Simultaneous image fusion and denoising based on multiscale transformation and sparse representation

Tahiatul Islam
Roll: 1109056
Supervisor: Dr. Sheikh Md. Rabiul Islam
Associate Professor, ECE, KUET

Department of Electronics & Communication Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
May, 2016

Acknowledgement

With due homage and honor, I want to express my gratitude to Almighty.

I express my indebtedness to my dedicated supervisor Dr. sheikh Md. Rabiul Islam (Associate Professor, Department of Electronics and Communication Engineering, KUET) for his necessary guidance, suggestions and encouragement to all phases of my work. I am also very grateful to our Departmental Head Dr. Md. Faruque Hossain (Professor, Department of Electronics and Communication Engineering, KUET). I would also like to thank all the teachers who've been much helpful and encouraging in completion of our thesis work.

Author

Abstract

Image fusion is one of the currently active area in image processing society.The goal of image fusion is to increase the total information in a single image from number of source images.In previous year different method has been proposed in fusing image and each of them has its own pros & cons. Multi-scale transform (MST) and sparse representation (SR) based method found most effective in image fusion.In my work i propose a method of image fusion which take the advantage of both MST and SR in fusing image which also able to remove noise introduced in the source images.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

In recent years, image fusion has become an important issue in image processing community. The target of image fusion is to generate a composite image by integrating the complementary information from multiple source images of the same scene. For an image fusion system, the input source images can be acquired from either different types of imaging sensors or a sensor whose optical parameters can be changed, and the output called fused image will be more suitable for human or machine perception than any individual source image. Image fusion technique has been widely employed in many applications such as computer vision, surveillance, medical imaging, and remote sensing.

As shown in fig 1.1 a,b shows the image whose center and boarder are blared respectively and c shows its fused image which contains the total information contain in source image a & b.

I proposed a technique of image fusion based on MST and SR. In MST image is decomposed into band of low pass and high pass based on the decomposition method and sparse representation sparsely represent an image patch by a dictionary matrix. I combine above two method to take advantage in image fusion. Theoretical description of MST and SR will discuss in second chapter.

(a) (b) (c)

Figure 1.1: (a),(b)Sensor image from two source.(c)Fused image of a,b.

1.2 Motivation

The major, recurrent theme throughout this work is my search for a good fusion model for image fusion of multifocous and natural images. In addition, I seek a model

for image fusion based on sparse representation that is capable of approximating the images with only fewer non zero coefficient. The hope is that I would find out a simple model that can fuse multiple images. Finally, the computation time for the model would be small and fusing the images would be easy enough.

1.3 Problem Statement

Being able to fuse multiple image is a very challenging task, but it could have great impact, for instance when taking image by digital camera it focus one object while another in different distance remain blurred or unfocused as shown in figure **??** .My task is to fuse image from different source sensor and generate an image that is fully focus for every object at relative distance in the fused image.

Figure 1.2: Example of a multifocus image.

1.4 Objectives

- First objective is to find efficient method to fuse images.
- Next objective is to remove any artifact produced by fusing transform coefficient.
- To remove noise introduced in source image from fused image.

Chapter 2

Multi-Scale Transformation

Multi-scale transform (MST) theories are the most popular tools used in various image fusion scenarios such as multi-focus image fusion, visible-infrared image fusion, and multimodal medical image fusion. Classical MST-based fusion methods include pyramid-based ones like Laplacian pyramid (LP)[2] ratio of low-pass pyramid (RP)[3] and gradient pyramid (GP)[4],wavelet-based ones like discrete wavelet transform (DWT)[5],stationary wavelet transform (SWT) [6] and dual-tree complex wavelet transform (DTCWT) [7], and multi-scale geometric analysis (MGA)-based ones like curvelet transform (CVT)[8] and nonsubsampled contourlet transform (NSCT) [9].
In the following section i am going to discuss some of above featured transformation.

2.1 Laplacian pyramid Transformation

Pyramid, or pyramid representation, is a type of multi-scale signal representation developed by the computer vision, image processing and signal processing communities, in which a signal or an image is subject to repeated smoothing and subsampling. Pyramid representation is a predecessor to scale-space representation and multiresolution analysis.

2.1.1 Pyramid Generation

There are two main types of pyramids: lowpass and bandpass.
A lowpass pyramid is made by smoothing the image with an appropriate smoothing filter and then subsampling the smoothed image, usually by a factor of 2 along each coordinate direction. The resulting image is then subjected to the same procedure, and the cycle is repeated multiple times. Each cycle of this process results in a smaller image with increased smoothing, but with decreased spatial sampling density (that is, decreased image resolution). If illustrated graphically, the entire multi-scale representation will look like a pyramid, with the original image on the bottom and each cycle's resulting smaller image stacked one atop the other.

A bandpass pyramid is made by forming the difference between images at adjacent levels in the pyramid and performing some kind of image interpolation between adjacent levels of resolution, to enable computation of pixelwise differences.
A variety of different smoothing kernels have been proposed for generating pyramids

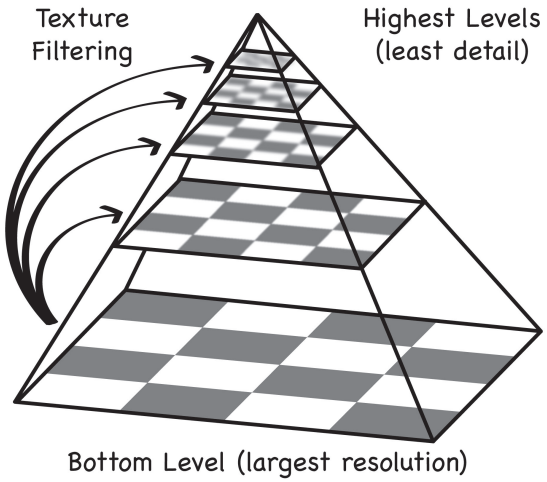

Figure 2.1: Laplacial pyramid decomposition level

2.2 Wavelet Transformation

I start this section by introducing the specific concepts related to the wavelet transform, so that the reader can understand the basic concepts associated with this transform.

The analysis and synthesis procedures lead to the pyramid-structured wavelet decomposition [9].The 1-D multiresolution wavelet decomposition can be easily extended to two dimensions by introducing separable 2-D scaling and wavelet functions as the tensor products of their 1-D complements. Hence, we obtain

$$\begin{aligned} \phi_{LL}(x,y) = \phi(x)\,\phi(y)\,\psi_{LH}(x,y) = \phi(x)\,\psi(y) \\ \psi_{HL}(x,y) = \psi(x)\,\phi(y)\,,\psi_{HH}(x,y) = \psi(x)\,\psi(y) \end{aligned} \tag{2.1}$$

The 2-D wavelet analysis operation consists in =ltering and down-sampling horizontally using the 1-D lowpass =lterL(with impulse responses l(i)) and highpass =lterH(withimpulse responses h(j)) to each row in the imageI(x; y),producing the coeIcient matrices IL(x; y) and IH(x; y). Vertically =ltering and down-sampling follows, using the lowpass and highpass =lters LandHto each column in IL(x; y) andIH(x; y) and produces four subimagesILL(x; y),ILH(x; y), IHL(x; y) and IHH(x; y) for one level of decomposition.ILL(x; y) is a smooth subimage corresponding to the low-frequency band of the MSD and can be considered as a smoothed and subsampled version of the original imageI(x; y), i.e. it represents the coarse approximation of I(x; y).ILH(x; y),IHL(x; y) andIHH(x; y) are detail subimages, which represent the horizontal, vertical and diagonal directions of the imageI(x; y).

Fig **??** depicts one stage in a multiresolution pyramid decomposition of the input imageI(x; y). In order to illustrate the examples of this section, we have used the Haar wavelet transform, although any other set of wavelets could be used. Hence,L$\equiv(1=\sqrt{2})[1;1]$ and H$\equiv(1=\sqrt{2})[1;-1]$.The detailed 2-D pyramid decomposition algorithm, can be expressed as follows: Let I(x; y) be the original image of size M$\times$N, l(i) the analysis lowpass coefficients of a specific wavelet basis, i =0;1,2.....Nh$-$1, where Nl is the support length of the filter L, h(j) the analysis high pass coefficients of a specific wavelet

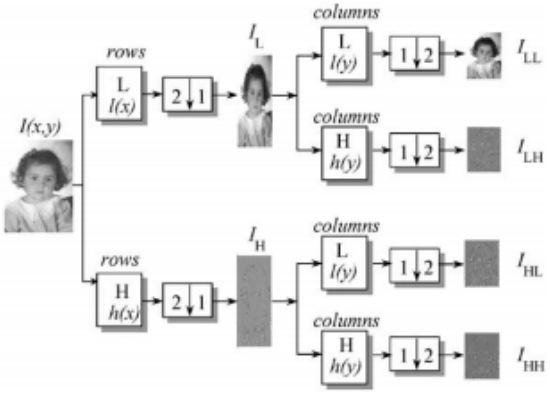

Figure 2.2: One stage of 2-D DWT multiresolution image decomposition (forward wavelet analysis).

basis,j=0,1....Nh−1, where Nh is the support length of the filter H. Then,

$$\begin{aligned} I_L(x,y) &= \frac{1}{N_l}\sum_{i=0}^{N_l-1} I(i)\,.I\left((2x+i)\,modM, y\right), I_H(x,y) \\ &= \frac{1}{N_h}\sum_{j=0}^{N_h-1} h(j)\,.I\left((2x+j)\,modM, y\right) \end{aligned} \tag{2.2}$$

for x=0,1,2.....M/2−1 and y=0,1,2.....N−1

$$\begin{aligned} I_{LL}(x,y) &= \frac{1}{N_l}\sum_{i=0}^{N_l-1} I(i)\,.I\left(x,(2x+i)\,modN\right), I_{LH}(x,y) \\ &= \frac{1}{N_h}\sum_{j=0}^{N_h-1} h(j)\,.I\left(x,(2y+j)\,modN\right) \end{aligned} \tag{2.3}$$

$$\begin{aligned} I_{HL}(x,y) &= \frac{1}{N_l}\sum_{i=0}^{N_l-1} I(i)\,.I_H\left(x,(2y+i)\,modN\right), I_{HH}(x,y) \\ &= \frac{1}{N_h}\sum_{j=0}^{N_h-1} h(j)\,.I_H\left(x,(2y+j)\,modN\right) \end{aligned} \tag{2.4}$$

for x=0,1,2.....m/2-1 and y=0,1,2.....N/2-1

The 2-D pyramid algorithm can iterate on the smooth subimage $I_{LL}(x,y)$ to obtain four coefficient matrices in the next decomposition level and so on. This is illustrated in Figs.**??** and which correspond to one-and two-level image decompositions, respectively.

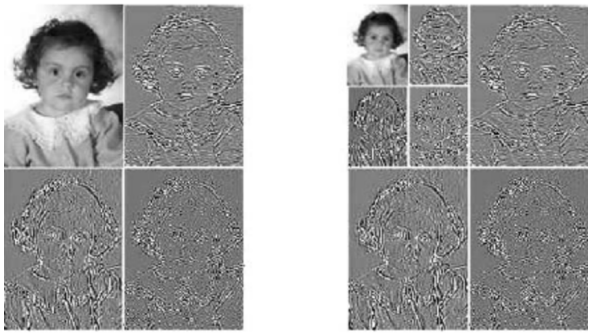

Figure 2.3: A representation of (a) one-level and (b) two-level image decomposition.

Some wavelet-based applications do not require all coefficients, only the most relevant. So an additional procedure can be carried out to eliminate non-significant coefficient by thresholding, since these have a magnitude close to zero. After thresholding, only the desired coefficients remain.The threshold value can be chosen as $T = \sigma\sqrt{2\log n/\sqrt{n}}$ in where σ is the standard deviation of the coefficients and n is the total size of samples. Another possibility is to fix T in order to replace a percentage of the coefficients with the smallest magnitude to zero. Obviously, the cancellation of coefficients implies a loss of information.

The inverse 2-D wavelet transform can be implemented using a backward 2-D pyramid algorithm. The 2-D wavelet synthesis operation consists in up-sampling and filtering vertically using the 1-D synthesis lowpass filter L(with impulse responses l(i)) and highpass filter H(with impulse responses h(j)) for each column in the subimage. Horizontal up-sampling and =ltering then follows, using the lowpass Land highpass filter H, for each row of the reversed image. Fig. **??** shows one stage in a wavelet reconstruction

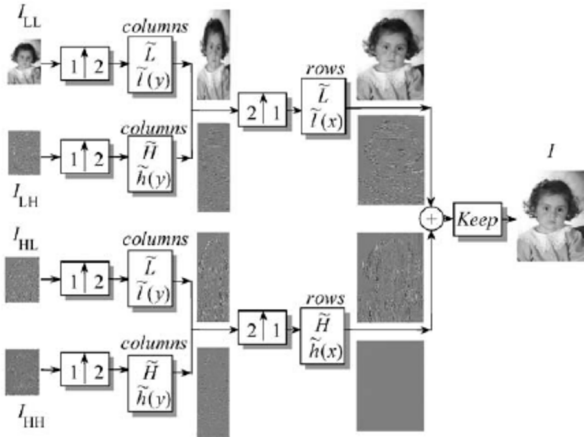

Figure 2.4: One stage of 2-D DWT multiresolution image reconstruction (backward wavelet synthesis).

2.3 Fast Discrete Curvelet Transforms

Despite considerable success, intense research in the last few years has shown that classical multiresolution ideas are far from being universally effective. Indeed, just as people recognized that Fourier methods were not good for all purposes and consequently introduced new systems such as wavelets researchers have sought alternatives to wavelet analysis. In signal processing for example, one has to deal with the fact that interesting phenomena occur along curves or sheets, e.g. edges in a two-dimensional image. While wavelets are certainly suitable for dealing with objects where the interesting phenomena, e.g. singularities, are associated with exceptional points, they are ill-suited for detecting, organizing, or providing a compact representation of intermediate dimensional structures. Given the significance of such intermediate dimensional phenomena, there has been a vigorous research effort to provide better adapted alternatives by combining ideas from geometry with ideas from traditional multiscale analysis

2.3.1 Why a discrete curvelet transform?

Curvelets are interesting because they efficiently address very important problems where wavelet ideas are far from ideal. I give three examples:

- Optimally sparse representation of objects with edges. Curvelets provide optimally sparse representations of objects which display curve-punctuated smoothness except for discontinuity along a general curve with bounded curvature. Such representations are nearly as sparse as if the object were not singular and turn out to be far more sparse than the wavelet decomposition of the object.

- Optimally sparse representation of wave propagators. Curvelets may also be a very significant tool for the analysis and the computation of partial differential equations. For example, a remarkable property is that curvelets faithfully model the geometry of wave propagation. Indeed, the action of the wave-group on a curvelet is well approximated by simply translating the center of the curvelet along the Hamiltonian flows. A physical interpretation of this result is that curvelets may be viewed as coherent waveforms with enough frequency localization so that they behave like waves but at the same time, with enough spatial localization so that they simultaneously behave like particles

- Optimal image reconstruction in severely ill-posed problems. Curvelets also have special micro local features which make them especially adapted to certain reconstruction problems with missing data.

2.3.2 Digital Curvelet Transform via Wrapping

The 'wrapping' approach assumes the same digital coronization as in Section 3.1, but makes a different, somewhat simpler choice of spatial grid to translate curvelets at each scale and angle. Instead of a tilted grid, we assume a regular rectangular grid

and define 'Cartesian' curvelets in essentially the same way as before,

$$c(j,l,k) = \int \hat{f}(\omega)\, U_j\left(S^{-1}\omega\right) e_{i<b\omega>} \tag{2.5}$$

The difficulty behind this approach is that, in the frequency plane, the window $U_{j,l}[n1,n2]$ does not fit in a rectangle of size $2^j 2^{j/2}$, aligned with the axes, in which the 2D IFFT could be applied to compute 2.5. After discretization, the integral over ω becomes a sum over n1, n2 which would extend beyond the bounds allowed by the 2-D IFFT. The resemblance of 2.5 with a standard 2D inverse FFT is in that respect only formal.hanged.

To understand why respecting rectangle sizes is a concern, we recall that $U_{j,l}$ is supported in the parallelepipedal region

2.3.3 Architecture of the FDCT

- Apply the 2D FFT and obtain Fourier samples $\hat{f}[n1,n2], -n/2 \leq n1,n2 \leq n/2$
- For each scale j and angle, form the product $U_{j,l}\hat{f}[n1,n2]$
- Wrap this product around the origin and obtain
- Apply the inverse 2D FFT to each $f_{j,l}$, hence collecting the discrete coefficients $c^D(j,l,k)$

Chapter 3

Sparse Representation of signal

There are various methods available to implement image fusion. Basically, these methods can be categorized into two categories. The first category is the spatial domain-based methods, which directly fuse the source images into the intensity values . The other category is the transformed domain-based methods, which fuse images with certain frequency or time frequency transforms Assuming that F(·) represents the "fusion operator," the fusion methods in the spatial domain can be summarized as

$$I_F = F(I_1, I_2 \dots I_k) \tag{3.1}$$

The simplest fusion method in spatial domain just takes the pixel-by-pixel average of the source images. However, this method often leads to undesirable side effects, such as reduced contrast . If the source images are not completely registered, then a single pixel-based method, such as spatial gradient (SG) based method, always results in artifacts in the fused image. Therefore, some more reasonable methods were proposed to fuse source images with divided blocks or segmented regions instead of single pixels. However, the blockbased fusion methods usually suffer from blockness in the fused image. For the region-based method, the source images are first segmented, and the obtained regions are then fused using their properties, such as spatial frequency or SG. The segmentation algorithms, usually complicated and time consuming, are of vital importance to the fusion quality.A more popular method that has been explored in recent years is by using multiscale transforms The transformed domain-based methods can be summarized as

$$I_F = T^{-1}(F(T(I_1), T(I_2) \dots T(I_j)) \tag{3.2}$$

Because the fused image obtained by transform domain-based algorithms is globally created, a little change in a single coefficient of the fused image in the transformed domain may cause all the pixel values to change in spatial domain. As a result, undesirable artifacts may be produced in the fusion process using the multiresolution transform-based methods in some cases Obviously, effectively and completely extracting the underlying information of the original images would make the fused image more accurate. Different from multiscale transformations, the sparse representation using an overcomplete dictionary that contains prototype signal atoms describes signals by sparse linear combinations of these atoms [**?**] Two main characteristics of sparse representation are its overcompleteness and sparsity [**?**]. Overcompleteness means that the number of basis atoms in the dictionary exceeds the number of image pixels or signal dimensions. The overcomplete dictionary that contains rich transform bases allows for more stable and meaningful representation of signals. Sparsity means that the coefficients corresponding to a signal are sparse, that is to say, only "a few descriptions"

can describe or capture the significant structure information about the object of interest. Benefiting from its sparsity and overcompleteness, sparse representation theory has successfully been applied in many practical applications, including compression, denoising, feature extraction, classification, and so on [**?**]. Recent studies have shown that common image features can also be accurately described by only a few coefficients or "a few descriptions" [**?**]. In general, sparse representation is a global operation, in the sense that it is based on the gray-level content of an entire image. However, the image fusion quality depends on the accurate representation of the local salient features of source images. Therefore, a "sliding window" technique is adopted to achieve better performance in capturing local salient features and keeping shift invariance.

3.1 Sparse Representaion

SR assumes that the signal $Y \in \mathbf{R}^n$ can be represented as a linear combination of given atoms.These atoms consist of an overcomplete dictionary $D \in \mathbf{R}^{n*k}, within << K$. The representation of y may either be exact $D\theta = y$ or approximate $|D\theta - y| \leq \epsilon$ where ϵ is the specified error.The details about dictionary D is discuss in following section. The vector $\theta \in \mathbf{R}^k$ is the coefficients of the signal. The coefficients with the fewest number of nonzero coefficients is certainly an appealing representation. Exact determination of the sparsest representation proves to be an non-deterministic polynomial time-hard problem. Approximate solutions are considered instead. Basically, two approaches were proposed in previous researches. One approach is based on greedy algorithms such as Matching Pursuit (MP) and Orthogonal Matching Pursuit (OMP).All of these approaches are based on replacing the l_0 with the l_p -norm,where l_p (0¡p≤1) is defined as $\| \theta \|_p = (\sum_i | \theta_i |^p)^{1/p}$

Let $Y_i \in \mathbf{R}_{n*L} \, (i = 1, 2, ...\lambda)$ denote the L signals with dimension from the i'th sensor. We can represent Y_i with common component $\Theta c \in \mathbf{R}^{K*L}$ and innovation component $\Theta_i^u \in \mathbf{R}^{k*L}$of sparse coefficient matrix,and noise $n_i \in \mathbf{R}^{n*L}$.

$$Y_i = Y^c + Y_i^u = D\Theta^c + D\Theta_i^u + n_i \tag{3.3}$$

The concatenated source images matrix can be represented sparsely by the concatenated coefficient matrix.

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_\Lambda \end{pmatrix} = \begin{pmatrix} D & D & 0 & \dots & 0 \\ D & 0 & D & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D & 0 & 0 & \dots & D \end{pmatrix} \begin{pmatrix} \Theta^c \\ \Theta_l^u \\ \vdots \\ \Theta_\Lambda^u \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_\Lambda \end{pmatrix} \tag{3.4}$$

where $D \in \mathbf{R}^{n*K}$is the overcomplete dictionary which shared by both the common component and the innovation component. Let $Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_\Lambda \end{pmatrix}$

$$D = \begin{pmatrix} D & D & 0 & \dots & 0 \\ D & 0 & D & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D & 0 & 0 & \dots & D \end{pmatrix}, \Theta = \begin{pmatrix} \Theta^c \\ \Theta_l^u \\ \vdots \\ \Theta_\Lambda^u \end{pmatrix}, n = \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_\Lambda \end{pmatrix}$$

Then, Eq 3.4 can be rewritten as follows

$$Y = D\Theta + n \tag{3.5}$$

A crucial step of the SR/JSR is the selection of such a dictionary D. Predefined basis functions like curvelets, bandlets, variants of wavelets, etc., can be used. However, the success of such prespecified dictionaries is often limited by their suitability in capturing the structure in the signals under consideration. For example, image contents have SR over wavelet dictionary, but audio signals are better represented by sinusoids. A more generalized approach is to learn the basis vectors that are specialized in representing the signal in question. Several algorithms available in the literature deal with this problem, such as the K-SVD algorithm, the method of optimal directions (MOD), and the Majorization method [9, 12, 13] The K-SVD algorithm is slow due to the SR computation and SVD operation exists at its each iteration.
The MOD does the most natural thing as $\hat{D} = argmin_D \parallel Y - DX \parallel_F^2 = YX^T\left(XX^T\right)^{-1}$ However, XX^T may not be always with full rank. The Majorization method, for the dictionary update, is slow due to using the"Landweber"update (which is a gradient descent update) as described in Ref [14]

3.2 SR for Image Fusion

Since the sparse representation globally handles an image, it cannot directly be used with image fusion, which depends on the local information of source images. In my method, we divide the source images into small patches and use the fixed dictionary D with small size to solve this problem. In addition, a sliding window technique is adopted to make the sparse representation shift invariant, which is of great importance to image fusion.
We assume that source image I is divided into many image patches. As shown in Fig.**??**, to facilitate the analysis, the $\left(j^{th}\right)$ patch with size $n \times n$ is lexicographically ordered as a vector V^j. Then, V^j can be expressed as

$$V^j = \sum_{t=1}^{T} S^j(t)\, d_t \tag{3.6}$$

where d_t is an atom from a given overcomplete dictionary, and $D\,[[d_1 \dots d_t \dots d_T]$, which contains T atoms.$S_j = [s_j(1) \dots s_j(t) \dots s_j(T)]$ is the sparse representation

Assume that the vectors responding to all the patches in image I are constituted into one matrix V. Then, Vcan be expressed as

$$D = [d_1, d_2 \dots d_T] \begin{pmatrix} S^1(1) & S^2(1) & \dots & S^j(1) \\ S^1(2) & S^2(2) & \dots & S^j(2) \\ \vdots & \vdots & \vdots & \vdots \\ S^1(T) & S^2(T) & \dots & S^j(T) \end{pmatrix} \tag{3.7}$$

where J is the number of image patches. Let $S = [s^1, s^2, ..., s^j]$. Then, eq. 3.7 can be expressed as

$$V = DS \tag{3.8}$$

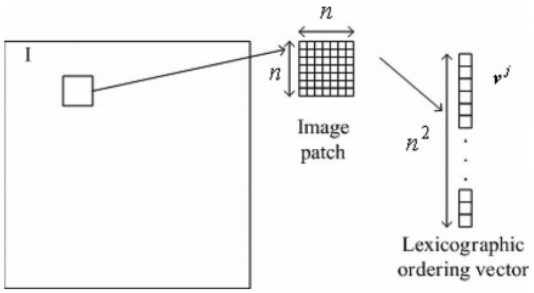

Figure 3.1: Selected image patch and its lexicographic ordering vector.

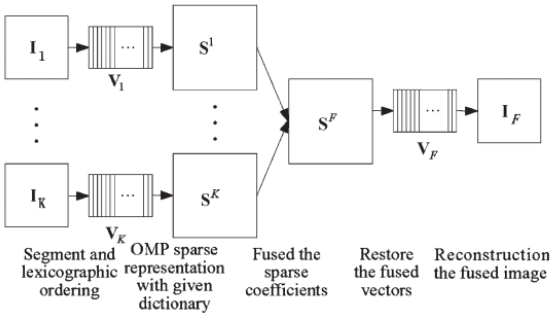

Figure 3.2: Schematic diagram of the proposed SR-based fusion method.

where S is a sparse matrix.

3.3 Training Dataset

The Dictionary D is learnt from a database consisting of different type of multifocus images like medical image, infrared images, etc.

I first randomly assign value in dictionary D of size $64 * 256$ from the image patches of training dataset image then i do sparse coding to get the sparse matrix of signal then i run K-SVD algorithm to update the dictionary.

The training data consists of 100,000 $8 * 8$ patches, which are randomly sampled from a database of 40 high-quality natural images. The dictionary size is set to 256 and the iteration number of K-SVD is fixed to 180.

Chapter 4

Image Denoising

4.1 Introduction

In many applications, image denoising is used to produce good estimates of the original image from noisy observations. The restored image should contain less noise than the observations while still keep sharp transitions (i.e. edges).

Wavelet transform, due to its excellent localization property, has rapidly become an indispensable signal and image processing tool for a variety of applications, including compression and denoising [21, 22, 23]. Wavelet denoising attempts to remove the noise present in the signal while preserving the signal characteristics, regardless of its frequency content. It involves three steps:a linear forward wavelet transform, nonlinear thresholding step and a linear inverse wavelet transform.

Wavelet thresholding (first proposed by Donoho [21, 22, 23]) is a signal estimation technique that exploits the capabilities of wavelet transform for signal denoising. It removes noise by killing coefficients that are insignificant relative to some threshold, and turns out to be simple and effective, depends heavily on the choice of a thresholding parameter and the choice of this threshold determines, to a great extent the efficiency of denoising. Researchers have developed various techniques for choosing denoising parameters and so far there is no "best" universal threshold determination technique.

In my method i decompose the image by MST used in the transformation then apply BayesShrink soft thresolding to remove the insignificant coefficient the The aim of this project was to study various thresholding techniques such as SureShrink [21], VisuShrink [23] and BayesShrink [22] and determine the best one for image denoising.

4.2 Hard and soft thresholding

Hard and soft thresholding with threshold ¸ are defined as follows:
The hard thresholding operator is defined as:

$$D(U,T) = \begin{cases} T & \text{if } U > T \\ 0 & \text{if } otherwise \end{cases} \tag{4.1}$$

The soft thresholding operator on the other hand is defined as:

$$D(U,T) = sgn\,(U) \max\,(0, \mid U \mid -\lambda) \tag{4.2}$$

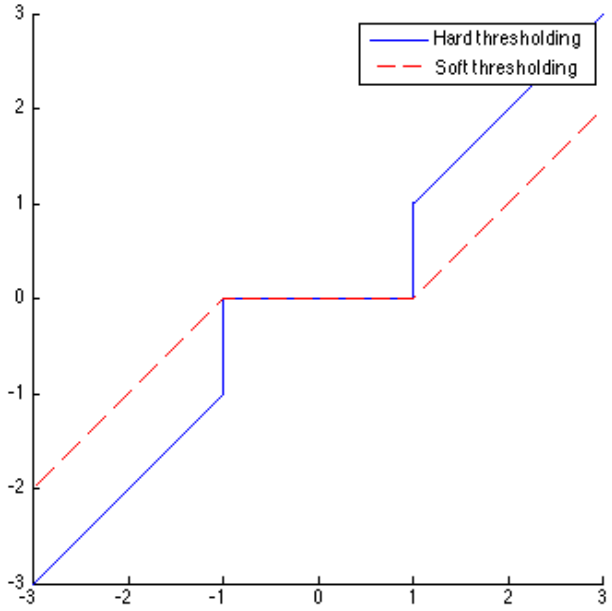

Figure 4.1: soft and hard thresholding

Hard threshold is a “keep orkill” procedure and is moreintuitively appealing. The transfer function of the same is shown in Fig 4.1. The alternative, soft thresholding (whose transfer function is also shown in Fig 4.1), shrinks coefficients above the threshold in absolute value. While at first sight hard thresholding may seem to be natural, the continuity of soft thresholding has some advantages. It makes algorithms mathematically more tractable [23]. Moreover, hard thresholding does not even work with some algorithms such as the GCV procedure . Sometimes, pure noise coefficients may pass the hard threshold and appear as annoying ’blips’ in the output. Soft thesholding shrinks these false structures.

4.3 Image denoising using thresholding

As one may observe, threshold selection isan important question when denoising. A small threshold may yield a result close to the input, but the result may still be noisy. A large threshold on the other hand, producesa signal with a large number of zero coefficients. This leads to a smooth signal.Paying too much attention to smoothness, however, destroys details and in image processing may cause blur and artifacts.
The problem boils down to finding an optimal threshold such that the mean squared error between the signal and its estimate is minimized. The wavelet decomposition of an image is done as follows: In the first level of decomposition, the image is split into 4 subbands, namely the HH, HL, LH and LL subbands. The HH subband gives the diagonal details of the image; the HL subband gives the horizontal features while the LH subband represents the vertical structures. The LL subband is the low resolution residual consisting of low frequency components and it is this subband which is further

split at higher levels of decomposition. The different methods for denoising investigate differ only in the selection of the threshold.
I use the above perception of denoising for every MST i used. So the basic procedure is

- Calculate the MST coefficient of the image
- Threshold the coefficient
- Inverse transform the coefficient to get the denoised image.

4.3.1 BayesShrink

BayesShrink is an adaptive data-driven threshold for image denoising via wavelet soft-thresholding.The threshold is driven in a Bayesian framework, and we assume generalized Gaussian distribution (GGD) for the wavelet coefficients in each detail subband and try to find the threshold T which minimizes the Bayesian Risk.It is found that BayesShrink performs better than SureShrink in terms of MSE. The reconstruction using BayesShrink is smoother and more visually appealing than one obtained using SureShrink.

The BayesShrink threshold is given by :

$$\hat{T}\left(\hat{\sigma_I}\right) = \frac{\hat{\sigma_n}}{\hat{\sigma_I}} \tag{4.3}$$

where σ_n and σ_I are noise and signal standard deviations respectively.To estimate the noise varianceσ_n^2 from the subband details, the median estimator is used on the 1-D subband coefficients:

$$\hat{\sigma_n} = median\left(\mid Details \mid\right)/0.6745 \tag{4.4}$$

the observed signal S is considered to be $S = I + n$ and signal (I) and noise (n)are assumed to be independent.Therefore,

$$\sigma_S^2 = \sigma_I^2 + \sigma_n^2 \tag{4.5}$$

where $\sigma_S^2$is the variance of the observed signal. So $\sigma_I^2$is estimated by:

$$\hat{\sigma_I} = \sqrt{\max\left(\left(\hat{\sigma_S^2} - \hat{\sigma_n^2}\right), 0\right)} \tag{4.6}$$

Chapter 5

Propsed method

5.1 Proposed fusion framework

the schematic diagram of the proposed fusion framework is shown in Fig5.1 . In the Fig5.1 only the fusion of two source images is considered while the proposed framework can be straightforwardly extended to fuse more than two images. The detailed fusion scheme contains the following four steps

1. **MST decompositon:** Perform a specific MST on the two source images $\{I_A, I_B\}$to obtain their low-pass bands $\{L_A, L_B\}$ and high-pass bands which are uniformly denoted as $\{H_A, H_B\}$
2. **Thresholding:** Perform thresholding as obtain in eq.4.3 on the low pass and high pass band to remove the unnecessary coefficient from the decomposition.
3. **Low-pass fusion:**
   - Apply the sliding window technique to divide $\{I_A, I_B\}$ into image patches of size $\sqrt{n} \times \sqrt{n}$ from upper left to lower right with a step length of s pixels. Suppose that there are T patches denoted as $\{P_A^i\}_{i=1}^T$ and $\{P_B^i\}_{i=1}^T$ in L_A and L_B respectively respectively
   - For each position i,rearrange $\{P_A^i, P_B^i\}$ into column vectors rearrange $\{\hat{V}_A^i, \hat{V}_B^i\}$ by
   $$\hat{V}_A^i = V_A^i - \bar{V}_A^i.1 \tag{5.1}$$
   $$\hat{V}_B^i = V_B^i - \bar{V}_B^i.1 \tag{5.2}$$
   whre 1 denotes an all=one valued $n \times 1$ vector $\bar{V}_A^i$ and $\bar{V}_B^i$ are mean values of all the elements in V_A^i and V_B^i respectively.
   - Calculate the sparse coefficient vectors $\{\alpha_A^i, \alpha_B^i\}$ of $\{\hat{V}_A^i, \hat{V}_B^i\}$using the orthogonal matching pursuit (OMP) algorithm [24] by
   $$\alpha_A^i = \arg\min \parallel \alpha \parallel_0 \; st \parallel \hat{V}_A^i - D\alpha \parallel_2 < \epsilon \tag{5.3}$$
   $$\alpha_B^i = \arg\min \parallel \alpha \parallel_0 \; st \parallel \hat{V}_B^i - D\alpha \parallel_2 < \epsilon \tag{5.4}$$
   where $\mathbf{D}$ is the learned dictionary.

- Merge $\alpha_A^i and \alpha_B^i$ with the "max-L1" rule to obtain the fused sparse vector

$$\alpha_F^i = \begin{cases} \alpha_A^i & \text{if } \| \alpha_A^i \|_1 > \| \alpha_B^i \|_1 \\ \alpha_B^i & \text{otherwise} \end{cases} \tag{5.5}$$

  The fused result of $V_A^i and V_B^i$ is calculated by

$$V_F^i = D\alpha_F^i + \bar{V}_F^i.1 \tag{5.6}$$

  where the merged mean value $\bar{V}_F^i$ is obtained by

$$\bar{V}_F^i = \begin{cases} \bar{V}_A^i & \text{if } \alpha_F^i = \alpha_B^i \\ \bar{V}_B^i & \text{otherwise} \end{cases} \tag{5.7}$$

- Iterate the above process for all the source image patches in $\{P_A^i\}_{i=1}^T$ and $\{P_B^i\}_{i=1}^T$to obtain all the fused vectors $\{V_F^i\}_{i=1}^T$. Let L_f denotes the low-pass fused result. For each V_F^ireshape it into a patch $p_F^i$and then plug p_F^iFinto its original position in L_F.As patches are overlapped, each pixel's value in L_F is averaged over its accumulation times

4. **High-pass fusion :** Merge $H_A and H_B$to obtain H_F with the popular "max-absolute" rule using the absolute value of each coefficient as the activity level measurement. Then, apply the consistency verification scheme to ensure that a fused coefficient does not originate from a different source image from most of its neighbors. This can be implemented via a small majority filter

5. **MST reconstruction** Perform the corresponding inverse MST over $L_F and H_F$ to reconstruct the final fused image I_F

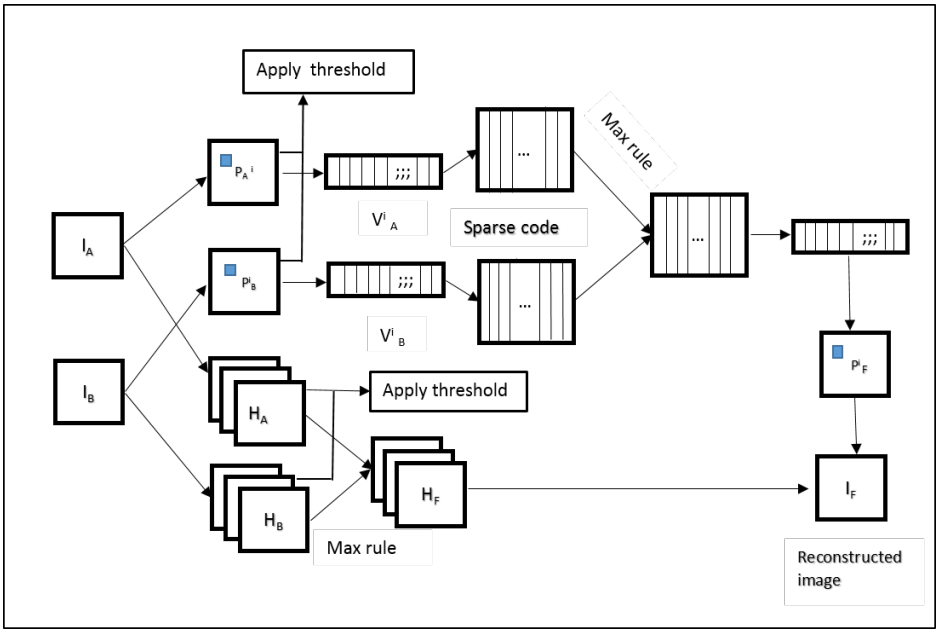

Figure 5.1: The schematic diagram of the proposed fusion framework

5.2 Advantage over the MST based methods

For the conventional MST-based image fusion methods (the high-pass bands are merged with the "max-absolute" rule while the low-pass bands are fused using the "averaging" rule), there are two main drawbacks as follows

The first one is the loss of contrast. Since most energy of an image is contained in the low-pass band (even though the decomposition level is set to 4 according to the analysis , the "averaging" fusion rule tends to lose some energy in the source images. For multi-focus image fusion, this phenomenon is not obvious because the source images are captured from the same type of sensors. However, for the fusion of multimodal images such as visible-infrared and medical images, the fused results of the MST-based methods are often in low contrast. This is mainly because different imaging modalities reflect different physical attributes, so a same region in different source images may have different brightness. For example,Fig. **??** shows a pair of computed tomography (CT) and magnetic resonance (MR) images. It can be seen that the CT image mainly focuses on dense structures like bones, while the MR image provides excellent soft-tissue details. When the "averaging" rule is used for low-pass fusion, the energy contained in those regions will be lost to a large extent. As a result, the contrast of those regions in the fused image will decrease a lot after MST reconstruction

The second one is the difficulty in selecting the MST decomposition level. On one hand, to ensure enough spatial details can be extracted from the source images, the decomposition level cannot be too small such as 1 or 2. On the other hand, Li et al.[25] experimentally verified that when the decomposition level is too large, one coefficient in the low-pass band have an impact on a large set of pixels in the fused image, so an error in the low-pass band (mainly caused by noise or mis-registration between the source images) will lead to serious artificial effects. Moreover, when the decomposition level becomes larger, the quality of high-pass fusion is also more sensitive to noise and mis-registration. Therefore, when the source images are not precisely registered, the decomposition level cannot be too large. Particularly, for multi-focus image fusion, due to the different imaging parameters (e.g. focal length) for multiple source images, the locations of object edges in different source images are often not exactly the same for their different sharpness. A typical example is shown in Fig.5.2 Between the two source images, both the borders and numbers of the two clocks in the scene have different sharpness, so it is practically impossible to make an accurate registration. Thus, a compromise on decomposition level should be made for the consideration of extracting enough spatial details and being robust to mis-registration.

As a smart blending approach, the SR-based image fusion scheme is combined into the MST-based fusion methods to overcome the above two defects. In the proposed framework, the SR-based scheme is employed to fuse the MST low-pass bands.In Section2, after applying the "max-L1" rule in Eq.5.4, we transfer the energy in source images to the fused image by Eq.5.7.Therefore, the contrast in the fused image is improved. For the second defect, by extracting spatial details in low-pass band with the SR-based fusion scheme, the decomposition level can be set less than 4 for multi-focus image fusion to make the method more robust to mis-registration. Thus, the difficulty

in determining decomposition level can be well solved.

5.3 Advantages over the SR-based method

The conventional SR-based image fusion method[14] mainly has the following three defects.

The first one is the fine details in source images like textures and edges tend to be smoothed for the following two reasons. First, the signal representation ability of the dictionary may be not sufficient for fine details, which means that the reconstruction result is not approximate to the input signal. As we know, the representation ability of the over-completed dictionary relies much on the number of atoms in it, but a dictionary with a large size will directly increase the computational cost. More importantly, the study in[22]shows that a highly redundant dictionary may lead to potential visual artifacts in the reconstruction result, especially when the input signal is corrupted by noise. Thus, a compromise on dictionary size is usually required. A typical example is that the dictionary size is 256 when the input signal is 64 dimensional 8×8. Second, the usage of sliding window technique may also cause smoothness. The step length of the sliding window is usually set to 1 when fusing images directly in spatial domain to avoid blocking effects. However, when the adjacent patches are greatly overlapped, some details in the fused image will be smoothed.

The second one is the "max-L1" rule may cause spatial inconsistency in the fused image when the source images are captured by different imaging modalities. As mentioned before, for multimodal image fusion, a region may be very bright in one source image while very dark in another, but the region in both of them may be very "flat" with few fine details. Note that although a region in each of two source images is visually "flat", there still exists little difference between the two source images in terms of variance, and the difference is usually consistent over all the patches in that region. That is to say, if one patch in the region of source image A has a larger variance than the corresponding patch in source image B, then most of the other patches in that region of source image A also tend to have larger variances than the corresponding patches in source image B. However, since the difference is very tiny, the "max-L1" fusion rule will become very sensitive to the random noise in spatial domain because a small change of value at a pixel may influence the fusion result of several patches. As a result, the fused patches in that region may originate from different source images, which will lead to spatial inconsistency in the fused image. Since the SR-based method handles patches in spatial domain, the impact of high-frequency noise is considerable.

The third one is the low computational efficiency. Since the sliding window's step length should be small enough, the sparse coding technique is performed on a large number of image patches. For instance, when the patch size is 8×8 and the step length is set to 1, there are 62001 patches to be processed for a source image of size 256×256 . In this case, it usually takes several minutes to fuse two source images with the SR-based method

The proposed fusion framework can effectively overcome the above three defects of the SR-based method. In our fusion framework, the high-frequency spatial information is separated by performing MST and extracted by the "max-absolute" rule. Meanwhile, the representation ability of the dictionary is enough to satisfy the reconstruction accuracy for low-frequency components. Furthermore, we will show in the next section that the sliding window's step length in low-pass bands can be set larger than that in spatial domain. Therefore, the inclination of SR-based method to smooth fine details can be prevented. For the second defect, without high-frequency details, the random noise can be effectively eliminated, so the probability that the patches in a "flat" region originate from different source images will decrease to a large extent, leading to better spatial consistency. Finally, the computational efficiency can also be improved by the proposed framework because the number of patches required to be processed with the sparse coding technique is greatly reduced. For one thing, the step length can be set larger. For another, the low-pass bands of many MSTs such as LP and DWT have smaller size relative to the original image.

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

Figure 5.2: Two pairs of source images. (a,b)multi-focus images , and (c,d) Medical images .

Chapter 6

Experiment setup and fusion evaluation metrics

6.1 Source images

As shown in Fig.6.1, source images grouped into three categories are employed to verify the effectiveness of the proposed fusion framework. Among them, there are multi-focus images Figs.6.1,visible-infrared images Fig. 6.1 and medical images 6.1. For each pair, the two source images are assumed to be pre-registered in our study.

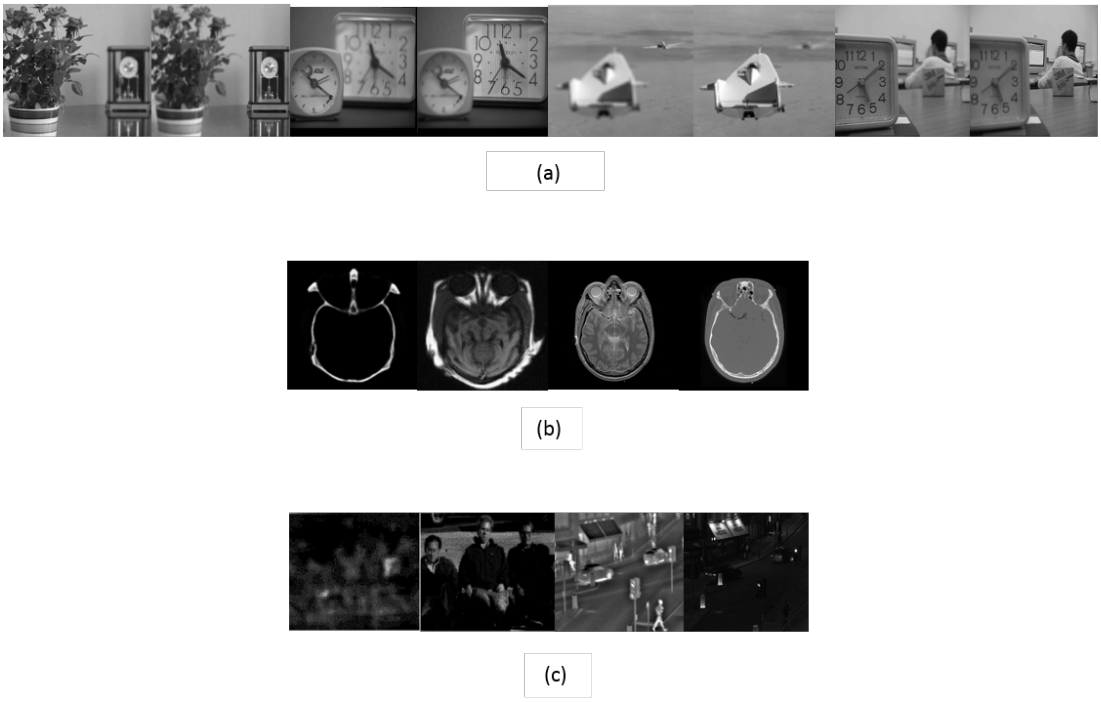

Figure 6.1: The source images used in our experiments. (a) Multi-focus images,(b)medical images and (c) visible-infrared images

6.2 Objective evaluation metrics

It is not an easy task to quantitatively evaluate the quality of a fused image since the reference image (ground truth) does not exist in practice. In recent years, many fusion metrics have been proposed, but none of them is universally believed to be always more

reasonable than others for various fusion scenarios. Thus, it is usually necessary to apply several metrics to make a comprehensive evaluation. In this work, five popular metrics, which are briefly introduced as follows, are employed to quantitatively evaluate the performances of different fusion methods. Uniformly, let A and B denote two source images of size $H \times W$while F represents the fused image.The different metrics are

- Standard deviation (SD). The SD of the fused image is defined as

$$SD = \sqrt{\frac{1}{H \times W} \sum_{x=1}^{H} \sum_{y=1}^{w} (F(x,y) - \mu)^2} \tag{6.1}$$

  where μ is the mean value

- **Entropy (EN)** The EN of the fused image is defined as

$$EN = -\sum_{l=0}^{L-1} P_F(l) \, log_2 P_F(l) \tag{6.2}$$

  where L is the number of gray level and P_F is the normalized histogram of the fused image. In our experiments,L is set to 256.EN is used to measure the amount of information in the fused image

- The gradient based fusion metric Q_G proposed by Xydeas and Petrovic[23]. It is calculated by

$$Q_G = \frac{\sum_{x=1}^{H} \sum_{y=1}^{W} \left(Q^{AF}(x,y) \, W^A(x,y) + Q^{BF}(x,y) \, W^B(xx,y) \right)}{\sum_{x=1}^{H} \sum_{y=1}^{W} \left(W^A(x,y) + W^B(x,y) \right)} \tag{6.3}$$

  The Q_G is a popular fusion metric which computes the amount of gradient information injected into the fused image from the source images

- **PSNR** can reflect the quality of reconstruction. The larger the PSNR is, the less the image distortion is.

$$PSNR = 10log\left(255^2/MSE\right) \tag{6.4}$$

- **Structural similarity (SSIM)** Given the two source images X(X=a orb)and the fused image f,the size of the images are all M×N, let X and f denote the mean of X, f, let $\sigma_x^2$and σ_{Xf} be the variance of X and covariance of X,f, respectively,

$$\sigma_x^2 = \frac{1}{MN-1} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(X(m,n) - \bar{X} \right)^2 \tag{6.5}$$

$$\sigma_{xf} = \frac{1}{MN-1} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(X(m,n) - \bar{X} \right) \left(f(m,n) - \bar{f} \right) \tag{6.6}$$

  Since image signals are generally non-stationary, it is appropriate to measure the number Q_0over local regions and then combine the different results into a single measure. A sliding window w is used in the images.$Q_0(x, f|w))$is computed in each window. Then compute the whole image metric $Q_0(x, f)$

Chapter 7

Experimental result

7.1 Fusion Evaluation measurement

I now will list value of different evaluation metrics stated in chapter:6 for different value of sigma for different type of image given in Fig.6.1.First i will go with multifocus image and fusion result will shown for pair of image and evaluation metrics will give in table for different method

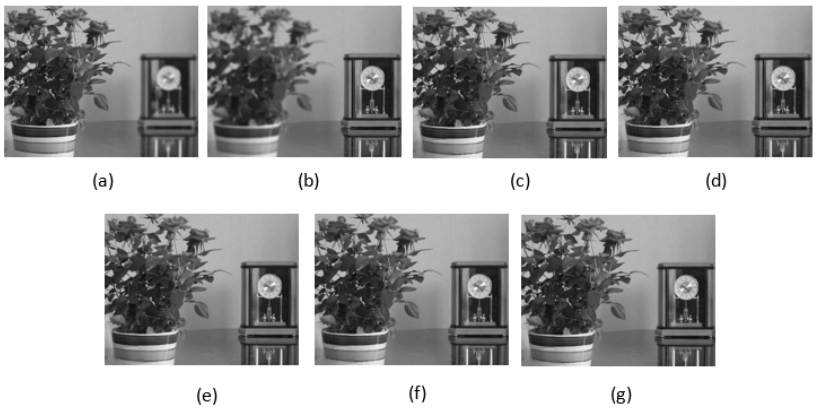

Figure 7.1: Example of multifocus image fusion (a)(b) is source image (c)(d)(e)(f)(g) are fusion result for LP-SR, DWR-SR, DTCWT-ST, CVT-SR, NSCT-SR respectivley

The following table show the evaluation metrics for the above images is given in tabel7.1

method	SD	MI	PSNR	Q_e	SSIM
LP-SR	53.1210	7.0147	43.2496	0.7121	0.9216
DWT_SR	52.8203	6.6233	42.2866	0.6845	0.9109
DTCWT_SR	52.4403	6.5610	43.2982	0.7021	0.9233
CVT_SR	52.7584	7.1643	43.3189	0.7044	0.9237
NSCT_SR	51.9634	6.2550	43.2760	0.6922	0.9254

Table 7.1: Evaluation metrics for different method

The evaluation metrics for different method for medical image is given in table7.2

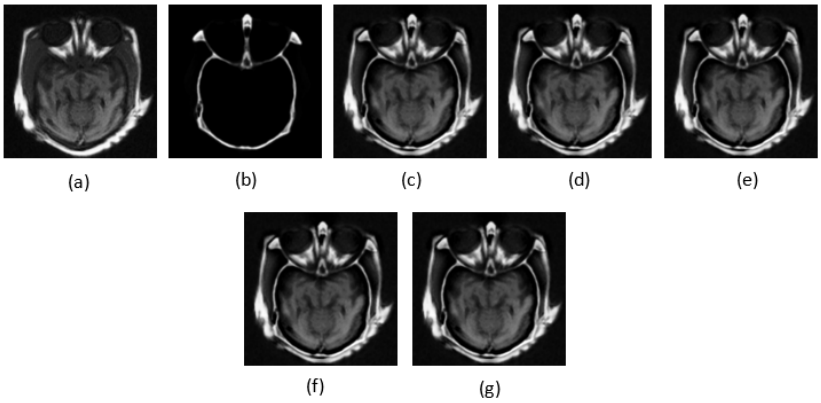

Figure 7.2: Example of image fusion (a)MR(b)CT are source images (c)(d)(e)(f)(g) are fusion result for LP-SR, DWR-SR, DTCWT-ST, CVT-SR, NSCT-SR respectivley

method	SD	MI	PSNR	Q_E	SSIM
LP-SR	60.7776	3.1260	42.7206	0.7682	0.7654
DWT_SR	51.7858	2.2327	39.6078	0.5645	0.5449
DTCWT_SR	53.5718	2.0315	43.1238	0.6159	0.7004
CVT_SR	54.4996	1.9097	43.5369	0.5713	0.6811
NSCT_SR	56.8796	3.0118	42.8384	0.7329	0.7637

Table 7.2: Evaluation metrics for different method for medical image

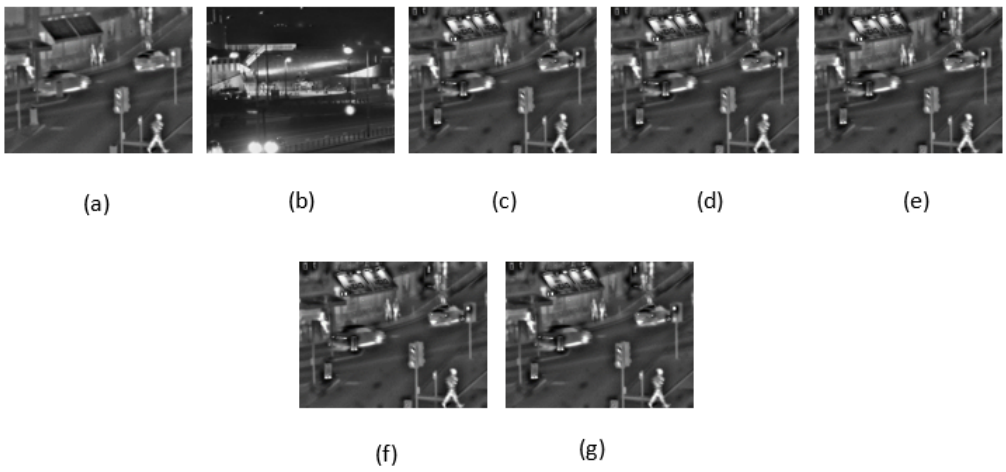

Figure 7.3: Example of image fusion (a)infrared image(b)visible image are source images (c)(d)(e)(f)(g) are fusion result for LP-SR, DWR-SR, DTCWT-ST, CVT-SR, NSCT-SR respectivley

method	SD	MI	PSNR	Q_g	SSIM
LP-SR	38.3217	2.2888	43.9515	0.6649	0.8529
DWT_SR	39.1248	2.5406	42.4839	0.6220	0.7867
DTCWT_SR	35.8982	2.7566	42.5670	0.6214	0.8637
CVT_SR	38.4855	1.5682	42.5671	0.5554	0.8151
NSCT_SR	36.9915	3.3075	43.9515	0.6423	0.8645

Table 7.3: Evaluation metrics of different method for infrared image

The evaluation metrics for different method for visible infrared image is given in table7.3

Since now we give the experimental result that gives without the noise.When we add noise then how the **PSNR** value varies with different values of **Sigma** is comparatively shown in fig7.4 for different **MST** method.I also put the value of psnr in Tabel.7.4 for multifocus image in Table7.5 for medical image and in Table**??** for infrared image.

method	Sigma=0	Sigma=10	Sigma=20	Sigma=30	Sigma=40
LP-SR	40.7547	32.1050	29.6455	28.8260	28.4305
DWT-SR	40.5786	32.1275	29.6915	28.8177	28.4180
CWT-SR	40.8904	32.7951	30.0288	29.0551	28.5810
CVT-SR	40.8469	32.6136	29.9358	28.9948	28.5211
NSCT-SR	40.9464	32.6726	29.9727	29.0035	28.5524

Table 7.4: PSNR variation for different value of sigma for different method for multifocus image

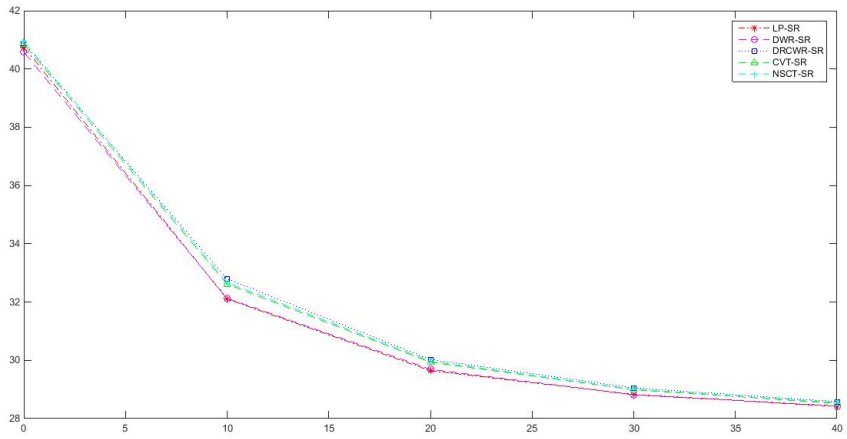

Figure 7.4: PSNR variation for different value of sigma for multifocus image of first image pair shown in figure 6.1

method	Sigma=0	Sigma=10	Sigma=20	Sigma=30	Sigma=40
LP-SR	42.7206	32.8634	30.3254	29.4772	29.0671
DWT-SR	39.6078	33.2129	30.6802	29.6775	29.2918
CWT-SR	43.1238	33.4460	30.6783	29.7806	29.2870
CVT-SR	43.5369	33.5767	30.7231	29.7802	29.2357
NSCT-SR	42.8384	33.3979	30.6726	29.7330	29.2600

Table 7.5: PSNR variation for different value of sigma for different method for medical image

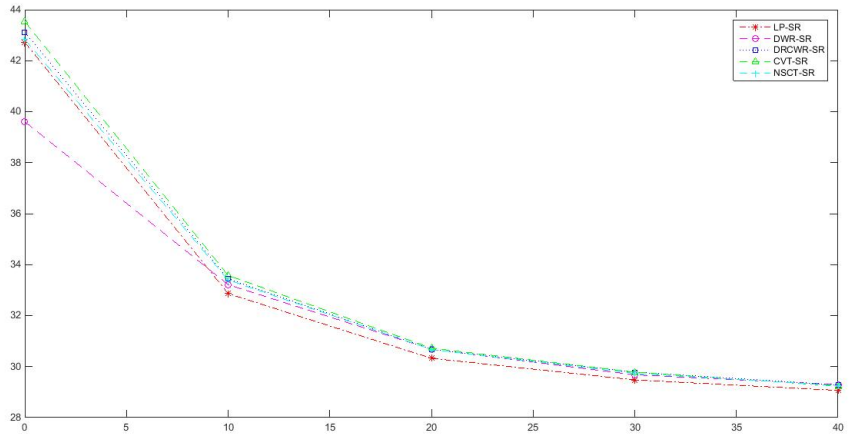

Figure 7.5: PSNR variation for different value of sigma for medical image of first medical image image pair shown in figure 6.1

7.2 Discussion

for each type of image fusion, I take the related contents in Tables 7.1-7.5 into consideration together and seek out some common regularities among the six MST s used in the proposed framework.

From the data obtained from table7.1-7.5 and graph7.4-7.5 we see that NSCT-SR based image fusion gives better result than other method.We perform this task on about 50 pair of image and this method gives the better result than any others method.

Chapter 8

Conclusion and Future plan

I proposed a general method for image fusion using MST and sparse representation.It gave us the result which overcome the disadvantage introduced in MST and SR method individually. Training time for Dictionary is relatively high.We use a dataset consisting different type of image and which contains about 40 images.For thresolding the image we use bayesShrik thresholding which take the threshold value in different direction of image under certain transformation domain, giving a good threshold result there for the image is noiseless in great measure. In the future, this model can be enhanced and cope with different efficient transformation domain which can take the geometry of image in better way. So it will be interesting to see how advancement is made in this approach.

Bibliography

[1] P. Burt, E. Adelson, The laplacian pyramid as a compact image code, IEEE Trans.Commun. 31 (4) (1983) 532–540..

[2] P. Burt, E. Adelson, The laplacian pyramid as a compact image code, IEEE Trans.Commun. 31 (4) (1983) 532–540

[3] A. Toet, Image fusion by a ratio of low pass pyramid, Pattern Recogn. Lett. 9(4)(1989) 245–253.

[4] V. Petrovic, C. Xydeas, Gradient-based multiresolution image fusion, IEEE Trans. Image Process. 13 (2) (2004) 228–237.

[5] H. Li, B. Manjunath, S. Mitra, Multisensor image fusion using the wavelet transform, Graph. Models Image Process. 57 (3) (1995) 235–245

[6] M. Beaulieu, S. Foucher, L. Gagnon, Multi-spectral image resolution refinement using stationary wavelet transform, in: Proceedings of 3rd IEEE International Geoscience and Remote Sensing Symposium, 2003, pp. 4032–4034

[7] J. Lewis, R. OCallaghan, S. Nikolov, D. Bull, N. Canagarajah, Pixel- and region-based image fusion with complex wavelets, Inform. Fusion 8 (2) (2007) 119–130.

[8] F. Nencini, A. Garzelli, S. Baronti, L. Alparone, Remote sensing image fusion using the curvelet transform, Inform. Fusion 8 (2) (2007) 143–156

[9] M. Aharon, M. Elad, and A. Bruckstein,"K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation,"IEEE Trans. Signal Process.54(11), 4311–4322(2006)

[10] T. Guha and R. K. Ward,"Learning sparse representations for human action recognition," IEEE Trans. Pattern Anal. Mach. Intell 34(8),1576–1588 (2012)

[11] M. S. Lewicki and T. J. Sejnowski,"Learning overcomplete representations,"Neural Comput.12(2), 337–365 (2000)

[12] K. Engan, S. O. Aase, and J. H. Husoy,"Method of optimal directions for frame design,"in Int. Conf. Audio, Speech and Signal Processing, Vol. 5, pp. 2443–2446, IEEE Press, Phoenix, AZ (1999).

[13] M. Yaghoobi, T. Blumensath, and M. E. Davies,"Dictionary learning for sparse approximations with the majorization method,"IEEE Trans. Signal Process.57(6), 2178–2191 (2009)

[14] L. Landweber,"An iterative formula for fredholm integral equations of the first kind,"Am. J. Math.73(3), 615–624 (1951)

[15] B. Yang and S. Li,"Multifocus image fusion and restoration with sparse representation,"IEEE Trans. Instrum. Meas. 59(4), 884–892 (2010).

[16] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse coding for natural images,"Nature, vol. 381, no. 6583, pp. 607–609, Jun. 1996

[17] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations,"Constr. Approx., vol. 13, no. 1, pp. 57–98, Mar. 1997.

[18] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,"IEEE Trans.Signal Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[19] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted norm minimization algorithm,"IEEE Trans. Signal Process., vol. 45, no. 3, pp. 600–616, Mar. 1997.

[20] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process.,vol. 15, no. 2, pp. 3736–3745, Dec. 2006.

[21] Iain M.Johnstone David L Donoho. Adapting to smoothness via wavelet shrinkage.Journal of the Statistical Association, 90(432):1200–1224, Dec 1995.

[22] David L Donoho. Ideal spatial adaptation by wavelet shrinkage. Biometrika,

[23] David L Donoho. De-noising by soft thresholding. IEEE Transactions on Information Theory, 41(3):613–627, May 1995.

[24] S. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Trans. Signal Process. 41 (12) (1993) 3397–3415.

[25] S. Li, B. Yang, J. Hu, Performance comparison of different multi-resolution transforms for image fusion, Inform. Fusion 12 (2) (2011) 74–84.