and Dragotti P.L. "Directionlets: anisotropic multi directional representation with separable filtering". *IEEE Transactions Image Processing*, page 1916–1933, 2006.

[103] Vladan.Velisavljevi, Baltasar.Beferull-Lozano1, Martin.Vetterli1, and Pier.Luigi.Dragotti. "discrete multi-directional wavelet bases". In *International Conference on Image processing*, volume 1, pages 1025–8, 2003.

[104] Vrigkas.M, Nikou.C, and Kondi.L.P. "On the improvement of image registration for high accuracy super-resolution". *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),pp:981 - 984,May 2011*.

[105] M.Vetterli V.Velisavljevic, B.Beferull-Lozano and P.L.Dragotti. "approximation power of directionlets". *Proceedings of the IEEE International Conference on Image Processing,Genova, Italy*, September 2005.

[106] Wei.Zhang and Wai-Kuen.Cham. "A single image based blind super-resolution approach". *15th IEEE International Conference on Image Processing,*, 4:329–332, 2008.

[107] WenYi.Zhao and Harpreet.S. Sawhney. "Is super-resolution with optical flow feasible?". *Computer Vision — ECCV ,Lecture Notes in Computer Science*, 2350:599–613, 2002.

[108] William.T.Freeman, Thouis.R.Jones, and Egon.C.Pasztor. "Example-based super-resolution". *IEEE,Image-Based Modeling and Rendering, and Lighting*, 82:56–65, 2002.

[109] Wim.Sweldens. The lifting scheme: A new philosaphy in biorthogonal wavelet constructions. *Disc. Math*, 109:13–26, 1992.

[110] X.H.Yuan, Z.D.Zhu, , and G. Zhang. "Target-aided sar image intelligent compression". *Progress In Electromagnetics Research B,vol.20*, 20:285–302, 2010.

[111] Xiaofeng.Ren, Charless.C.Fowlkes, and Jitendra.Malik. "Learning probabilistic models for contour completion in natural images". *Int J Comput Vision*, 77:47–63, 2007.

[112] Xinbo.Gao, Kaibing.Zhang, Dacheng.Tao, and Xuelong.Li. "Joint learning for single-image super-resolution via a coupled constraint". *IEEE Transactions on Image Processing*, 157:469–480, 2009.

[113] Yacov.Hel-Or and Daniel.Keren. "Demosaicing of color images using steerable wavelets". *HP Labs Israel,Tech. Rep.HPL-2002-206R1*.

[114] Yi.Tang, Pingkun.Yan, Yuan.Yuan, and Xuelong.Li. "Single-image super-resolution via local learning". *International Journal of Machine Learning and Cybernetics*, pages 15–23, 2011.

[115] Yu.Hu, Kin-Man.Lam, Guoping.Qiu, and Tingzhi.Shen. "From local pixel structure to global image super-resolution: A new face hallucination framework". *IEEE Transactions on Image Processing*, 20:433–445, Feb.2011.

[116] Zheng.Zhihui, Wang.Bo, and Sun.Kang. "Single remote sensing image super-resolution and denoising via sparse representation". *Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM)*, pages 1–5, 2011.

[117] Zhiwei.Xiong and Xiaoyan.Sun. "super-resolution for low quality thumbnail images". *IEEE International Conference on Multimedia and Expo*, pages 181 – 184, 2008.

[118] Zomet, Hebrew, and Peleg.S. "Multi-sensor super-resolution". *Proceedings. Sixth IEEE Workshop on Applications of Computer Vision,(WACV 2002)*, pages 27–31, 2002.