

A Neuro-Fuzzy Approach for Medical Image Fusion

Sudeb Das*, and Malay Kumar Kundu, *Senior Member, IEEE*,

Abstract—This paper addresses a novel approach to the multi-sensor, multimodal medical image fusion (MIF) problem, employing multiscale geometric analysis of non-subsampled contourlet transform and fuzzy-adaptive reduced pulse-coupled neural network (RPCNN). The linking strengths of the RPCNNs' neurons are adaptively set by modelling them as the fuzzy membership values, representing their significance in the corresponding source image. Use of RPCNN with less complex structure and having less number of parameters, leads to computational efficiency, an important requirement of point-of-care (POC) health care technologies. The proposed scheme is free from the common shortcomings of the state-of-the-art MIF techniques: contrast reduction, loss of image fine details and unwanted image degradations etc. Subjective and objective evaluations show better performance of this new approach compared to existing techniques.

Index Terms—Image fusion, artificial neural network, fuzzy logic, medical imaging, image analysis.

I. INTRODUCTION

THE rapid and significant advancements in medical imaging technologies and sensors, lead to new uses of medical images in various healthcare and bio-medical applications including diagnosis, research, treatment and education etc. Different modalities of medical images 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) and magnetic resonance angiography (MRA) etc., provide high-resolution images with excellent anatomical detail and precise localization capability. Whereas, 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, useful for detecting cancer and related metabolic abnormalities. A single modality of medical image cannot provide comprehensive and accurate information. Therefore, it is necessary to correlate one modality of medical image to another to obtain the relevant information. Moreover, the manual process of integrating several modalities of medical images is rigorous, time-consuming, costly, subject to human error, and requires years of experience. Therefore, automatically combining multimodal medical images through image fusion (IF) has become the focus of imaging research and processing [1], [2].

In recent years, many IF and medical image fusion (MIF) techniques have been proposed by various researchers. It has been found that the pixel-level spatial domain methods usually

lead to contrast reduction. Approaches based on intensity-hue-saturation (IHS), principal component analysis (PCA), and the Brovey transform offer better results, but suffer from spectral degradation. Pyramidal IF schemes fail to introduce any spatial orientation selectivity in the decomposition process, and hence often cause blocking effects [3], [4]. Most of the above mentioned schemes are modality specific with its own limitations [5]. The problem with widely used Wavelet transform (WT) is that it can preserve spectral information efficiently but cannot express spatial characteristics well [6]–[8]. As a result, WT based fusion schemes fail to preserve the salient features of the source images efficiently, but introduce artifacts and inconsistencies in the fused results [9]. Recently, to overcome these problems, many improved IF/MIF methods based on multiscale geometric analysis (MGA) tools (like curvelet, contourlet, ripplelet etc.) are proposed [9], [10]. However, measuring the importance/contribution of individual source image in the fused image, and finding effective way of combining them is still an open problem.

Pulse-coupled neural network (PCNN) is a novel neural network model that simulates the processing mechanism of cat's visual cortex and characterized by global coupling and pulse synchronization of neurons [11]–[13]. PCNN and its modified versions have been used in the IF/MIF domains by various researchers [3], [5], [14], [15]. Even though, PCNN based IF/MIF methods outperform other conventional schemes, but these approaches also suffer from several shortcomings: contrast reduction, loss of image fine details and unwanted degradations etc. Moreover, PCNN has several parameters with complex structures, and optimal estimation of these parameters is a major limitation to automatization and generalization of PCNN. In most of the PCNN based IF/MIF techniques, these parameters are kept same and set as a constant. But, according to human visual system (HVS), the responses to a region with notable features are stronger than a region with non-notable features. Thus, the parameters of PCNN's neurons should be related to the importance (significance) of the features of either the corresponding pixel (in spatial domain) or coefficient (in transform domain) of the image [16]–[18]. But, the problem remains: how to measure the importance (significance) of the pixel (coefficient) in the corresponding image. Therefore, we not only need a way to adaptively and automatically set the values of the parameters of the PCNN, but also to make the fusion scheme free from the common problems faced by the existing techniques [19]. This paper explores the use of fuzzy logic for building a simultaneous fusion and enhancement technique based on the HVS model to address these above mentioned problems. In this regard, the main contributions of this article are as follows: (1) A novel MIF scheme employing non-subsampled contourlet transform (NSCT) and reduced PCNN (RPCNN) with adaptive

Manuscript received Month XX, 2013; revised Month XX, 2013. Asterisk indicate corresponding author.

*S. Das and M. K. Kundu are with the Machine Intelligence Unit, Indian Statistical Institute, 203 B.T.Road, Kolkata-108, India e-mail: (see to.sudeb@gmail.com, malay@isical.ac.in).

linking strengths based on the corresponding image's local features. (2) Following the subjectivity of HVS, fuzzy logic is used to enable the proposed scheme to produce high quality fused images with higher contrast, more clarity and more useful subtle detail information. (3) Without involving any prior training and trials the less complex RPCNN having less number of parameters leads to computational efficiency, which is suitable for real time image processing applications (like point-of-care (POC) health care technologies). Both subjective and objective performance evaluations are made and verified in the paper. Performance comparisons with state-of-the-art schemes show that the proposed method performs better.

The rest of the paper is organized as follows. NSCT is described in Section II. In Section III, the RPCNN model is briefly reviewed. Section IV, presents the proposed MIF scheme. Experimental results and comparisons are given in Section V, and we draw conclusion in Section VI.

II. NON-SUBSAMPLED CONTOURLET TRANSFORM

NSCT is a fully shift-invariant, multiscale, and multi-direction expansion with fast implementability [20]. NSCT achieves the shift-invariance property (not present in contourlet transform (CNT)) by using non-subsampled pyramid filter bank (NSP) and non-subsampled DFB (NSDFB).

A. Non-subsampled Pyramid Filter Bank (NSP)

NSP ensures the multiscale property of NSCT, and has no down-sampling or up-sampling, and hence is shift-invariant. It is constructed by iterated two channel non-subsampled filter bank, and one low-frequency and one high-frequency image can be produced at each NSP decomposition level. The subsequent NSP decomposition stages are carried out to decompose the low-frequency component available iteratively to capture the singularities in the image. As a result, NSP can result in $k + 1$ sub-images, which consist of one low-frequency image and k high-frequency images whose sizes are all the same as the source image; k denotes the number of decomposition levels.

B. Non-subsampled Directional Filter Bank (NSDFB)

The NSDFB is constructed by eliminating the downsamplers and upsamplers of the DFB [20]. This results in a tree composed of two-channel non-subsampled filter banks (NSFBs). The NSDFB allows direction decomposition with l stages in high-frequency images from NSP at each scale, and produces 2^l directional sub-images with the same size as the source image. Thus, NSDFB enables NSCT with the multi-direction property and provides more precise directional detail information. The outputs of the first level and second level filters are combined to get the directional frequency decompositions. The synthesis filter bank is obtained similarly. All filter banks in the NSDFB tree structure are obtained from a single NSFB with fan filters. To obtain multi directional decomposition, the NSDFBs are iterated, and to get the next level decomposition, all filters are up sampled by a quincunx matrix given by $QM = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

NSCT is obtained by combining the NSP and the NSDFB as described by the Fig. 1(a). The resulting filtering structure

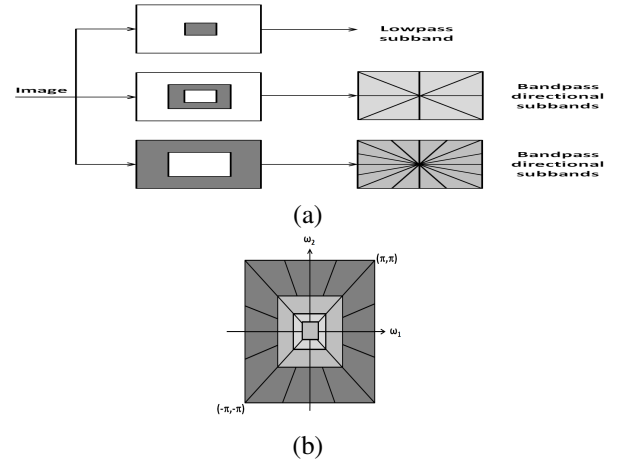


Fig. 1: NSCT (a) NSFB structure that implements the NSCT. (b) Frequency partitioning obtained with the proposed structure.

approximates the ideal partition of the frequency plane displayed in Fig. 1(b). It must be noted that different from the contourlet expansion, the NSCT has a redundancy given by $R = \sum_{j=0}^l 2^{2j}$, where, 2^{2j} is the number of directions at scale j . For further information on NSCT see [20].

In the proposed method the decomposition parameter of NSCT is set at $levels = [1, 2, 2]$ and we use 'pyrex' and 'vk' as the pyramid filter and orientation filter, respectively. With this decomposition configuration, the number of subband images obtained is 11. Each obtained subband is of the same size as the source medical image. This helps in finding the relationship between different subbands, which is good for designing the fusion rule and helpful for avoiding the pseudo-Gibbs phenomenon. Moreover, NSCT has better frequency selectivity and regularity than the other MGA tools and is capable of capturing the fine details present in the image. Furthermore, NSCT provides a sparse representation of signals and structurally conforms to the frequency sensitivity distribution of the HVS. These facts motivate us to utilize NSCT to develop our proposed MIF scheme.

III. REDUCED PULSE COUPLED NEURAL NETWORK

PCNN is a single layered, $2D$, laterally connected neural network of pulse coupled neurons [12], [21]. The neuron consists of an input part (dendritic tree), linking part and a pulse generator as shown in Fig. 2. The neuron receives the input signals from feeding and linking inputs. Considering the applications of multimodal MIF, and in order to reduce the computational complexity, we use a RPCNN adapted slightly from [22]:

$$F_{i,j}[n] = S_{i,j} \quad (1)$$

$$L_{i,j}[n] = \sum_{k,l} W_{i,j,k,l} Y_{i,j}[n-1] \quad (2)$$

$$U_{i,j}[n] = F_{i,j}[n](1 + \beta L_{i,j}[n]) \quad (3)$$

$$Y_{i,j}[n] = \begin{cases} 1, & U_{i,j}[n] > T_{i,j}[n-1] \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$T_{i,j}[n] = e^{-\alpha T} T_{i,j}[n-1] + V_T Y_{i,j}[n] \quad (5)$$

where, the indexes i and j refer to the pixel (coefficient) location in the image. k and l refer to the dislocation in a

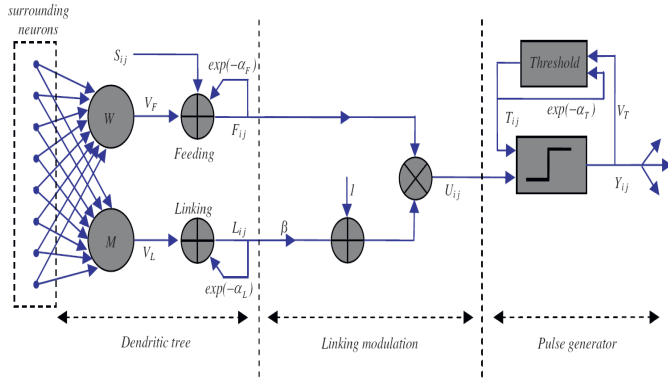


Fig. 2: Structure of PCNN.

symmetric neighborhood around one pixel (coefficient), and n denotes the current iteration. F and L are called feeding and linking input, respectively. $W_{i,j,k,l}$ represents the synaptic weight coefficient and $S_{i,j}$ denotes the external stimulus. The linking modulation is given in Eq.(3), where $U_{i,j}[n]$ is the internal state of the neuron and β is the linking strength parameter. The pulse generator determines the firing events in the model in Eq.(4). The value of $Y_{i,j}[n]$ depends on the internal state and threshold. The dynamic threshold of the neuron is given in Eq.(5), where V_T and α_T are normalized constant and time constant, respectively.

Compared to the 9 parameters of the standard PCNN model, the RPCNN contains only 4 key parameters: $W_{i,j,k,l}$, $\beta_{i,j}$, V_T and α_T (for details see [12], [21], [22]). Moreover, $W_{i,j,k,l}$ is usually kept unchanged and we set these to the reciprocal of square distance between two pixels (coefficients). Among the remaining three parameters, the linking coefficient β can vary the weighting of the linking channel in the internal activity, and hence is application dependant. Keeping this in mind, we propose to adaptively set the values of the linking strengths β , based on fuzzy logic approach, and set the values of the other two parameters, heuristically.

IV. PROPOSED METHOD

In the proposed method, coefficients of both low frequency subbands (LFSs) and high frequency subbands (HFSs) are fused in a similar way using RPCNNs with fuzzy-adaptive linking strengths. The notations used are as follows: $I = (X, Y, R)$ where X , Y , R represents the two source images and the resultant fused image, respectively. The value $B_{s,d}^I(i, j)$ indicates a coefficient of the subband B of the image I at the scale s ($= 1, \dots, S$) and direction d , where S is the coarsest scale, and (i, j) denotes the spatial location of the coefficient in the subband. The method can be easily extended to more than two images.

A. Fuzzy Adaptive Linking Strength

From PCNN related literature we know that the linking strength (β) reflects the pixel's (coefficient) characteristics, and should be adaptive to the importance (significance) of the corresponding pixel (coefficient). Moreover, from the HVS model related literature, it has been found that the contrast enhancement mechanism and incremental visual threshold can be effectively model as a non-linear system, which following

the HVS decide visually significant or insignificant pixels with respect to its neighbors [17], [23], [24]. The uncertainty exists in deciding the visual quality (significance) of the image's pixel (coefficient) and the subjectivity of the HVS response is successfully handled by fuzzy logic approaches [18], [25].

Keeping these in mind, we propose to use a novel fuzzy based technique to adaptively set the value of β , by estimating each coefficient's significance (importance) in the corresponding image. If a coefficient's "local average energy" is large or its "local information entropy" is large, then the coefficient has more importance in the image. We consider $LAE_{s,d}^I(i, j)$ and $LIE_{s,d}^I(i, j)$ as the representations of a coefficient's "local average energy" and its "local information entropy", respectively. LAE gives information about the existence of edges, contours and textures in an image. Similarly, LIE indicates the complexity or unpredictability of a local region. Regions corresponding to high signal complexity tend to have flatter distributions hence higher entropy and these regions are considered to be the important regions (edges, contours and texture information) of the image [26]. We compute two fuzzy membership values corresponding to each coefficient $B_{s,d}^I(i, j)$ using the 'S-type' membership function. Considering, $\mu_1(B_{s,d}^I(i, j))$ and $\mu_2(B_{s,d}^I(i, j))$ as the fuzzy membership values associated with $LAE_{s,d}^I(i, j)$ and $LIE_{s,d}^I(i, j)$, respectively, we compute, $\mu(B_{s,d}^I(i, j))$ as the membership value associated with the coefficient's larger "local average energy" or larger "local information entropy". This $\mu(B_{s,d}^I(i, j))$, reflecting the importance of the coefficient $B_{s,d}^I(i, j)$ in the corresponding image I , is used as the linking strength $\beta_{i,j}^{s,d,I}$.

For a coefficient $B_{s,d}^I(i, j)$, $LAE_{s,d}^I(i, j)$ and $LIE_{s,d}^I(i, j)$ are computed according to the Eq. (6) and Eq. (7), respectively, considering a window of size $M \times N$ centered around the coefficient:

$$LAE_{s,d}^I(i, j) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N B_{s,d}^I(m, n)^2 \quad (6)$$

$$LIE(B_{s,d}^I(i, j)) = - \sum p(B_{s,d}^I(i, j)) \log_2 p(B_{s,d}^I(i, j)) \quad (7)$$

where, $p(B_{s,d}^I(i, j))$ is the probability of occurrence of the coefficient $B_{s,d}^I(i, j)$.

The fuzzy membership values $\mu_1(B_{s,d}^I(i, j))$ and $\mu_2(B_{s,d}^I(i, j))$ are computed as follows:

$$\mu_1(B_{s,d}^I(i, j)) = \begin{cases} 0, & LAE_{s,d}^I(i, j) \leq a_1 \\ 2 \left(\frac{LAE_{s,d}^I(i, j) - a_1}{c_1 - a_1} \right)^2, & a_1 \leq LAE_{s,d}^I(i, j) \leq b_1 \\ 1 - 2 \left(\frac{LAE_{s,d}^I(i, j) - a_1}{c_1 - a_1} \right)^2, & b_1 \leq LAE_{s,d}^I(i, j) \leq c_1 \\ 1, & LAE_{s,d}^I(i, j) \geq c_1 \end{cases} \quad (8)$$

and

$$\mu_2(B_{s,d}^I(i, j)) = \begin{cases} 0, & LIE_{s,d}^I(i, j) \leq a_2 \\ 2 \left(\frac{LIE_{s,d}^I(i, j) - a_2}{c_2 - a_2} \right)^2, & a_2 \leq LIE_{s,d}^I(i, j) \leq b_2 \\ 1 - 2 \left(\frac{LIE_{s,d}^I(i, j) - a_2}{c_2 - a_2} \right)^2, & b_2 \leq LIE_{s,d}^I(i, j) \leq c_2 \\ 1, & LIE_{s,d}^I(i, j) \geq c_2 \end{cases} \quad (9)$$

where, $b_1 = average(LAE_{s,d}^I)$, $c_1 = b_1 + max(|b_1 - max(LAE_{s,d}^I)|, |b_1 - min(LAE_{s,d}^I)|)$, $a_1 = 2b_1 - c_1$, and similarly, $b_2 = average(LIE_{s,d}^I)$, $c_2 = b_2 + max(|b_2 - max(LIE_{s,d}^I)|, |b_2 - min(LIE_{s,d}^I)|)$, $a_2 = 2b_2 - c_2$. Here, b_i is the cross-over point, c_i is the shoulder point and a_i is the

feet point of S type membership curve, $i = 1, 2$ (considering two source images).

The linking strength $\beta_{i,j}^{s,d,I}$ corresponding to the coefficient $B_{s,d}^I(i, j)$ is then computed as follows:

$$\beta_{i,j}^{s,d,I} = \max(\mu_1(B_{s,d}^I(i, j)), \mu_2(B_{s,d}^I(i, j))) \quad (10)$$

B. Algorithm

Assuming that the medical images to be fused are co-registered to ensure that the corresponding pixels are aligned, here we outline the salient steps of the proposed MIF method:

- 1) Decompose the registered source medical images X and Y by NSCT to get the LFSs and HFSs.
- 2) Compute the linking strengths $\beta_{i,j}^{s,d,I}$, $I = (X, Y)$ as described in Section IV-A.
- 3) Input the coefficients of the subbands to motivate the RPCNNs and generate pulse of neurons using Eqs.(1)–(4), and compute the firing times $T_{i,j}^{s,d,I}[n]$, $I = (X, Y)$ by Eq.(5).
- 4) At $n = N$ (total number of iterations), determine the fused coefficient $B_{s,d}^R(i, j)$ following the fusion rule:

$$B_{s,d}^R(i, j) = \begin{cases} B_{s,d}^X(i, j), & T_{i,j}^{s,d,X}[N] \geq T_{i,j}^{s,d,Y}[N] \\ B_{s,d}^Y(i, j), & \text{otherwise} \end{cases} \quad (11)$$

- 5) Apply inverse NSCT on the fused coefficients to get the final fused medical image R .

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

V. EXPERIMENTAL RESULTS AND COMPARISONS

To evaluate the performance of the proposed technique, extensive experiments were carried out on 5 different groups of medical images combining different modalities (CT/MRI, T1_W_MR/MRA, FLAIR/DW, FDG_PET/MR and FDG_PET/CT). Each group contains at least 7 image pairs: one pair from one patient and one slice per image. For simplicity, we term the proposed neuro-fuzzy hybrid MIF method based on NSCT as NFHF-NSCT.

A. Experimental Setup

We implemented the proposed technique in MATLAB, and experiments were done on a PC with 2.66 GHz CPU and 4 GB RAM. Parameters of PCNN were set as $k \times l = 3 \times 3$, $\alpha_\theta = 0.2$, $V_\theta = 20$, $W = [0.7071 \ 1 \ 0.7071; 1 \ 0 \ 1; 0.7071 \ 1 \ 0.7071]$ and $N = 200$. The size of the window for computing the local average energy and the local information entropy, was set as 3×3 .

The selected quantitative criterions used in the objective analysis are described next briefly. Standard deviation (STD) is used to measure the image contrast (high STD means better contrast). To measure the image information content entropy (EN) is used. If entropy of fused image is higher than source images then it indicates that the fused image contains more information. Spatial frequency (SF) can be used to measure the overall activity and clarity level of an image. Larger SF value denotes better fusion result. Mutual information (MI) measures the degree of dependence of the two images. A larger MI implies better quality [27]. The other objective IF performance measures used in the experiments are $Q^{AB/F}$ and

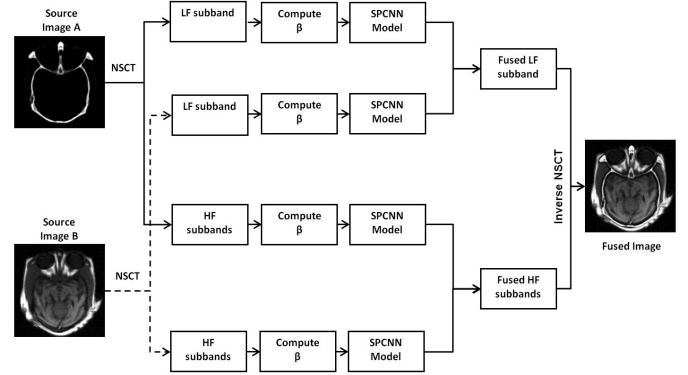


Fig. 3: Block Diagram of the proposed MIF method.

Q_0 [28]. Values of both of these two measures should be near to 1, which indicate better fused image.

The visual and quantitative results for 5 pairs of source images from 5 different combinations described previously, are given in this section. For simplicity, we have termed the five pairs of source medical images as ‘Group 1’ to ‘Group 5’, and these are shown in the first two columns of Fig. 4. In ‘Group 1’ the CT image in Fig. 4(a1) shows the bone structure, and the MRI image in Fig. 4(b1) displays the soft tissue information. The T1-weighted MR image in Fig. 4(a2) contains the soft tissues, and a lesion in the brain, but the vascular nature of the lesion is not clear. The vascular nature of the lesion is evident in MRA image of Fig. 4(b2), but the tissue information is low. The fluid attenuated inversion recovery (FLAIR) MR image of Fig. 4(a3) shows symmetrical signal hyper-intensity of the occipito-parietal cortical ribbon, and the diffusion-weighted (DW) image of Fig. 4(b3) shows increased signal in the areas of the FLAIR abnormality. In Fig. 4(a4) the coronal F-18 fluorodeoxyglucose (FDG) PET image provides the metabolic information, whereas, the coronal MR-T1 image of Fig. 4(b4) shows the structural information. The FDG-PET image of Fig. 4(a5) shows a lesion in the right lung that shows increased FDG uptake, and the CT image of Fig. 4(b5) shows the structural information with exact location of the lesion within the right lung. We have compared our proposed method with five state-of-the-art MIF schemes both subjectively and objectively. The performance of the proposed NFHF-NSCT method, is also compared with the effectiveness of other MGA-tools such as contourlet (NFHF-CNT) and curvelet (NFHF-CVT). Even though we provide the visual and quantitative results for 5 pairs of medical images, for the other source images we got similar results.

B. Subjective Analysis and Discussion

An expert radiologist was asked to subjectively evaluate the effectiveness of the proposed method. Both the fused images obtained by our proposed method, and the fused images obtained by the compared schemes were shown to the radiologist. According to the clinician opinion, it can be seen from the given results of Fig. 4, that apart from our proposed method (Fig. 4: j1-j5) and the schemes of [6] (Fig. 4: g1-g5) and [14] (Fig. 4: e1-e5), all the other compared techniques suffer from the problem of contrast reduction. Moreover, he found that the fused images obtained by the schemes of [7]

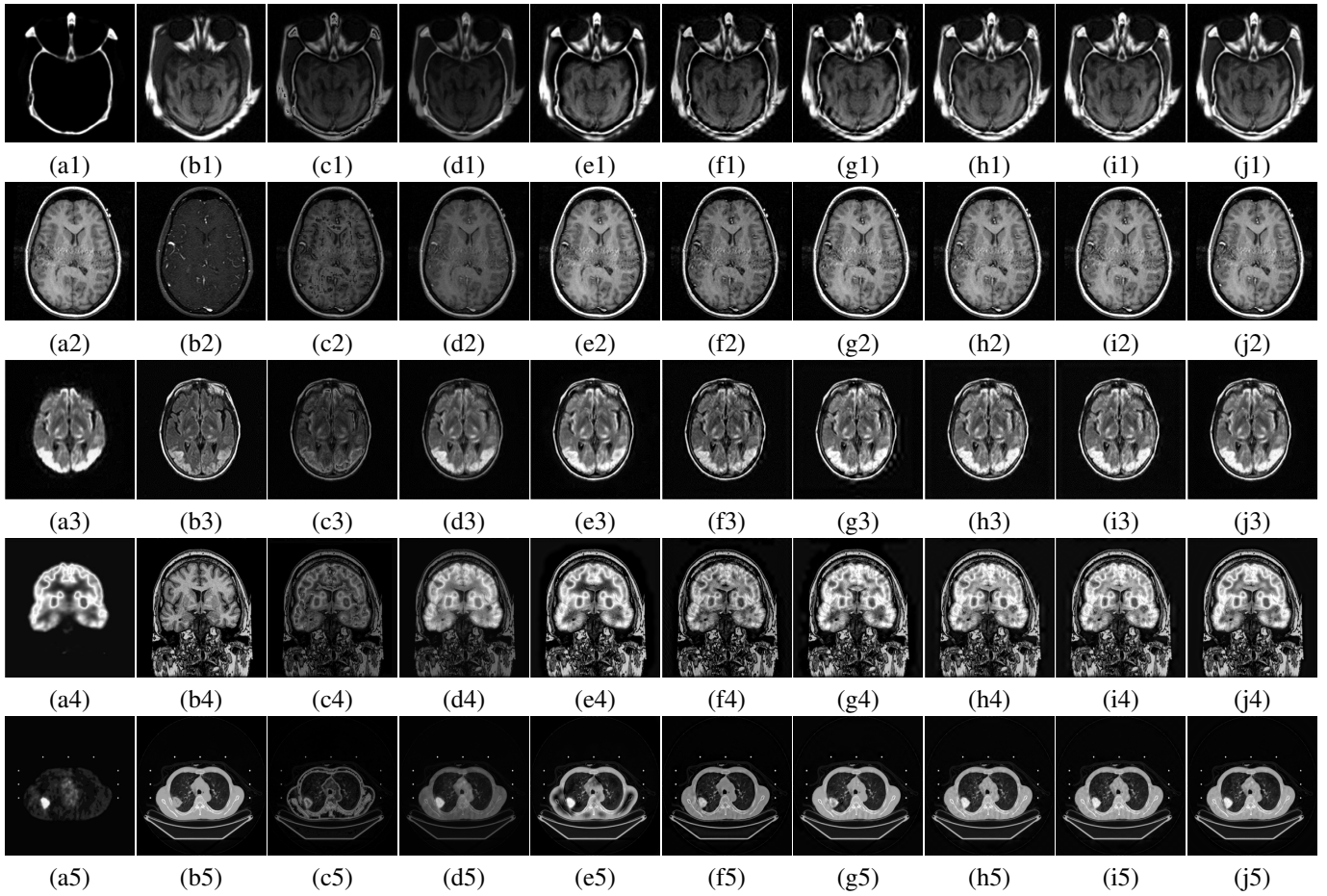


Fig. 4: Visual results for the five pairs (ak, bk) of source images, ($k = 1, 2, 3, 4, 5$). Fused images obtained: $c1-c5$ by scheme [7], $d1-d5$ by scheme [5], $e1-e5$ by scheme [14], $f1-f5$ by scheme [9], $g1-g5$ by scheme [6], $h1-h5$ by NFHF-CNT, $i1-i5$ by NFHF-CVT, $j1-j5$ by proposed scheme.

(Fig. 4: $c1-c5$), [14] (Fig. 4: $e1-e5$), [9] (Fig. 4: $f1-f5$) and [6] (Fig. 4: $g1-g5$) have lost large amount of image details. Furthermore, he observed that the methods of [9] (Fig. 4: $f1-f5$) and [6] (Fig. 4: $g1-g5$) suffer from the problem of blocking effect (as evident from the lower portions of the images) and contain unwanted image degradations. Whereas, in his opinion, the fused image Fig. 4(j1) obtained by our proposed scheme for ‘Group 1’ source images, contains both the bone structure (from the CT image of Fig. 4(a1)) and the soft tissue information (from the MRI image of Fig. 4(b1)). The lesion and its vascular nature along with the soft tissue information are evident in the fused image Fig. 4(j2) of ‘Group 2’. Both the complementary information from the source images of ‘Group 3’ can be clearly seen in the fused image Fig. 4(j3). For ‘Group 4’, the fused image shown in Fig. 4(j4) contains the metabolic information from the FDG-PET image of Fig. 4(a4) and the structural information from the T1-weighted MR image of Fig. 4(b4). The fused image Fig. 4(j5) of ‘Group 5’ shows both the structural information (exact location) and the metabolic activity of the lesion in the same image. Finally, after careful inspection of all the resultant images, the clinician conformed to the effectiveness of the proposed scheme. He found that the fused images obtained by the proposed scheme, are clearer, informative and have

higher contrast than the source medical images. For evaluating the efficiency of NSCT over CNT and CVT, we also showed the zoomed in versions of the fused images to the clinician. The zoomed in version resultant images of ‘Group 1’ images, obtained by NFHF-CNT, NFHF-CVT and NFHF-NSCT are shown in Fig. 5. The clinician conformed that even though, the fused images obtained by NFHF-CNT and NFHF-CVT look similar to the fused images produced by our proposed technique. But, both of these transforms cause blurring of edges and unwanted image degradations as shown in Fig. 5. We can clearly see from the resultant images given in Figs. 4 and 5 that the proposed MIF method results in low contrast reduction, high clarity and high information content. The proposed MIF scheme also causes less unwanted degradations in the fused images, as well as is free from the problem of blocking effect.

C. Objective Analysis and Discussion

For the five pairs of source medical images the detailed quantitative evaluations are given in Table I and Fig.6. Columns 3 to 5 in the Table I, show the SF, EN and STD of the source medical images, respectively. The values of these quantitative measures of the fused images obtained by the proposed technique are given in columns 6 to 8 of Table I. The ‘**bold**’ values indicate the highest values in Table I, for

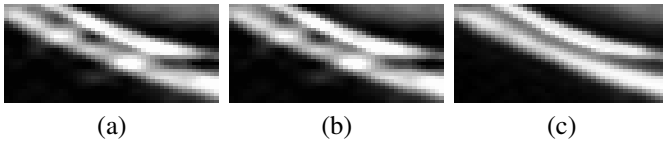


Fig. 5: Performance comparison of CNT, CVT and NSCT. Zoomed in versions of ‘Group 1’ fused images of Fig. 4: (a) $h1$ for CNT, (b) $i1$ for CVT and (c) $j1$ for NSCT.

TABLE I: Performance Evaluation of the Proposed Scheme

Group	Source Image			Fused Image			
	No.	SF	EN	STD	SF	EN	STD
1	a1	4.4316	1.7126	44.7519	7.2512	6.7918	64.6989
	b1	6.2600	5.6013	58.8283			
2	a2	7.7005	4.1524	69.1972	7.9600	6.3514	69.1150
	b2	6.4901	4.3310	25.5812			
3	a3	10.4970	2.4263	59.7992	10.5607	5.8155	64.4903
	b3	12.7502	2.6375	49.6101			
4	a4	5.1728	3.2840	67.1263	9.5685	6.6329	74.2056
	b4	9.3992	5.3682	64.4280			
5	a5	2.8705	1.9766	19.8552	5.8448	4.9837	56.5273
	b5	5.6831	5.0498	56.8748			

that quantitative measure. The highest values of SF for the images of Groups 1,2,4 and 5 indicate that the fused images obtained by our proposed method have more activity and clarity level than the source images. Similarly the highest values of EN for the fused images show that the fused images obtained by the proposed scheme, have more information content than the source images. We can also see from Table I that the STD values of the fused images for 3 out of 5 source image combinations are higher than their corresponding source images. This shows that the proposed method produces higher contrast fused images. Only in case of image groups 2 and 5, the STD values of one of the corresponding source images (Fig. 4(a2) and Fig. 4(b5)) are greater than the fused images. It may be because of the fact that the other source images (Fig. 4(b2) and Fig. 4(a5)) of the image groups 2 and 5, have very low contrast (indicated by low STD values), causing the fused images to have lower STD values (lower by very small amount). Therefore, it is evident from Table I, that the fused images obtained by the proposed method are more clear, informative and have higher contrast which is helpful in visualization and interpretation.

The quantitative performance comparisons of our proposed method against some of the existing MIF schemes are given in the Fig. 6 in form of bar charts. It is clear from Fig. 6 that the proposed scheme has all the largest objective evaluations (5 out of 6), which is obviously better than the other methods. The highest values of SF indicate that the fused images obtained by our proposed method, have more activity and clarity level than the other schemes. Similarly the highest 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 other compared methods. It is also clear that the method [5] is better than our proposed scheme in view of the quantitative measure MI. It may be because that the method of [5] is based on a multi-channel PCNN in spatial domain, which can retain the source image information in the fused image better than our proposed scheme. But, our proposed scheme

is superior for all other quantitative measures. Moreover, the RPCNN used in the proposed technique having less complex structure and parameters is computationally efficient than the original PCNN. This helps in reducing the computational cost of the overall system. Specifically, the proposed technique requires approximately 45 to 55 seconds to fuse a pair of source medical images of size 512×512 , irrespective of their modalities. Therefore, it is obvious from Table I and Fig. 6, that the fused images obtained by the proposed MIF method based on hybrid neuro-fuzzy technique and NSCT, are more clear, informative and have higher contrast than the existing mentioned MIF methods.

VI. CONCLUSION

We propose a novel MIF method based on a hybrid neuro-fuzzy approach in NSCT domain. To overcome the drawbacks of traditional MIF schemes, and to integrate as much information as possible into the fused images we exploit the advantages of NSCT, RPCNN and fuzzy logic. Linking strengths of the neurons in the RPCNNs are adaptively computed based on the fuzzy characteristics of the image, which results in high quality fused images. 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. The effects of different fusion rules, as well as new techniques to compute the parameters of the neurons of the PCNN, are some of the future scopes of the proposed technique.

ACKNOWLEDGMENT

We would like to thank Machine Intelligence Unit, Indian Statistical Institute, Kolkata-108 (Internal Academic Project) for providing facilities to carry out this work. We are grateful to Dr. Pradip Kumar Das (Medicare Images, Asansol-4, West Bengal) for providing us with some of the source medical images and for the subjective evaluation of the fusion results. We also like to thank <http://www.imagefusion.org/> and <http://www.med.harvard.edu/aanlib/home.html> for providing us the source medical images.

REFERENCES

- [1] B. Solaiman, R. Debon, F. Pipelier, J. M. Cauvin, and C. Roux, “Information fusion: Application to data and model fusion for ultrasound image segmentation,” *IEEE TBME*, vol. 46, no. 10, pp. 1171–1175, 1999.
- [2] 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.
- [3] S. Das and M. K. Kundu, “NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency,” *Medical and Biological Engineering and Computing*, vol. 50, no. 10, pp. 1105–1114, 2012.
- [4] V. S. Petrovic and S. Xydeas, “Gradient-based multiresolution image fusion,” *IEEE TIP*, vol. 13, no. 2, pp. 228–237, 2004.
- [5] Z. Wang and Y. Ma, “Medical image fusion using m-PCNN,” *Information Fusion*, vol. 9, no. 2, pp. 176–185, April 2008.
- [6] 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.
- [7] 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.
- [8] G. Pajares and J. M. de la Cruz, “A wavelet-based image fusion tutorial,” *Pattern Recognition*, vol. 37, no. 9, pp. 1855–1872, 2004.

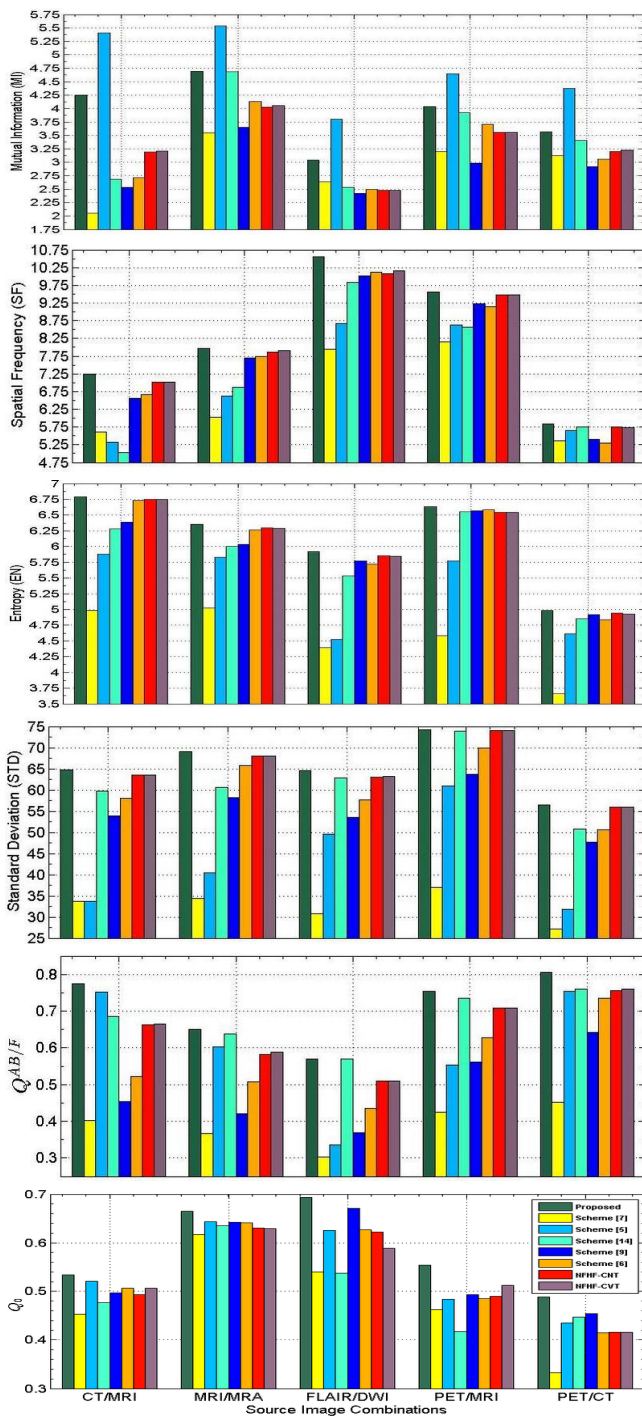


Fig. 6: Objective Performance Comparisons

[9] 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.

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

[11] 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.

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

[13] J. L. Johnson, "Time signatures of images," in *Proc. of IEEE Neural Networks*, vol. 2, 1994, pp. 1279-1284.

[14] 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 nonsubsamped contourlet transform domain," *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1508-1514, 2008.

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

[16] J. M. Kinsler, "Foveation by a pulse-coupled neural network," *IEEE NN*, vol. 10, no. 3, pp. 621-626, 1999.

[17] M. K. Kundu and S. K. Pal, "A note on the transformation between intensity and gray level," *Pattern Recognition Letters*, vol. 8, pp. 257-269, 1988.

[18] M. K. Kundu and S. K. Pal, "Automatic selection of object enhancement operator with quantitative justification based on fuzzy set theoretic measure," *Pattern Recognition Letters*, vol. 11, pp. 811-829, 1990.

[19] J. M. Kinsler, "Pulse coupled image fusion," *Optical Engineering*, vol. 36, no. 3, pp. 737-742, 1997.

[20] A. da Cunha, J. Zhou, and M. Do, "The nonsubsamped contourlet transform: Theory, design, and applications," *IEEE TIP*, vol. 15, no. 10, pp. 3089-3101, 2006.

[21] J. Johnson and M. Padgett, "PCNN models and applications," *IEEE TNN*, vol. 10, no. 3, pp. 480-498, may 1999.

[22] J. A. Karvonen, "Baltic sea ice SAR segmentation and classification using modified pulse-coupled neural networks," *IEEE TGRS*, vol. 42, no. 7, pp. 1566-1574, 2004.

[23] G. Buchsbaum, "An analytical derivation of visual nonlinearity," *IEEE TBME*, vol. 27, no. 5, pp. 237-242, 1980.

[24] M. K. Kundu and S. K. Pal, "Thresholding for edge detection using human psycho visual phenomena," *Pattern Recognition Letters*, vol. 4, pp. 433-441, 1986.

[25] H. Cheng and H. Xu, "A novel fuzzy logic approach to contrast enhancement," *Pattern Recognition*, vol. 33, no. 5, pp. 809-819, 2000.

[26] T. Kadir and M. Brady, "Saliency, scale and image description," *International Journal of Computer Vision*, vol. 45, no. 2, pp. 83-105, 2001.

[27] 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.

[28] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronic Letters*, vol. 36, no. 4, pp. 308-309, 2000.



Sudeb Das received the M.C.A. degree from WBUT, India, in 2008. He is currently working toward his Ph.D. from Calcutta University in Computer Sci. and doing his research in Indian Statistical Institute, Kolkata, India. He has contributed about 15 research papers to well-known and prestigious archival journals, international refereed conferences. His research areas include Pattern Recognition, Digital Watermarking, CBIR, biomedical imaging and Soft computing etc.



Malay Kumar Kundu (M'90-SM'99) received his B. Tech., M. Tech. and Ph.D (Tech.) degrees in Radio physics and Electronics all are from the University of Calcutta. Currently he is a full professor in the Machine Intelligence Unit of the Indian Statistical Institute, Kolkata, India. He had been the head of the Machine Intelligence Unit from September 1993 to November 1995 and Professor In-charge (Chairman) of the Computer and Communication Sciences Division of the Institute during 2004 to 2006. He is a Fellow of a The International Association for Pattern Recognition, USA (FIAPR), Indian National Academy of Engineering (FNAE), National Academy of Sciences (FNASc.), India and the Institute of Electronics and Telecommunication Engineers (FIETE), India. A senior member of the IEEE, USA and the founding life member and Vice President of the Indian Unit for Pattern Recognition and Artificial Intelligence (IUPRAI). He was selected as INAE Distinguished Professor in 2013. He received the prestigious VASVIK award for industrial research in the field of Electronic Sciences and Technology for the year 1999. He also received the Sir. J. C. Bose memorial award of the Institute of Electronics and Telecommunication Engineers (IETE), India in the year 1986. His current research interest includes Image/video processing and analysis, soft computing, computer vision, machine intelligence and data security. He has contributed 3 book volumes, about 140 research papers in well known and prestigious archival journals, international refereed conferences and in the edited monograph volumes. He is the holder of nine U.S patents.