CHAPTER 2

LITERATURE SURVEY

Nowadays, with the rapid development in high-technology and modern instrumentations, medical imaging has become a fundamental component of a large number of applications, including diagnosis, research, and treatment. In order to support more accurate clinical information for physicians to deal with medical diagnosis and evaluation, multimodality medical images are needed, such as X-ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Magnetic Resonance Angiography (MRA), and Positron Emission Tomography (PET) images, etc (Maes et al 2003). These multimodality medical images usually provide complementary and occasionally contradictory information. For example, the CT image can provide dense structures like bones and implants with less distortion, but it cannot detect physiological changes, while the MRI image can provide normal and pathological soft tissues information, but it cannot support the bones information. In this case, only one kind of image may not be sufficient to provide accurate clinical requirements for the physicians. Therefore, the fusion of the multimodal medical images is necessary and it has become a promising and very challenging research area in recent years (Barra & Boire 2001, Zhu & Cochoff 2006).

Image fusion can be broadly defined as the process of combining multiple input images or some of their features into a single image without the introduction of distortion or loss of information (Petrovic & Zydes 2004). The aim of image fusion is to integrate complementary as well as redundant
information from multiple images to create a fused image output. So, the new image generated should contain a more accurate description of the scene than any of the individual source images and is more suitable for human visual and machine perception or further image processing and analysis tasks (Zhang & Lohmann 1999). For medical image fusion, the fusion of images can often lead to additional clinical information not apparent in the separate images. Another advantage is that it can decrease the storage cost by storing just the single fused image instead of multi-source images.

So far, many techniques for image fusion have been proposed in the literature and a thorough overview of these methods can be viewed in reference (Wang et al 2000). According to the stage at which the combination mechanism takes place, the image fusion methods can be generally grouped into three categories, namely, pixel level or sensor level, feature level, and decision level (Shivappa et al 2008). Since the pixel level fusion has the advantage that the images used contain the original measured quantities, and the algorithms are computationally efficient and easy to implement, the most image fusion applications occupy pixel level based methods (Redondo et al 2009).

The simplest way of image fusion is to take the average of the two images pixel by pixel. However, this method generally leads to undesirable side effect such as reduced contrast (Li & Yang 2008). More robust algorithm for pixel level fusion is the weighted average approach. In this method, the fused pixel is estimated as the weighted average of the corresponding input pixels. However, the weight estimation regularly requires a user-specific threshold. Other methods have been developed, such as intensity-hue-saturation (IHS), principal component analysis (PCA), and the Brovey transform. These techniques are easy to understand and implement. However, although the fused images obtained by these methods have high spatial
quality, they usually suffer from spectral degradation, i.e., they can yield high spatial resolution fused image, but they ignore the high quality of spectral information which is especially crucial for remote sensing image fusion (Pradhan et al 2006). Artificial neural network (ANN) has also been introduced to make image fusion, as seen in (Li et al 2002). However, the performance of ANN depends on the sample images and this is not an appealing characteristic. Yang et al. used a statistical approach to fuse the images (Yang et al 2002) however, in his method the distortion is modeled as a mixture of Gaussian probability density functions which is a limiting assumption. Due to the real-world objects usually contain structures at many different scales or resolutions and multiresolution or multiscale approaches can provide a means to exploit this fact, the multiresolution techniques have then attracted more and more interest in image fusion.

The multiresolution techniques occupy two kinds, one is pyramid transform another is wavelet transform. In the pyramid fusion, the input images are first transformed into their multiresolution pyramid representations. The fusion process then creates a new fused pyramid from the input image pyramids in a certain fusion rule. The fused image is finally reconstructed by performing an inverse multiresolution transform. Examples of this approach include the Laplacian pyramid (Burt & Adelson 1983), the gradient pyramid (Burt & Koleszynski 1993), the contrast pyramid (Toet 1989a), the ratio-of-low-pass pyramid (Toet 1989b) and the morphological pyramid (Toet 1989). However, for the reason of the pyramid method fails to introduce any spatial orientation selectivity in the decomposition process, the above mentioned methods often cause blocking effects in the fusion results (Li et al 1995). Matsopoulos et al. earlier applied the morphological pyramid method to fuse the MRI and CT images (Matsopoulos & Marshall 1995), but this method can occasionally create many undesired edges. Another family of the multiresolution fusion
techniques is the wavelet based method, which usually used the discrete wavelet transform (DWT) in the fusion. Since the DWT of image signals produces a nonredundant image representation, it can provide better spatial and spectral localization of image information as compared to other multiresolution representations. The research results reveal that DWT schemes have some advantages over pyramid schemes such as increased directional information, no blocking artifacts that often occur in pyramid fused images, better signal to-noise ratios and so on (Li et al. 2002).

Therefore, the wavelet based method has been popular widely used for image fusion (Li et al. 1995, Chipman et al. 1995, Zhang et al. 1999, Pu et al. 2000, Ma et al. 2005, Acerbi-Jr et al. 2006), and two detailed surveys can be seen in (Pajares et al. 2004, Amolins et al. 2007). Although there are considerable wavelet based fusion works today, most of them concerned on remote images, multifocus images, and infrared images, while fewer work has been done for medical images. Yu et al. (2001) fused the medical images by the wavelet based method with a maximum-selection fusion rule (Yu et al. 2001), which is similar to Burt’s method (Burt et al. 1993). However, this method suffers from the noise and artifacts as they tend to have higher contrast. Qu et al. used the modulus maxima selection criteria for the wavelet transform coefficients in the medical image fusion (Qu et al. 2001). The disadvantage of this method is that they consider only wavelet coefficients (pixel) values while making decisions about constructing the fused image (Garg et al. 2006).

CT image provides clear bones information but no soft tissues information, while contrast to CT image the MRI image provides clear soft tissues information but no bones information. That is to say, the same object in the two medical images appears very distinctly. Hence, when the two images are decomposed by wavelet transform, the approximation image (low
frequency band) and the detail image (high frequency bands) may have very
different physical meaning. Based on this and the above analysis, a new
fusion rule to perform the wavelet coefficients which treats the low frequency
band and high frequency bands with different fusion schemes separately. The
coefficients of low frequency band are chosen by a visibility based selection
scheme, while the coefficients of the high frequency bands are performed
with a maximum window-based variance selection scheme.

The fusion of multimodal medical images plays an important role
in many clinical applications for they can support more accurate information
than any individual source image.

The fusion of MRI image and CT image of the same organ is to
obtain a single image containing as much information as possible about that
organ for diagnosis. Algorithms such as the intensity, hue and saturation
(IHS) algorithm and the wavelet fusion algorithm have proved to be
perfectly registered images from multiple sources to produce a high quality
fused image with spatial and spectral information. Fused image can
significantly benefit medical diagnosis and also the further image processing
(Kiran kumar, 2009, Siyue Chen et al 2011). Wavelet transform the other
hand, has the ability to perform local analysis for revealing various aspects of
data like trends, breakdown points, discontinuities in higher derivatives, and
self similarities (Zhang-Shu Xiao & Chong-Xun Zheng 2009). A major
drawback of two-dimensional wavelets is their limited capability in capturing
directional information which has a significant role in analysis of the images,
including feature extraction and classification. To overcome this deficiency,
researchers have recently come up with a new family of wavelet methods that
can capture the intrinsic geometrical structures such as curvelet transform
(Kiran kumar, 2009) and contourlet transform curvelets are very successful in
detecting image activities along curves, while analyzing images at multiple scales, locations, and orientations. Several fusion algorithms have been proposed extending from the simple averaging to the curvelet transform.

Contourlet transform proposed by Do and Vetterli uses a structure similar to that of curvelets except at discrete domain. The contourlet expansion is composed of basis images oriented at various directions in multiple scales, with flexible.

The Ridgelet was introduced as a sparse expansion for functions on continuous spaces that are smooth away from discontinuities along lines (Minh N. Do & Martin Vetterli 2003). The Ridgelet Transform belongs to the family of discrete transforms employing basis functions. To facilitate its mathematical representation, it can be viewed as a wavelet analysis in the Radon domain. The Radon transform itself is a tool of shape detection. So, the Ridgelet Transform is primarily a tool of ridge detection or shape detection of the objects in an image (Mamatha & Gayatri 2012).

Contourlet transform proposed by Do and Vetterli is obtained by combining the laplacian pyramid with a directional filter bank. Contourlet transform provides a flexible multi-resolution, local and directional expansion for images. Contourlet transform Curvelets are very successful in detecting image activities along curves, while analyzing images at multiple scales, locations, and orientations (Minh N. Do et al 2005, Srinivasa rao et al 2007). Image fusion algorithms can be categorized into different levels: low, middle, and high; or pixel, feature, and Symbolic levels (Al-Azzawi et al 2009). There is several wavelet fusion rules, which can be used for the selection of the wavelet coefficients from the wavelet transform of the images to be fused. The frequently used rule is the maximum frequency rule which selects the coefficients that have the maximum absolute values (Mamatha & Gayatri 2012).
The multi-resolution and multidirectional characteristics ensure the transform a good ability in handling the abundant texture information in natural images. However, the contourlet transform is not shift invariant because of the sub sampled filter structure. Thus, the contourlet transform will cause the visual artifact in image processing applications. Hence, a geometric image transform constructed by combining wavelet transform and directional filter banks called Wavelet Based Contourlet Transform (WBCNT) is proposed (Ramin Eslami et al 2004).

Removing noise from the original signal is still a challenging problem for researchers. Digital images play an important role both in daily life applications such as satellite television, Magnetic Resonance Imaging (MRI), Computer Tomography (CT) as well as in areas of research and technology such as geographical information systems and astronomy. Data sets collected by image sensors are generally contaminated by noise. Imperfect instruments, problems with the data acquisition process, and interfering natural phenomena can all degrade the data of interest. Furthermore, noise can be introduced by transmission errors and compression. Thus, denoising is often a necessary and the first step to be taken before the images data is analyzed. It is necessary to apply an efficient denoising technique to compensate for such data corruption. Image denoising still remains a challenge for researchers because noise removal introduces artifacts and causes blurring of the images. Most of the natural images are assumed to have additive random noise which is modeled as a Gaussian. Speckle noise (Guo et al 1994) is observed in ultrasound images whereas Rician noise (Robert D. Nowak 1999) affects MRI images.

Image Denoising has remained a fundamental problem in the field of image processing. Wavelets give a better performance in image denoising due to properties such as sparsity and multiresolution structure. With Wavelet
Transform gaining popularity in the last two decades various algorithms for
denoising in wavelet domain were introduced. The focus was shifted from the
Spatial and Fourier domain to the Wavelet transform domain. Ever since
Donoho’s Wavelet based thresholding approach was published in 1995, there
was a surge in the denoising papers being published. Although Donoho’s
concept was not revolutionary, his methods did not require tracking or
correlation of the wavelet maxima and minima across the different scales as
proposed by Mallat. Thus, there was a renewed interest in wavelet based
denoising techniques since Donoho (Mallat et al 1992) demonstrated a simple
approach to a difficult problem. Researchers published different ways to
compute the parameters for the thresholding of wavelet coefficients. Data
adaptive thresholds (Imola K. Fodor et al 2003) were introduced to achieve
optimum value of threshold. Later efforts establish that substantial
improvements in perceptual quality could be obtained by translation invariant
methods based on thresholding of an Undecimated Wavelet Transform
(Coifman et al 1995). These thresholding techniques were applied to the non-
orthogonal wavelet coefficients to reduce artifacts. Multi wavelets were also
used to achieve similar results. Probabilistic models using the statistical
properties of the wavelet coefficient seemed to outperform the thresholding
techniques and gained ground.

There are two basic approaches to image denoising, spatial filtering
methods and transform domain filtering methods. A traditional way to remove
noise from image data is to employ spatial filters. Spatial filters can be further
classified into non-linear and linear filters. With non-linear filters, the noise is
removed without any attempts to explicitly identify it. Spatial filters use a low
pass filtering on groups of pixels with the assumption that the noise occupies
the higher region of frequency spectrum. Usually spatial filters remove noise
to a reasonable extent but at the cost of blurring images which in turn makes
the edges in pictures invisible. In recent years, a variety of nonlinear median-
type filters such as weighted median (Yang et al 1995), rank conditioned rank selection (Hardie et al 1994), and relaxed median (Ben Hamza et al 1999) have been developed to overcome this drawback. In linear filters the wiener filtering (Jain 1989) method requires the information about the spectra of the noise and the original signal and it works well only if the underlying signal is smooth. Wiener method implements spatial smoothing and its model complexity control correspond to choosing the window size. To overcome the weakness of the Wiener filtering, Donoho and Johnstone proposed the wavelet based denoising scheme in (David et al 1994, 1995).

The transform domain filtering methods can be subdivided according to the choice of the basis functions. The basis functions can be further classified as data adaptive and non-adaptive. Non-adaptive transforms are discussed first since they are more popular. Spatial-frequency filtering refers use of low pass filters using Fast Fourier Transform (FFT). In frequency smoothing methods (Jain 1989) the removal of the noise is achieved by designing a frequency domain filter and adapting a cut-off frequency when the noise components are decorrelated from the useful signal in the frequency domain. Filtering operations in the wavelet domain can be subdivided into linear and nonlinear methods. Linear filters such as Wiener filter in the wavelet domain yield optimal results when the signal corruption can be modeled as a Gaussian process and the accuracy criterion is the mean square error (MSE) (Choi et al 1998, Strela 2001).

The most investigated domain in denoising using Wavelet Transform is the non-linear coefficient thresholding based methods. The procedure in which small coefficients are removed while others are left untouched is called Hard Thresholding. Wavelet transform using soft thresholding was also introduced in (Donoho 1995). Similar to soft thresholding, other techniques of applying thresholds are semi-soft
thresholding and Garrote thresholding (Imola K. Fodor et al 2003). Most of the wavelet shrinkage literature is based on methods for choosing the optimal threshold which can be adaptive or non-adaptive to the image. VISUSHrink (David L. Donoho et al 1994) is non-adaptive universal threshold, which depends only on number of data points. It has asymptotic equivalence suggesting best performance in terms of MSE when the number of pixels reaches infinity. SUREShrink uses a hybrid of the universal threshold and the Stein’s Unbiased Risk Estimator (SURE) threshold and performs better than VISUSHrink. BayesShrink (Simoncelli et al 1996, Chipman et al 1997) minimizes the Bayes’ Risk Estimator function assuming Generalized Gaussian prior and thus yielding data adaptive threshold. BayesShrink outperforms SUREShrink most of the times. Cross Validation (Marteen Jansen 2000) replaces wavelet coefficient with the weighted average of neighborhood coefficients to minimize generalized cross validation (GCV) function providing optimum threshold for every coefficient.

In Non-orthogonal Wavelet Transform, Undecimated Wavelet Transform (UDWT) has also been used for decomposing the signal to provide visually enhanced solution. Since UDWT is shift invariant it avoids visual artifacts such as pseudo-Gibbs phenomenon. Though the improvement in results is much higher, use of UDWT adds a large overhead of computations thus making it less feasible. In (Lang et al 1995) normal hard/soft thresholding was extended to Shift Invariant Discrete Wavelet Transform. In (Cohen et al 1999) Shift Invariant Wavelet Packet Decomposition (SIWPD) is exploited to obtain number of basis functions. Wavelet Coefficient Model focuses on exploiting the multiresolution properties of Wavelet Transform. This technique identifies close correlation of signal at different resolutions by observing the signal across multiple resolutions. This method produces
excellent output but is computationally much more difficult and expensive. The modeling of the wavelet coefficients can either be deterministic or statistical.

The Deterministic method of modeling involves creating tree structure of wavelet coefficients with every level in the tree representing each scale of transformation and nodes representing the wavelet coefficients. This approach is adopted in (Baraniuk 1999), Lu et al (1992), tracked wavelet local maxima in scale-space, by using a tree structure. Other denoising method based on wavelet coefficient trees is proposed by Donoho (1997). A good review of statistical properties of wavelet coefficients can be found in Buccigrossi et al (1999) and Romberg et al (2001).

Without noise or distortion, the reconstruction from the transform expansion by both the contourlets and the discrete wavelets are perfect with zero MSE. However, due to the fading channel, the image recovery from the distorted coefficients may not be perfect.

The multi-resolution and multidirection characteristics ensure the transform a good ability in handling the abundant texture information in natural images. However, the contourlet is not shift invariant because of the sub sampled filter structure (Candes et al 1999). Thus, the contourlet will cause the visual artifact in image denoising applications. Hence, a geometric image transform constructed by combining wavelet transform and directional filter banks called Wavelet Based Contourlet Transform (WBCT).

The combination of good property of the ridgelet transform with the approximate shift invariant property of the wavelets can be made. The wavelet transform can be applied to the entire image or partitioning the image into a number of overlapping squares and thus applying the ridgelet transform to each square continued, so that the Wavelet Based Ridgelet Transform
(WBRT) can obtained. In the noised image, (Minh N. Do et al 2003) Wavelet transform can be used to get the information at the homogeneous areas and Curvelet transform is applied to extract the edge information which called Wavelet based Curvelet transform (WBCT).

The overhead of decompression is enormous. Today’s sophisticated algorithms need between 150 to 300 instructions per pixel for decompression. The processing of compression in compressed form reduces the amount of data that must be processed and avoids complex compression and decompression. Decreasing data volume has the side effect of increasing data locality and thus more efficiently uses processor cache, which further improves performance (Vetterli et al 2005). Therefore advanced techniques for the coding of the residual signal usually provide little additional compression as compared to traditional techniques, and additional complexity often does not justify this improvement. In recent years interest in multimedia has generated a lot of research in the area of image coding in academics and industry alike and several successful standards are involved. They address a wide range of applications having different requirements in terms of bit rate, picture quality, complexity, error resilience and delay as well as improved compression ratios, peak signal to noise ratio, root mean square error.

Compression improvement of up to fifty percent over the best previous standards is the primary motivation for advancing the new recommended techniques like wavelet based linking transforms. Wavelet based linking transform coding is more robust under transmission and decoding errors and also facilitates progressive transmission of images because of their inherent Multiresolution nature. Wavelet based linking coding schemes are especially suitable for applications where scalability and tolerance is needed.
The compression ratio obtainable from lossy compression can significantly exceed that obtainable from lossless compression. The primary tradeoff concerns requires for reproducibility versus the storage and transmission requirement. Lossy compression mainly consists of decorrelation and quantization stages that reduce the image size by permanently eliminating certain information. The decorrelation stage of the image compression algorithm is usually done by a transformation from one space to another space to facilitate compaction of information.

The Ridgelet Transform (Candes 1999) was developed over several years in an attempt to break an inherent limit plaguing wavelet denoising of images. This limit arises from the frequency depicted fact that the two-dimensional (2-D) Wavelet Transform of images exhibits large wavelet coefficients to represent the image edges. A basic model for calculating ridgelet coefficients is to view ridgelet analysis as a form of wavelet analysis in the radon domain. The ridgelet representation solves the problem of sparse approximation of smooth objects with straight edges. As a consequence the Curvelet Transform (Candes 1999) has been introduced. Curvelet transform is based on multiscale ridgelets combined with a spatial band pass filtering operation. Curvelet Transform was initially developed in the continuous-domain via multiscale filtering followed by a block Ridgelet Transform on each band pass image.

Do et al (2000), an attempt has been made to use Ridgelet Transform for image compression. However, in image processing, edges are typically curved rather than straight and ridgelet alone cannot yield efficient representation. But, if one uses a sufficient fine scale to capture curved edges, such an edge gets almost straight. Therefore ridgelets are deployed in a restricted manner.
Wavelet-based methods have expanded in the field of still image and video compression. They offer the advantage of a better tradeoff between complexity, compression and quality over the traditional Discrete Cosine Transform based methods. However, for image compression, Wavelet Transform has a problem with the orientation selectively because it provides local frequency representation of image regions over a range of spatial scales, and, therefore, it does not represent two dimensional singularities effectively. Various lossless and lossy image coding techniques have been developed (Sayood 2000). In a map of the large wavelet coefficients, one can see the edges of the images repeated at scale after scale. This effect means that many wavelet coefficients are required to reconstruct the edges in an image properly. Reducing the number of coefficients will introduce artifacts on the edges of the reconstructed image (Do et al 2000).

One approach is the use of multiresolution transforms that are free from blocking effect artifacts such as in case of the Discrete Cosine Transform, which is used in the JPEG (baseline) industry standard (Sayood 2000). By the use of the Wavelet Transform, the corresponding co-efficient of the different decomposition levels correlate and show a characteristic trend.

Later, a proposal was made for the second-generation Curvelet Transform (Candes 2002) that was defined via frequency partitioning without using Ridgelet Transform. Both curvelet constructions require a rotation operation for the frequency decomposition, which ensures the construction in the continuous-domain. For discrete images, sampled on a rectangular grid, the discrete implementation of the curvelet transform is very challenging.

During the past two decades, image compression has developed from a mostly academic Rate-Distortion field (Tunel et al 2003), into a highly commercial business. Therefore, a new image representation method was introduced known as the Contourlet transform (Do et al 2003). The
theory was started with a description of the transform in the discrete domain and then proves its convergence to an expansion in the continuous domain. Thus a discrete domain multiresolution and multidirectional expansion is constructed. This is the way through which wavelets are derived from filter banks, but using non-separable filter banks. Due to the fast-iterated filter bank algorithm the construction results in a flexible multiresolution, local and directional image expansion using contour segments. However, Contourlet Transform has the adverse property of showing other types of artifacts due to the discrete approach. This Research work investigates the assessment of the Contourlet Transform for image compression by a combining Contourlet Transform and Wavelet Transform.