Fusion of Multimodality Medical Images Using Combined Activity Level Measurement and Contourlet Transform

Sudeb Das[1] and Malay Kumar Kundu[2] *Senior Member, IEEE*
Machine Intelligence Unit, Indian Statistical Institute
203 B.T.Road, Kolkata-108, India
Email: [1]to.sudeb@gmail.com, [2]malay@isical.ac.in

***Abstract*—In this paper, we propose a novel multimodality Medical Image Fusion (MIF) method, based on a novel combined Activity Level Measurement (ALM) and Contourlet Transform (CNT) for spatially registered, multi-sensor, multi-resolution medical images. The source medical images are first decomposed by CNT. The low-frequency subbands (LFSs) are fused using the novel combined ALM, and the high-frequency subbands (HFSs) are fused according to their 'local average energy' of the neighborhood of coefficients. Then inverse contourlet transform (ICNT) is applied to the fused coefficients to get the fused image. The performance of the proposed scheme is evaluated by various quantitative measures like Mutual Information (MI), Spatial Frequency (SF), and Entropy (EN) etc. Visual and quantitative analysis and comparisons show the effectiveness of the proposed scheme in fusing multimodality medical images.**

Keywords-Image Fusion, Contourlet Transform, Multiscale Geometric Analysis, Mutual Information

I. INTRODUCTION

In recent years, with the rapid and significant development in technology and modern instrumentations, medical imaging is taking on an increasingly critical and vital role in a large number of healthcare applications including diagnosis, research, and treatment etc. An increasing number of medical image modalities have become available, to provide support to the physicians on their clinical diagnosis. Different modalities of medical imaging reflect different information of human organs and tissues, and have their respective application ranges. For instance, structural images like Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Ultrasonography (USG), Magnetic Resonance Angiography (MRA) etc. provide high-resolution images with anatomical information. On the other hand, functional images such as Position Emission Tomography (PET), Single-Photon Emission Computed Tomography (SPECT) and functional MRI (fMRI) etc. provide low-spatial resolution images with functional information. A single modality of medical image cannot provide comprehensive and accurate information. As a result, combining anatomical and functional medical images to provide much more useful information through image fusion (IF), has become the focus of imaging research and processing [1].

So far, many IF techniques have been proposed by various researchers in the literature. Generally, IF methods can be classified into three categories based on the merging state: pixel or sensor level, feature level, and decision level. It has been found, that the pixel-level spatial domain IF methods usually leads to contrast reduction. Methods based on Intensity-Hue-Saturation (IHS), Principal Component Analysis (PCA), and the Brovey Transform offers better results, but suffers from spectral degradation [2] [3]. Many Multiresolution Analysis (MRA) based IF methods have been proposed to improve the fusion result. Pyramidal IF schemes such that the Laplacian pyramid, the gradient pyramid, the contrast pyramid, the ratio-of-low-pass pyramid and the morphological pyramid etc. fails to introduce any spatial orientation selectivity in the decomposition process, and hence often causes blocking effects [4]. The other MRA-tool, which has been used in IF extensively, is the Discrete Wavelet Transform (DWT) [5]. The problem with Wavelet Transform (WT) is that, it can preserve spectral information efficiently but cannot express spatial characteristics well. The isotropic WT is scant of shift-invariance and multi-directionality, and fail to provide an optimal expression of highly anisotropic edges and contours in images. So, WT based fusion scheme cannot preserve the salient features of the source images efficiently, and will probably introduce some artifacts and inconsistency in the fused results [6].

Recently, a theory called Multi-scale Geometric Analysis (MGA) for high-dimensional signals has been developed, and several MGA tools were proposed like Ridgelet, Curvelet, Contourlet and Ripplet etc. These MGA tools do not suffer from the problems of wavelet. Few MIF methods based on these MGA tools were also proposed to improve the fusion result [6] [7]. However, the importance/contribution of each source images in the fused image is unequal, thus how to measure it and combine the corresponding coefficients have become the most important problem in MIF methods based on MRA and/or MGA-tools. To handle this inequality regarding the importance/contribution of each of the source images in the fused image, different activity level measurements (ALMs) have been applied in the IF paradigm.

In this paper, we propose a novel multimodality Medical Image Fusion (MIF) method, based on CNT using a new combined ALM. We focus on creating the novel fusion rule

978-1-61284-860-0/11/$26.00 ©2011 IEEE

by estimating the importance of the CNT coefficients in the image subbands. The LFSs are fused using the proposed novel ALM, and the HFSs are fused according to their `local average energy' of the neighborhood of coefficients in the HFSs. Both visual and quantitative performance evaluations are made and verified in the paper. Performance comparisons of the proposed method with some of the existing MIF schemes show that the proposed method performs better.

The rest of the paper is organized as follows. CNT is described in Section 2. Section 3, presents the proposed MIF algorithm. Experimental results and comparisons are given in Section 4, and we draw conclusion in Section 5.

II. CONTOURLET TRANSFORM

The major drawback of DWT in two dimensions is their limited ability in capturing directional information. In light of this, Do and Vetterli [8] developed the CNT, based on an efficient two-dimensional multiscale and directional filter bank (DBF). CNT not only possess the main features of DWT, but also offer a high degree of directionality and anisotropy. It allows for different and flexible number of directions at each scale, while achieving nearly critical sampling. In addition, CNT uses iterated filter banks, which makes it computationally efficient (*O(N)* operations for an *N*-pixels image) [8].

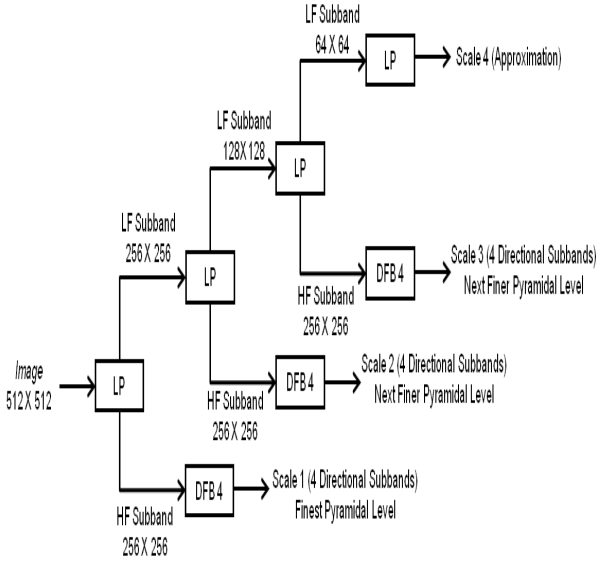


Figure 1. Flowchart of CNT for a 512 X 512 image.

CNT gives a multiresolution, local and directional expansion of image using Pyramidal Directional Filter Bank (PDFB). The PDFB combines Laplacian Pyramid (LP) which captures the point discontinuities, with a DFB which links these discontinuities into linear structures. Fig. 1, shows the flowchart of CNT for a 512 × 512 image. As shown in Fig. 1, first stage of CNT is LP decomposition and DFB is the second stage.

LP scheme is shown in Fig. 2. Here, the input image *x* is first lowpass filtered by analysis filter *H* and then down-sampled to produce a coarse approximation *a*. It is then interpolated and passed through the synthesis filter *G*. The resulting image is subtracted from the original image *x* to obtain the bandpass image *b*. This process can be iterated

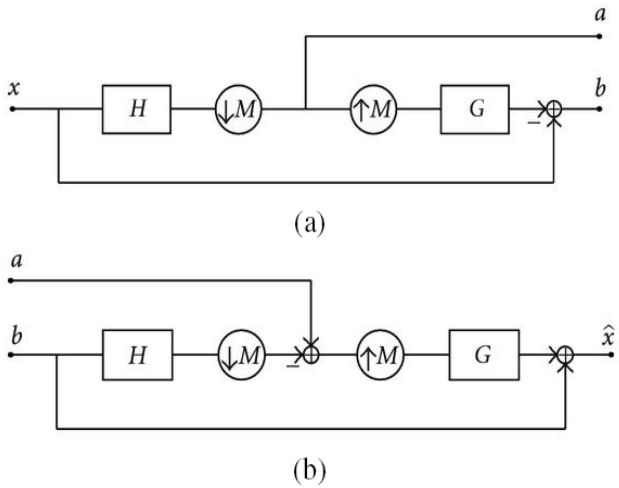


Figure 2. Laplacian pyramid scheme: (a) analysis and (b) synthesis.

repeatedly on the coarser version of the image *a*. LP is a multiscale decomposition of $L^2(R^2)$ into series of increasing resolution subspaces which are orthogonal complements of each other as follows [9]:

$$L^2(R^2) = V_{j_0} \oplus \left(\overset{-\infty}{\underset{j=J_0}{\oplus}} W_j\right) \tag{1}$$

where V_{j0} is the approximation subspace at the scale 2^{j0}, W_j is the detail in the finer scale 2^{j-1}. In the LP, each subspace W_j is spanned by a frame $\mu_{j,n}(t)_{n \in Z^2}$ that utilizes a uniform grid on R^2 of intervals $2^{j-1} \times 2^{j-1}$.

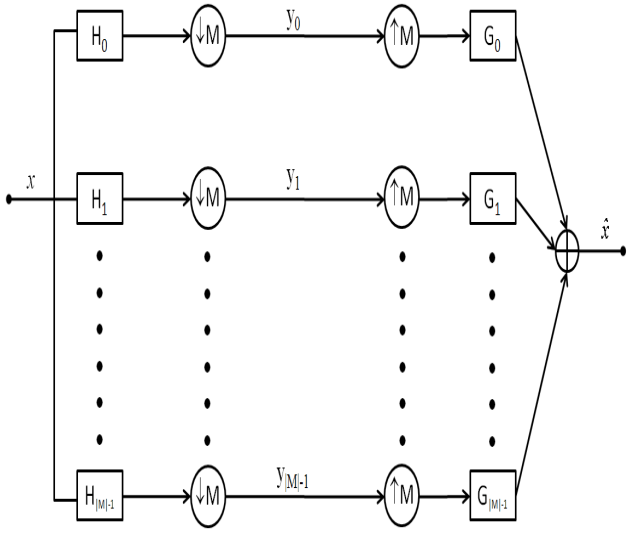


Figure 3. Construction of DBF.

In 1992, Bamberger and Smith constructed a 2D-DFB that can be maximally decimated while achieving perfect

reconstruction [10]. It is used in the second stage of CNT to link the edge points into linear structures, which involves modulating the input image and using quincunx filter banks (QFB) with diamond-shaped filters. An l-level tree-structured DFB is equivalent to a 2^l parallel channel filter bank with equivalent filters and overall sampling matrices as shown in Fig. 3. As shown in Fig. 3, corresponding to the subbands indexed, the equivalent analysis and synthesis filters are denoted using H_k and G_k, $0 \le k < 2^m$.

An l-level DFB generates a perfect directional basis for discrete signal in $l^2(Z^2)$ that is composed of the impulse responses of 2^l directional synthesis filters and their shift. They can be represented as follows:

$$g_k^{(l)}[n - S_k^{(l)} n]_{0 \le k < 2^l, n \in Z^2} \tag{2}$$

$$g[n] = \frac{2\pi}{n_1}[\psi(\frac{n_1(l+1)}{N} + n_2) - \psi(\frac{n_1 l}{N} + n_2)] \tag{3}$$

where $N = 2^{n-2}$ and $\psi(x)$ is similar to the common *sin* function.

$$\psi(x) = \frac{1 - \cos(\pi x)}{\pi x} \tag{4}$$

In CNT, applying an l_j-level DBF to the detail subspace W_j results in a decomposition with 2^{l_j} directional subspaces as follows:

$$W_j = \bigoplus_{k=0}^{2^{l_j}-1} W_{j,k}^{l_j} \tag{5}$$

A DFB is designed to capture the high frequency content like smooth contours and directional edges. Fig. 4, shows the frequency partition of CNT.

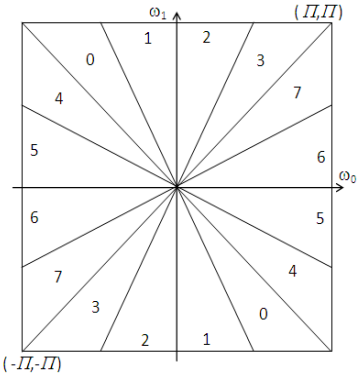


Figure 4. Frequency partition of CNT.

III. Proposed Method

The notations used in this section are as follows: A, B , F represents the two source images and the fused image, respectively. $T = (A,B,F)$. L^T_J indicates the LFS of the image T at the coarsest scale J. $H^T_{j,k}$ represents the HFS of the image T at scale j, (j = 1,....,J) and direction k . (m,n) denotes the spatial location of each coefficient. The method can be easily extended to more than two images.

A. Fusing Low Frequency Subband

To fuse the LFSs coefficients we have used a novel fusion rule based on a combined ALM, by estimating each coefficient contribution to the fused image. If "a coefficient is large" and "its local information entropy is large", then the coefficient may contribute more importance to the fused image. Considering $f_1(x(m,n))$ and $f_2(x(m,n))$ as the two simple ALMs representing "a coefficient is large", and "its local information entropy is large", respectively, we have created, $f(x(m,n))$ as the combined ALM representing "a coefficient is large and its local information entropy is large", which represents the importance of the coefficient to the fused image.

For a coefficient $x(m,n) \epsilon L^A_J$, $f_1(x(m,n))$ is defined as follows:

$$f_1(x(m,n)) = \frac{|x(m,n)|}{\max_{x(p,q) \in L_J^A} (|x(p,q)|)} \tag{6}$$

Similarly, $f_2(x(m,n))$ is defined as follows:

$$f_2(x(m,n)) = \frac{I(x(m,n))}{\min_{x(p,q) \in L_J^A} (I(x(p,q)))} \tag{7}$$

where, $I(x(m,n))$ denotes the 'local information entropy' of the coefficient $x(m,n)$. The 'local information entropy' of a coefficient is calculated considering a $M \times N$ neighborhood around the coefficient $x(m,n)$, and is defined as follows:

$$I(x(m,n)) = -\sum p(x(m,n)) \log_2 p(x(m,n)) \tag{8}$$

where, $p(x(m,n))$ is the probability of occurrence of the coefficient $x(m,n)$. We, then compute $f(x(m,n))$ as follows:

$$f(x(m,n)) = \min(f_1(x(m,n)), f_2(x(m,n))) \tag{9}$$

Similarly, for a coefficient $y(m,n) \epsilon L^B_J$, (j = 1, ..., J) we compute $f_1(y(m,n))$, $f_2(y(m,n))$ and $f(y(m,n))$. To get a fused coefficient $z(m,n)$ of the subband L^F_J we follow the given rule:

$$z(m,n) = \frac{(x(m,n) \bullet f(x(m,n))) + (y(m,n) \bullet f(y(m,n)))}{f(x(m,n)) + f(y(m,n))} \tag{10}$$

B. Fusing High Fequency Subband

The HFSs coefficients are fused according to their 'local average energy' of the neighborhood of the coefficient under consideration. For a coefficient $x(m,n) \epsilon H^A_{j,k}$, ($j$ = 1, ..., J); its' 'local average energy' $E_A(x(m,n))$ is computed as follows:

$$E_A(x(m,n)) = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} x(m,n)^2 \tag{11}$$

where, $M \times N$ is the size of the neighborhood around the coefficient $x(m,n)$.

Similarly for a coefficient $y(m,n) \in H^B_{j,k}$, $(j = 1, ..., J)$; its' 'local average energy' $E_B(y(m,n))$ is computed. To get a fused coefficient $z(m,n)$ of the subband $H^F_{j,k}$ we follow the following rule:

$$z(m,n) = \frac{(E_A(x(m,n)) \bullet x(m,n)) + (E_B(y(m,n)) \bullet y(m,n))}{E_A(x(m,n)) + E_B(y(m,n))} \quad (12)$$

C. Algorithm

The medical images to be fused must be registered to assure that the corresponding pixels are aligned. Here we outline the salient steps of the proposed MIF method:

1) Decompose the registered source medical images A and B by CNT to get the LFSs and HFSs.
2) Fused the coefficients of LFSs using the combined ALM described in Section III-A, to get the fused LFS.
3) Similarly to get the fused HFSs, fused the HFSs of the images A and B, using fusion rule described in Section III-B.
4) Apply inverse contourlet transform on the fused LFS and HFSs to get the final fused medical image.

The block diagram of the proposed MIF scheme is shown in Figure 5.

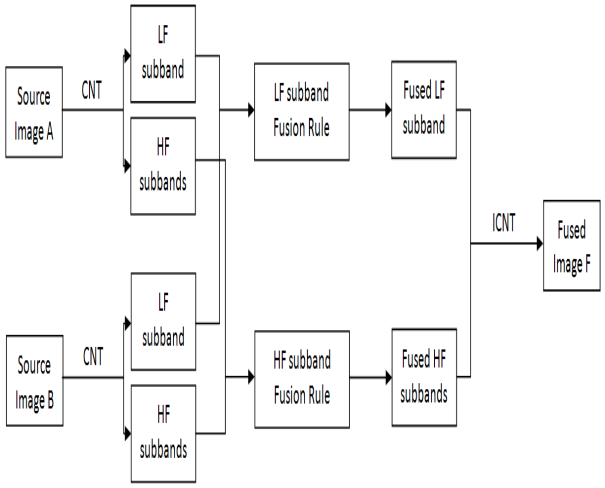


Figure 5. Block Diagram of the proposed MIF method.

IV. Experimental Results and Comparisons

To evaluate the performance of the proposed MIF method, extensive experiments were carried out on various modalities of medical images. Fig. 6(a)-(b) and Fig. 6(c)-(d) shows two different sets of source images used in the experiments, and are denoted by *IS1* and *IS2*, respectively. The CT image in Fig. 6(a) shows the bones and the MRI image in Fig. 6(b) displays the soft tissues information. The T1-weighted MR image in Fig. 6(c) of *IS2* contains the soft tissues but no illness information, and the MRA image in Fig. 6(d) shows the illness information (shown by the marked ellipse) but no soft tissues information. The decomposition parameter of CNT was *levels* = [1, 2, 4, 4]. To show the effectiveness of the proposed technique, visual as well as quantitative analysis were carried out. The selected quantitative criterions used in the experiments are as follows:

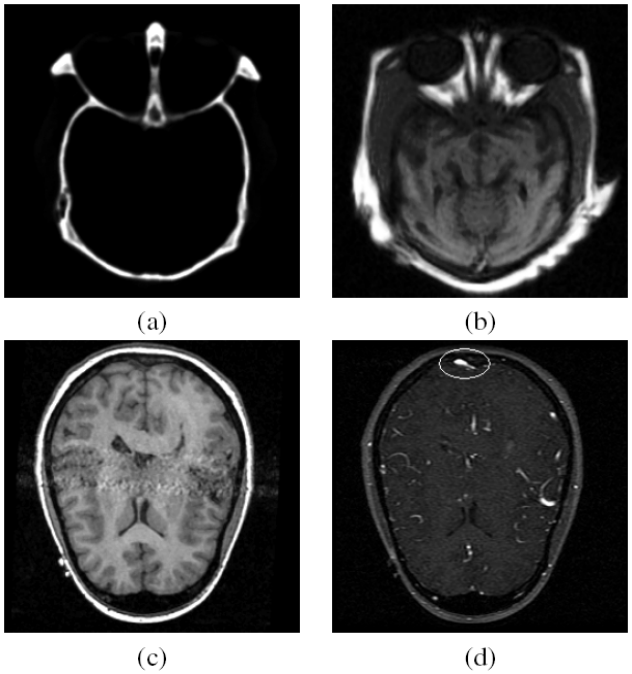


Figure 6. Source Images: (a) CT image (SF= 4:43, EN= 1:71, STD= 44:75); (b) MRI image (SF= 6:26, EN= 5:66, STD= 58:84); (c) T1-weighted MR image (SF= 7:70, EN= 4:15, STD= 69:20); (d) MRA image (SF= 6:49, EN= 4:33, STD= 25:58). ((a)-(b)) first set IS1 and ((c)-(d)) second set IS2 of images, respectively (downloaded from http://www.imagefusion.org/).

A. Standard Deviation (STD)

It measures the contrast in the fused image. An image with high contrast would have a high standard deviation.

$$STD = \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(F(m,n) - MEAN)^2} \quad (13)$$

where $M \times N$ denotes the size of the image and $F(m,n)$ indicates the gray-value of the pixel of image F at position (m,n) and

$$MEAN = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}|F(m,n)| \quad (14)$$

B. Entropy (EN)

The entropy of an image is a measure of information content. It is the average number of bits needed to quantize the intensities in the image. It is defined as

$$EN = -\sum_{g=0}^{L-1} p(g)\log_2 p(g) \quad (15)$$

where $p(g)$ is the probability of grey-level g , and the range of g is [0,.....,L-1]. An image with high information content would have high entropy. If entropy of fused image is higher than parent images then it indicates that the fused image contains more information.

C. Spatial Frequency (SF)

Spatial frequency can be used to measure the overall activity and clarity level of an image. Larger SF value denotes better fusion result:

$$SF = \sqrt{RF^2 + CF^2} \quad (16)$$

where RF is the row frequency and CF is the column frequency:

$$RF = \sqrt{\frac{1}{M(N-1)} \sum_{m=0}^{M-1} \sum_{n=0}^{N-2} (F(m,n+1) - F(m,n))^2} \quad (17)$$

and

$$CF = \sqrt{\frac{1}{(M-1)N} \sum_{m=0}^{M-2} \sum_{n=0}^{N-1} (F(m+1,n) - F(m,n))^2} \quad (18)$$

where $M \times N$ denotes the size of the image and $F(m,n)$ indicates the gray-value of the pixel of image F at position (m,n).

D. Mutual Information (MI)

It measures the degree of dependence of the two images. A larger measure implies better quality. Given two images x_F and x_R; MI is defined as [11]:

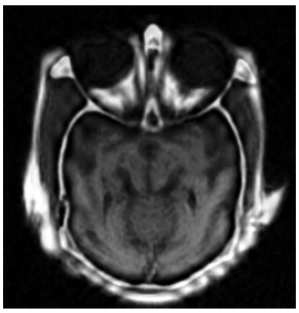
$$MI = I(x_A; x_F) + I(x_B; x_F) \quad (19)$$

where

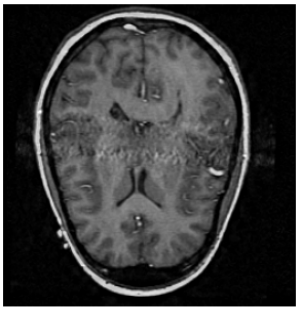
$$I(x_R; x_F) = \sum_{u=1}^{L} \sum_{v=1}^{L} h_{R,F}(u,v) \log_2 \frac{h_{R,F}(u,v)}{h_R(u) h_R(v)} \quad (20)$$

where h_R, h_F are the normalized gray level histograms of x_R and x_F, respectively. $h_{R,F}$ is the joint gray level histogram of x_R and x_F, and L is the number of bins. x_R and x_F correspond to the reference and fused images, respectively. $I(x_R;x_F)$ indicates how much information the fused image x_F conveys about the reference x_R. Thus, the higher the mutual information between x_F and x_R, the more likely x_F resembles the ideal x_R.

An expert clinician was asked to subjectively evaluate the effectiveness of the proposed MIF method. After careful manual inspection, the clinician conformed to the effectiveness of the proposed scheme. He found that the fused images obtained by the proposed MIF scheme, were clearer, informative and have higher contrast which is helpful in visualization as well as interpretation.



(a)



(b)

Figure 7. Results of the proposed MIF scheme (a) fused image of IS1; (b) fused image of IS2

Fig. 7, shows the resultant fused images obtained by our proposed MIF method for the two test sets of images *IS1* and *IS2*. It can be easily seen from the Fig. 7, that the fused images have much more salient features and detailed information than the source images of Fig. 6.

TABLE I: PERFORMANCE COMPARISONS USING *IS1*

Scheme	MI	SF	EN	STD
Scheme [6]	—	6.56	6.39	53.82
Scheme [5]	2.71	—	6.73	57.98
Scheme [12]	2.06	—	4.98	—
Our Scheme	**3.25**	**6.95**	**6.76**	**60.86**

The Table I, shows the performance comparisons of our method against some of the existing schemes using *IS1*. In Table I, '—' denotes the values which are not available. For the second set of test images *IS2* the values of the different quantitative measures are MI=3.75, SF=7.42, EN=6.30 and STD=53.95. The '**bold**' values indicate the highest values in the Table I. The higher value of SF indicates that the fused images obtained by our proposed method, have more activity and clarity level than the source images. Similarly the higher values of EN and STD for the fused images show that the fused images obtained by the proposed scheme, have more information, as well as higher contrast than the source images. So, it is clear from Table I that, the fused

images obtained by the proposed MIF method based on CNT are clearer, more informative and have higher contrast which is helpful in visualization and interpretation.

V. CONCLUSION

We propose a novel multimodality MIF method based on contourlet transform using a new combined ALM based fusion rule. The CNT is capable of resolving two dimensional singularities and representing image edges more efficiently, which makes the fused images clearer and more informative. To integrate as much information as possible into the fused images the new combined ALM based fusion rule is developed in such a way, which overcomes the drawbacks of traditional IF schemes. The proposed MIF method is analyzed both visually and quantitatively, and is compared with several existing IF techniques and the superiority of the proposed scheme is established. Experimental results show that the proposed method can preserve more useful information in the fused image with higher spatial resolution and less difference to the source images.

REFERENCES

[1] V. Barra and J. Y. Boire, "A general framework for the fusion of anatomical and functional medical images," NeuroImage, vol. 13, no. 3, pp. 410–424, 2001.

[2] J. Yonghong, "Fusion of landsat TM and SAR image based on principal component analysis," Remote Sensing Technology and Application, vol. 13, no. 1, pp. 46–49, 1998.

[3] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," in Proc. of Image Vision Computing, vol. 26, no. 7, 2008, pp. 971–979.

[4] H. Li, B. S. Manjunath, and S. K. Mitra, "Multi-sensor image fusion using the wavelet transform," in Proc. of CVGIP: Graphical Model and Image Processing, vol. 57, no. 3, 1995, pp. 235–245.

[5] Y. Yang, D. S. Park, S. Huang, and N. Rao, "Medical image fusion via an effective wavelet-based approach," EURASIP Journal on Advances in Signal Processing, vol. 2010, pp. 44:1–44:13, 2010.

[6] L. Yang, B. L. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," Neurocomputing, vol. 72, no. 1-3, pp. 203–211, 2008.

[7] S. Das, M. Chowdhury, and M. K. Kundu, "Medical image fusion based on ripplet transform type-I," Progress In Electromagnetics Research B, vol. 30, pp. 355–370, 2011.

[8] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," IEEE Transactions on Image Processing, vol. 14, no. 12, pp. 2091–2106, 2005.

[9] P. J.Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," IEEE Transactions on Communications, vol. 31, pp. 532–540, 1983.

[10] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," IEEE Transactions on Signal Processing, vol. 40, no. 4, pp. 882–893, 1992.

[11] G. H. Qu, D. L. Zhang, and P. F. Yan, "Information measure for performance of image fusion," Electronic Letters, vol. 38, no. 7, pp. 313–315, 2002.

[12] H. Tian, Y.-N. Fu, and P.-G. Wang, "Image fusion algorithm based on regional variance and multi-wavelet bases," in Proc. of 2nd Int. Conf. Future Computer and Communication, vol. 2, 2010, pp. 792–795.