Ripplet Based Multimodality Medical Image Fusion Using Pulse-Coupled Neural Network and Modified Spatial Frequency

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, a novel multimodality Medical Image Fusion (MIF) method, based on Ripplet Transform Type-I (RT) using Pulse-Coupled Neural Network (PCNN) is presented. The proposed MIF scheme exploits the advantages of both RT and PCNN to obtain better results. The source medical images are first decomposed by discrete RT (DRT). The low-frequency subbands (LFSs) are fused using the 'max selection' rule. For the fusion of high-frequency subbands (HFSs) a PCNN model is utilized. Modified Spatial Frequency (MSF) in DRT domain is input to motivate the PCNN and coefficients in DRT domain with large firing times are selected as coefficients of the fused image. Then inverse DRT (IDRT) 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.**

***Index Terms*—Image Fusion, Ripplet Transform, Multi-scale Geometric Analysis, PCNN, Mutual Information**

I. Introduction

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, that has been used in IF extensively is the Discrete Wavelet Transform (DWT) [5], [6]. 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 multidirectionality, 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 [7].

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 and Contourlet etc. These MGA tools does not suffer from the problems of wavelet. Few MIF methods based on these MGA tools were also proposed to improve the fusion result [7], [8]. We have used RT in the proposed method, because it is capable of resolving two dimensional ($2D$) singularities and representing image edges more efficiently. 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.

PCNN is a visual cortex-inspired neural network and characterized by the global coupling and pulse synchronization of neurons [9], [10]. PCNN and it's modified versions have been used in the IF paradigm by various researchers [11], [12], [13], [14]. The results of various research works show that PCNN outperforms the other conventional image fusion methods. However, in most of the existing IF methods based

978-1-4577-0792-6/11/$26.00 ©2011 IEEE

on PCNN, the value of single pixel (coefficient) in spatial or transform domain is used to motivate one neuron. But a simple use of pixels (coefficients) in spatial or transform domain is not effective enough because humans are sensitive to edges and directional features. So, in this proposed method modified spatial frequency (MSF) in the DRT domain is used to motivate the PCNN. The LFSs are fused using the 'max selection' rule, and the HFSs are fused using the PCNN and MSF. 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. RT is described in Section 2. In Section 3, the standard PCNN model is briefly reviewed. Section 4, presents the proposed MIF algorithm. Experimental results and comparisons are given in Section 5, and we draw conclusion in Section 6.

II. Ripplet Transform Type-I (RT)

Conventional transforms like FT and WT suffer from discontinuities such as edges and contours in images. To address this problem, Jun Xu et al. proposed a new MGA-tool called RT. RT is a higher dimensional generalization of the Curvelet Transform (CVT), capable of representing images or $2D$ signals at different scales and different directions. To achieve anisotropic directionality, CVT uses a parabolic scaling law. From the perspective of microlocal analysis, the anisotropic property of CVT guarantees resolving $2D$ singularities along C^2 curves. On the other hand, RT provides a new tight frame with sparse representation for images with discontinuities along C^d curves [15].

There are two questions regarding the scaling law used in CVT: 1) Is the parabolic scaling law optimal for all types of boundaries? and if not, 2) What scaling law will be optimal? To address these questions, Jun Xu et al. intended to generalize the scaling law, which resulted in RT. RT generalizes CVT by adding two parameters, i.e., support c and degree d. CVT is just a special case of RT with $c = 1$ and $d = 2$. The anisotropy capability of representing singularities along arbitrarily shaped curves of RT, is due to these new parameters c and d.

A. Continuous Ripplet Transform (CRT)

For a 2D integrable function $f(\overrightarrow{x})$, the CRT is defined as the inner product of $f(\overrightarrow{x})$ and ripplets $\rho_{a\,\overrightarrow{b}\,\theta}(\overrightarrow{x})$ as given below [15]

$$R(a,\,\overrightarrow{b},\theta) = \langle f, \rho_{a\,\overrightarrow{b}\,\theta}\rangle = \int f(\overrightarrow{x})\overline{\rho_{a\,\overrightarrow{b}\,\theta}(\overrightarrow{x})}d\overrightarrow{x} \quad (1)$$

where $R(a,\,\overrightarrow{b},\theta)$ are the ripplet coefficients and $\overline{(.)}$ denotes the conjugate operator. The ripplet function of the Eq.(1), is defined as

$$\rho_{a\,\overrightarrow{b}\,\theta}(\overrightarrow{x}) = \rho_{a\,\overrightarrow{0}\,0}(R_\theta(\overrightarrow{x} - \overrightarrow{b})) \quad (2)$$

where $\rho_{a\,\overrightarrow{0}\,0}(\overrightarrow{x})$ is the ripplet element function, $R_\theta = \begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix}$ is the rotation matrix, $\overrightarrow{x}$ and $\overrightarrow{b}$ are $2D$ vectors; b and θ denotes the position parameter and rotation parameter respectively. The element ripplet function is defined in frequency domain as

$$\widehat{\rho}_a(r,\omega) = \frac{1}{\sqrt{c}}a^{\frac{1+d}{2d}}W(a\cdot r)V(\frac{a^{\frac{1}{d}}}{c\cdot a}\cdot\omega) \quad (3)$$

where $\widehat{\rho}_a(r,\omega)$ are the fourier transform (FT) of $\rho_{a\,\overrightarrow{0}\,0}(\overrightarrow{x})$ in polar coordinate system, and a is the scale parameter. $W(r)$ is the 'radial-window' and $V(\omega)$ is the 'angular window'. These two windows have compact supports on $[\frac{1}{2}, 2]$ and $[-1, 1]$ respectively. They satisfy the following admissibility conditions:

$$\int_{\frac{1}{2}}^{2} W^2(r)\frac{dr}{r} = 1 \quad (4)$$

$$\int_{-1}^{1} V^2(t)dt = 1 \quad (5)$$

These two windows partition the polar frequency domain into 'wedges' as shown in Fig. 1(a).

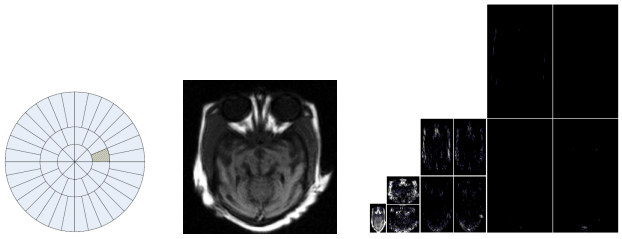
The ripplets functions decay very fast outside the elliptical effective region, which has the following property for its length and width: $width \approx length^d$. Length and width is the major and minor axis of the ellipse respectively. The customizable effective region tuned by support c and degree d bespeaks the most distinctive property of ripplets the general scaling. For $c = 1$, $d = 1$, both axis directions are scaled in the same way. So ripplet with $d = 1$ will not have the anisotropic behavior. For $d > 1$, the anisotropic property is reserved for RT. For $d = 2$, ripplets have parabolic scaling; for $d = 3$, ripplets have cubic scaling; and so forth.

The CRT can only capture the characteristics of high frequency components of $f(\overrightarrow{x})$, since the scale parameter a cannot take the value of infinity. So the 'full' CRT consists of fine scale RT and coarse scale isotropic WT. We can perfectly reconstruct the input function based on its ripplet coefficients. If $f(\overrightarrow{x}) \in L^2$ is a high-pass function i.e., its FT vanishes for $|\omega| < \frac{2}{a_0}$ and a_0 is a constant, then $f(\overrightarrow{x})$ can be reproduced by its RT through

$$f(\overrightarrow{x}) = \int R(a,\,\overrightarrow{b},\theta)\rho_{a\,\overrightarrow{b}\,\theta}(\overrightarrow{x})da\,d\overrightarrow{b}\,d\theta/a^3 \quad (6)$$

and a Parseval formula for f holds

$$||f||_{L_2}^2 = \int |R(a,\,\overrightarrow{b},\theta)|^2 da\,d\overrightarrow{b}\,d\theta/a^3 \quad (7)$$



(a) (b) (c)

Fig. 1. (a) The tiling of polar frequency domain. The shadowed 'wedge' corresponds to the frequency transform of the element function. Ripplet Transform: (b) MRI image (c) Different subbands after decomposition.

B. Discrete Ripplet Transform (DRT)

As digital image processing needs discrete transform instead of continuous transform, here we describe the discretization of RT [15].

The discretization of CRT is based on the discretization of the parameters of ripplet functions. a is sampled at dyadic

intervals. b and θ are sampled at equal-spaced intervals. a_j, $\overrightarrow{b}_k$ and θ_l substitute a, $\overrightarrow{b}$ and θ respectively, and satisfy that $a_j = 2^{-j}$, $\overrightarrow{b}_k = [c \cdot 2^{-j} \cdot k_1, 2^{-j/d} \cdot k_2]^T$ and $\theta_l = \frac{2\Pi}{c} \cdot 2^{-\lfloor j(1-1/d) \rfloor} \cdot l$, where $\overrightarrow{k} = [k_1, k_2]^T$, and $j, k_1, k_2, l \in \mathbb{Z}$. $(\cdot)^T$ denotes the transpose of a vector. $d \in \mathbb{R}$, since any real number can be approximated by rational numbers, so we can represent d with $d = n/m$, $n, m \neq 0 \in \mathbb{Z}$. Usually, we prefer $n, m \in \mathbf{N}$ and n, m are both primes. In the frequency domain, the corresponding frequency response of ripplet function is in the form

$$\widehat{\rho}_j(r,\omega) = \frac{1}{\sqrt{c}} a^{\frac{m+n}{2n}} W(2^{-j} \cdot r) V(\frac{1}{c} \cdot 2^{-\lfloor j \frac{m-n}{n} \rfloor} \cdot \omega - l) \quad (8)$$

where W and V satisfy the following admissibility conditions:

$$\sum_{j=0}^{+\infty} |W(2^{-j} \cdot r)|^2 = 1 \quad (9)$$

$$\sum_{l=-\infty}^{+\infty} |V(\frac{1}{c} \cdot 2^{-\lfloor j(1-1/d) \rfloor} \cdot \omega - l)|^2 = 1 \quad (10)$$

given c,d and j.

The 'wedge' corresponding to the ripplet function in the frequency domain is

$$H_{j,l}(r,\theta) = \{2^j \leq |r| \leq 2^{2j}, |\theta - \frac{\pi}{c} \cdot 2^{-\lfloor j(1-1/d) \rfloor} \cdot l| \leq \frac{\pi}{2} 2^{-j}\} \quad (11)$$

The DRT of an $M \times N$ image $f(n_1, n_2)$ will be in the form of

$$R_{j,\overrightarrow{k},l} = \sum_{n_1=0}^{M-1} \sum_{n_2=0}^{N-1} f(n_1, n_2) \overline{\rho_{j,\overrightarrow{k},l}(n_1, n_2)} \quad (12)$$

where $R_{j,\overrightarrow{k},l}$ are the ripplet coefficients.

The image can be reconstructed through Inverse Discrete Ripplet Transform (IDRT)

$$\widetilde{f}(n_1, n_2) = \sum_j \sum_{\overrightarrow{k}} \sum_l R_{j,\overrightarrow{k},l} \rho_{j,\overrightarrow{k},l}(n_1, n_2) \quad (13)$$

The Fig. 1(b), and Fig. 1(c), shows an MRI image and the subbands of the ripplet transformed MRI image after decomposition, respectively.

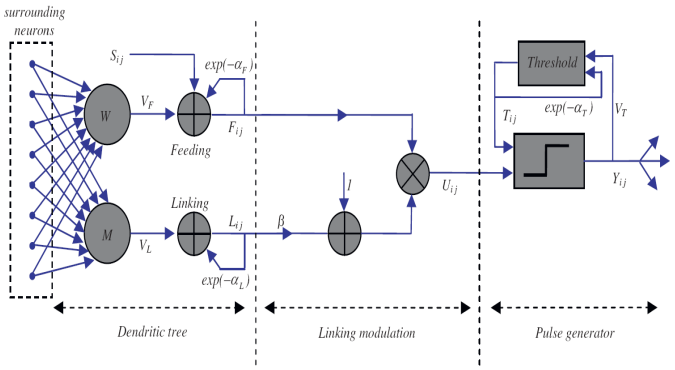


Fig. 2. Structure of PCNN.

III. Pulse Coupled Neural Network

PCNN is a single layered, two-dimensional, laterally connected neural network of pulse coupled neurons. The PCNN neurons structure is shown in Fig. 2. The neuron consists of an input part (dendritic tree), linking part and a pulse generator. The neuron receives the input signals from feeding and linking inputs. Feeding input is the primary input from the neurons receptive area. The neuron receptive area consists of the neighboring pixels of corresponding pixel in the input image. Linking input is the secondary input of lateral connections with neighboring neurons. The difference between these inputs is that the feeding connections have a slower characteristic response time constant than the linking connections. The standard PCNN model is described as iteration by the following equations [10], [16]:

$$F_{i,j}[n] = e^{-\alpha_F} F_{i,j}[n-1] + V_F \sum_{k,l} w_{i,j,k,l} Y_{i,j}[n-1] + S_{i,j} \quad (14)$$

$$L_{i,j}[n] = e^{-\alpha_L} L_{i,j}[n-1] + V_L \sum_{k,l} m_{i,j,k,l} Y_{i,j}[n-1] \quad (15)$$

$$U_{i,j}[n] = F_{i,j}[n](1 + \beta L_{i,j}[n]) \quad (16)$$

$$Y_{i,j}[n] = \begin{cases} 1, & U_{i,j}[n] > T_{i,j}[n] \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

$$T_{i,j}[n] = e^{-\alpha_T} T_{i,j}[n-1] + V_T Y_{i,j}[n] \quad (18)$$

In the Eq.(14) to Eq.(18), the indexes i and j refer to the pixel location in the image, k and l refer to the dislocation in a symmetric neighborhood around one pixel, and n denotes the current iteration (discrete time step). Here n varies from 1 to N (total number of iterations). The dendritic tree is given by Eqs.(14)-(15). The two main components F and L are called feeding and linking, respectively. $w_{i,j,k,l}$ and $m_{i,j,k,l}$ are the synaptic weight coefficients and S is the external stimulus. V_F and V_L are normalizing constants. α_F and α_L are the time constants; generally, $\alpha_F < \alpha_L$. The linking modulation is given in Eq.(16), where $U_{i,j}[n]$ is the internal state of the neuron and β is the linking parameter. The pulse generator determines the firing events in the model in Eq.(17). $Y_{i,j}[n]$ depends on the internal state and threshold. The dynamic threshold of the neuron is Eq.(18), where V_T and α_T are normalized constant and time constant, respectively.

IV. Proposed Method

The notations used in this section are as follows: A, B, R represents the two source images and the resultant fused image, respectively. $C = (A, B, R)$. L_G^C indicates the LFS of the image C at the coarsest scale G. $D_{g,h}^C$ represents the HFS of the image C at scale g, $(g = 1, ..., G)$ and direction h. (i, j) denotes the spatial location of each coefficient. The method can be easily extended to more than two images.

A. Fusing Low Frequency Subbands

The LFSs coefficients are fused using 'max selection' rule. According to this fusion rule, select the frequency coefficients from L_G^A or L_G^B with greater absolute value as the fused coefficients:

$$L_G^R(i,j) = \begin{cases} L_G^A(i,j), & |L_G^A(i,j)| \geq |L_G^B(i,j)| \\ L_G^B(i,j), & \text{otherwise} \end{cases} \quad (19)$$

B. Fusing High Frequency Subbands

The HFSs of the source images are fused using PCNN. As humans are sensitive to features such as edges, contours etc., so instead of using PCNN in DRT domain directly (i.e., using individual coefficients), modified spatial frequency (MSF) in DRT domain is considered as the image feature to motivate the PCNN.

1) Modified Spatial Frequency (MSF) : Spatial frequency (SF) proposed by Eskicioglu et al. is calculated by row and column frequency [17]. It reflects the whole activity level of an image, which means the larger the SF, the higher the image resolution. We have used a modified version of SF in the proposed MIF method. The MSF consists of row (RF), column (CF) and diagonal frequency (DF). For a $M \times N$ pixels image f the MSF is defined as

$$MSF = \sqrt{RF^2 + CF^2 + DF^2} \tag{20}$$

where,

$$RF = \sqrt{\frac{1}{M(N-1)} \sum_{m=1}^{M} \sum_{n=2}^{N} [f_{m,n} - f_{m,n-1}]^2} \tag{21}$$

$$CF = \sqrt{\frac{1}{(M-1)N} \sum_{m=2}^{M} \sum_{n=1}^{N} [f_{m,n} - f_{m-1,n}]^2} \tag{22}$$

and,

$$DF = P + Q \tag{23}$$

where,

$$P = \sqrt{\frac{1}{(M-1)(N-1)} \sum_{m=2}^{M} \sum_{n=2}^{N} [f_{m,n} - f_{m-1,n-1}]^2} \tag{24}$$

and,

$$Q = \sqrt{\frac{1}{(M-1)(N-1)} \sum_{m=2}^{M} \sum_{n=2}^{N} [f_{m-1,n} - f_{m,n-1}]^2} \tag{25}$$

2) Fusion Using DRT-MSF-PCNN: Let, $MSF_{i,j}^{g,h,C}$ be the modified spatial frequency corresponding to a coefficient $D_{g,h}^{C}(i,j)$, measured by using an overlapping window around the concerned coefficient where $C = (A, B)$. In order to reduce the computational complexity, we use a simplified PCNN:

$$F_{i,j}^{g,h,C}[n] = MSF_{i,j}^{g,h,C} \tag{26}$$

$$L_{i,j}^{g,h,C}[n] = e^{-\alpha_L} L_{i,j}^{g,h,C}[n-1] + V_L \sum_{k,l} W_{i,j,k,l}^{g,h,C} \; Y_{i,j,k,l}^{g,h,C}[n-1] \tag{27}$$

$$U_{i,j}^{g,h,C}[n] = F_{i,j}^{g,h,C}[n] * (1 + \beta L_{i,j}^{g,h,C}[n]) \tag{28}$$

$$\theta_{i,j}^{g,h,C}[n] = e^{-\alpha_\theta} \theta_{i,j}^{g,h,C}[n-1] + V_\theta Y_{i,j}^{g,h,C}[n-1] \tag{29}$$

$$Y_{i,j}^{g,h,C}[n] = \begin{cases} 1, & U_{i,j}^{g,h,C}[n] > \theta_{i,j}^{g,h,C}[n] \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

$$T_{i,j}^{g,h,C}[n] = T_{i,j}^{g,h,C}[n-1] + Y_{i,j}^{g,h,C}[n] \tag{31}$$

where, the feeding input $F_{i,j}^{g,h,C}$ is equal to the normalized modified spatial frequency $MSF_{i,j}^{g,h,C}$. The linking input $L_{i,j}^{g,h,C}$ is equal to the sum of neurons firing times in linking range. $W_{i,j,k,l}$ is the synaptic gain strength and subscripts k and l are the size of linking range in the PCNN. α_L is the decay constants. β is the linking strength, V_L and V_θ are the amplitude gains. $U_{i,j}^{g,h,C}$ is the total internal activity and $\theta_{i,j}^{g,h,C}$ is the threshold. If $U_{i,j}^{g,h,C}$ is larger than $\theta_{i,j}^{g,h,C}$, then the neuron will generate a pulse $Y_{i,j}^{g,h,C} = 1$, also called one firing time. The sum of $Y_{i,j}^{g,h,C} = 1$ in n iteration (namely the firing times), is used to represent the image information. Here, rather than $Y_{i,j}^{g,h,C}[n]$, we have analyzed $T_{i,j}^{g,h,C}[n]$, since neighboring coefficients with similar features represent similar firing times in a given iteration times.

C. Algorithm

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

1) Decompose the registered source medical images A and B by DRT to get the LFSs and HFSs.
2) Fused the coefficients of LFSs using the 'max selection' rule described in Section IV-A, to get the fused LFS.
3) Compute the MSF as described in Section IV-B1, using overlapping window on the coefficients in HFSs.
4) Input MSF of each HFSs to motivate the PCNN and generate pulse of neurons with Eqs.(26)–(30). and compute the firing times $T_{i,j}^{g,h,C}[n]$ by Eq.(31).
5) If $n = N$, then iteration stops. Then fuse the coefficients in the HFSs by the following fusion rule:

$$D_{g,h}^{R}(i,j) = \begin{cases} D_{g,h}^{A}(i,j), & T_{i,j}^{g,h,A}[N] \geq T_{i,j}^{g,h,B}[N] \\ D_{g,h}^{B}(i,j), & \text{otherwisee} \end{cases} \tag{32}$$

6) Apply inverse ripplet transform (IDRT) on the fused LFS and HFSs to get the final fused medical image.

The block diagram of the proposed MIF scheme is shown in Fig. 3.

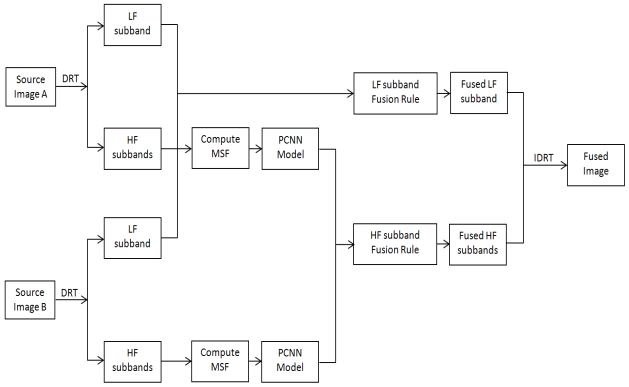


Fig. 3. Block Diagram of the proposed MIF method.

V. 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. 4(a)-(b) and Fig. 4(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. 4(a) shows the bones and the MRI image in Fig. 4(b) displays the soft tissues information. The T1-weighted MR image in Fig. 4(c) of $IS2$ contains the soft tissues but no illness information, and the MRA image in Fig. 4(d) shows the illness information (shown by the marked ellipse) but no soft tissues information. The decomposition parameter of DRT was $levels = [1, 2, 4, 4]$. Parameters of PCNN was set as $k \times l = 3 \times 3$, $\alpha_L = 0.06931$, $\alpha_\theta = 0.2$, $\beta = 0.2$, $V_L = 1.0$, $V_\theta = 20$, $W = \begin{bmatrix} 0.707 & 1 & 0.707 \\ 1 & 0 & 1 \\ 0.707 & 1 & 0.707 \end{bmatrix}$, and $N = 200$.

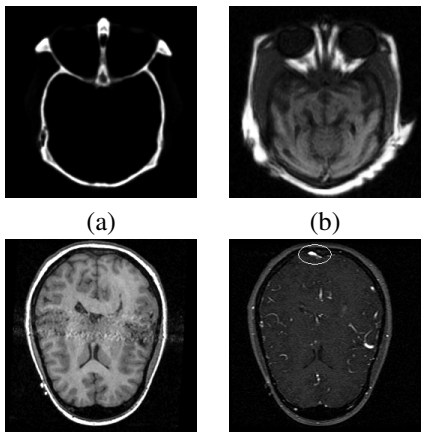


Fig. 4. Source Images: (a) CT image (SF= 4.43, EN= 1.71, STD= 44.75); (b) MRI image (SF= 6.26, EN= 5.66, STD= 58.84); (c) T1-weighted MR image (SF= 7.70, EN= 4.15, STD= 69.20); (d) MRA image (SF= 6.49, EN= 4.33, STD= 25.58). ((a)-(b)) first set $IS1$ and ((c)-(d)) second set $IS2$ of images, respectively (downloaded from http://www.imagefusion.org/).

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:

A. Standard Deviation (STD)

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

$$STD = \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(F(m,n) - MEAN)^2}, \quad (33)$$

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

$$MEAN = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}|F(m,n)|, \quad (34)$$

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

$$EN = -\sum_{g=0}^{L-1} p(g)\log_2 p(g), \quad (35)$$

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:

$$SF = \sqrt{RF^2 + CF^2}, \quad (36)$$

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

$$RF = \sqrt{\frac{1}{M(N-1)}\sum_{m=1}^{M}\sum_{n=2}^{N}(F(m,n) - F(m,n-1))^2}, \quad (37)$$

and

$$CF = \sqrt{\frac{1}{(M-1)N}\sum_{m=2}^{M}\sum_{n=1}^{N}(F(m,n) - F(m-1,n))^2}, \quad (38)$$

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 [18]:

$$MI = I(x_A; x_F) + I(x_B; x_F), \quad (39)$$

where

$$I(x_R; x_F) = \sum_{u=1}^{L}\sum_{v=1}^{L} h_{R,F}(u,v) log_2 \frac{h_{R,F}(u,v)}{h_R(u)h_F(v)}, \quad (40)$$

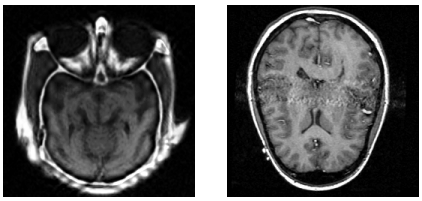
where h_R, h_F are the normalized gray level histograms of x_R and x_F, respectively. $h_{R,F}$ is the joint gray level histogram of x_R and x_F, and L is the number of bins. x_R and x_F correspond to the reference and fused images, respectively. $I(x_R; x_F)$ indicates how much information the fused 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

TABLE I
PERFORMANCE COMPARISONS USING $IS1$

Scheme	MI	SF	EN	STD
Scheme [7]	—	6.56	6.39	53.82
Scheme [5]	2.71	—	6.73	57.98
Scheme [6]	2.06	—	4.98	—
Our Scheme	**3.12**	**7.07**	**6.81**	**64.13**

the proposed scheme. He found that the fused images obtained by the proposed MIF scheme, were more clear, informative and have higher contrast which is helpful in visualization as well as interpretation.



(a) (b)

Fig. 5. Results of the proposed MIF scheme (a) fused image of $IS1$; (b) fused image of $IS2$

Fig. 5, 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. 5, that the fused images have much more salient features and detailed information than the source images of Fig. 4.

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= 4.19, SF= 8.12, EN= 6.46 and STD= 68.99. 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 more clear, informative and have higher contrast which is helpful in visualization and interpretation.

VI. CONCLUSION

We propose a novel multimodality MIF method based on ripplet transform using modified spatial frequency motivated PCNN. The DRT 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 low frequency source subbands are fused using 'max selection' rule, and PCNN is used to select the 'better' coefficients from the decomposed source high frequency subbands. To improve the result, instead of using single coefficient to motivate the PCNN, modified spatial frequency is used as the image feature to motivate the PCNN. 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.

ACKNOWLEDGMENT

We would like to thank Jun Xu and Depeng Wu (Dept. of Electrical and Computer Engineering, University of Florida, USA) for helping us in the implementation of Ripplet Transform.

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] 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.

[7] 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.

[8] 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.

[9] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, "Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex," *Neural Computation*, vol. 2, pp. 293–307, 1990.

[10] Z. Wang, Y. Ma, F. Cheng, and L. Yang, "Review of pulse-coupled neural networks," *Image Vision Comput.*, vol. 28, no. 1, pp. 5–13, 2010.

[11] Q. Liu, L. ping Xu, Y. Wang, Y. de Ma, and Q. Xie, "A novel algorithm of image fusion based on adaptive ULPCNN time matrix," in *Information Engineering (ICIE), 2010 WASE International Conference on*, vol. 1, aug. 2010, pp. 198 –202.

[12] Q. Xiao-Bo, Y. Jing-Wen, X. Hong-Zhi, and Z. Zi-Qian, "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain," *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1508–1514, 2008.

[13] Z. Wang, Y. Ma, and J. Gu, "Multi-focus image fusion using PCNN," *Pattern Recognition*, vol. 43, no. 6, pp. 2003–2016, June 2010.

[14] Z. Wang and Y. Ma, "Medical image fusion using m-PCNN," *Information Fusion*, vol. 9, no. 2, pp. 176–185, April 2008.

[15] J. Xu, L. Yang, and D. Wu, "Ripplet: A new transform for image processing," *Journal of Visual Communication and Image Representation*, vol. 21, no. 7, pp. 627–639, 2010.

[16] J. Johnson and M. Padgett, "PCNN models and applications," *Neural Networks, IEEE Transactions on*, vol. 10, no. 3, pp. 480 –498, may 1999.

[17] A. Eskicioglu and P. Fisher, "Image quality measures and their performance," *Communications, IEEE Transactions on*, vol. 43, no. 12, pp. 2959 –2965, Dec 1995.

[18] 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.