Med Biol Eng Comput (2012) 50:1105–1114
DOI 10.1007/s11517-012-0943-3

ORIGINAL ARTICLE

NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency

Sudeb Das · Malay Kumar Kundu

Received: 22 October 2011 / Accepted: 5 July 2012 / Published online: 24 July 2012

Abstract In this article, a novel multimodal medical image fusion (MIF) method based on non-subsampled contourlet transform (NSCT) and pulse-coupled neural network (PCNN) is presented. The proposed MIF scheme exploits the advantages of both the NSCT and the PCNN to obtain better fusion results. The source medical images are first decomposed by NSCT. The low-frequency subbands (LFSs) are fused using the 'max selection' rule. For fusing the high-frequency subbands (HFSs), a PCNN model is utilized. Modified spatial frequency in NSCT domain is input to motivate the PCNN, and coefficients in NSCT domain with large firing times are selected as coefficients of the fused image. Finally, inverse NSCT (INSCT) is applied to get the fused image. Subjective as well as objective analysis of the results and comparisons with state-of-the-art MIF techniques show the effectiveness of the proposed scheme in fusing multimodal medical images.

Keywords Image fusion · Pulse-coupled neural network · Multiscale geometric analysis · Medical imaging · NSCT

This work was supported by the Machine Intelligence Unit, Indian Statistical Institute, Kolkata-108 (Internal Academic Project).

S. Das (✉) · M. K. Kundu
Machine Intelligence Unit, Indian Statistical Institute,
203 B.T.Road, Kolkata 700108, India
e-mail: to.sudeb@gmail.com

M. K. Kundu
e-mail: malay@isical.ac.in

1 Introduction

Over the last few decades, medical imaging is playing an increasingly critical and vital role in a large number of healthcare applications including diagnosis, research, treatment, and education etc. To provide support to the physicians, various modalities of medical images have become available, reflecting different information of human organs and tissues and possessing their respective application ranges. For instance, structural medical images like magnetic resonance imaging (MRI), computed tomography (CT), ultrasonography (USG), magnetic resonance angiography (MRA), etc. provide high-resolution images with anatomical information. Whereas, functional medical images, such as position emission tomography (PET), Single-photon emission CT (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. Therefore, combining anatomical and functional medical images to provide much more useful information through image fusion (IF) has become the focus of imaging research [1].

So far, many IF techniques have been proposed by various researchers. It has been found that the pixel-level spatial domain IF methods usually lead to contrast reduction. Methods based on intensity-hue-saturation (IHS), principal component analysis (PCA), and the Brovey transform offer better results, but suffer from spectral degradation [25]. Pyramidal IF schemes such that the laplacian pyramid, gradient pyramid, contrast pyramid, ratio-of-low-pass pyramid, and the morphological pyramid, etc. fail to introduce any spatial orientation selectivity in the decomposition process, and hence often cause blocking effects [10]. The widely used discrete wavelet transform (DWT) can preserve spectral information efficiently but

cannot express spatial characteristics effectively [15, 24]. Therefore, DWT-based fusion schemes cannot preserve the salient features of the source images efficiently, and introduce artifacts and inconsistencies in the fused results [22]. Recently, several multiscale geometric analysis (MGA) tools were developed, such as Curvelet, Contourlet, NSCT, Ripplet, etc. which do not suffer from the problems of wavelet. Many IF and MIF methods based on these MGA tools were also proposed [3, 12, 13, 22].

PCNN is a visual cortex-inspired neural network characterized by the global coupling and pulse synchronization of neurons [6, 17]. It has been observed that PCNN-based IF schemes outperform the conventional IF methods [4, 5, 8, 11, 18, 23]. Even though there exists several IF schemes based on transform domain and PCNN, most of these methods suffer from various problems. Wang et al. [16] have proposed a fast MIF scheme based on a multi-channel PCNN (m-PCNN) model with easy extensibility capability, producing fused images with high information content, but suffering from the problems of contrast reduction and loss of image fine details. Bo et al. [19] have developed an IF method based on spatial frequency (SF) motivated PCNN in NSCT domain. It works well for multifocus IF and visible/infrared IF, but the absence of directional information in SF and the use of same fusion rule for both the subbands cause contrast reduction and loss of image details. The IF technique proposed by Xin et al. [20] based on dual-layer PCNN model with a negative feedback control mechanism in the NSCT domain has shown promising results in multifocus IF. Deepika et al. [4] have proposed a combined method of MIF and edge deduction based on NSCT and PCNN. This scheme also suffers from the problems of contrast reduction and unwanted image degradations. The technique proposed by Feng et al. [8] based on bi-dimensional empirical mode decomposition and m-PCNN shows good result in preserving the source images fine details in the fused image, but suffers from contrast reduction. In most of the existing IF methods based on PCNN, the value of a single pixel (coefficient) in spatial or transform domain is used to motivate one neuron [4, 16, 20]. But this simple use of pixels (coefficients) in spatial or transform domain is not effective enough, because humans are sensitive to edges and directional features. Moreover, it has also been found that using different fusion rules for different subbands result in better fused images.

The field of MIF is quite different from that of multifocus and visible/infrared IF. Most of the times, there are very subtle differences between the features of the source medical images. Special care has to be taken during the fusion process of these fine details. Therefore, we need a MIF scheme that can simultaneously handle the problems of contrast reduction, loss of image details and unwanted image degradations. The main contribution of our proposed MIF method is to use the shift-invariance, multi-scale and multi-directional properties of NSCT along with the modified spatial frequency (capable of capturing the fine details present in the image [26]) motivated PCNN in such a way that can capture the subtle differences and the fine details present in the source medical images that result in fused images with high contrast, clarity and information content.

2 Methods

2.1 Non-subsampled contourlet transform (NSCT)

In 2006, da Cunha et al. [2] proposed an overcomplete transform called the NSCT. NSCT is a fully shift-invariant, multiscale and multidirection expansion that has a fast implementation. The Contourlet transform (CT) is not shift invariant due to the presence of the down-samplers and up-samplers in both the Laplacian pyramid (LP) and directional filter bank (DFB) stages of CT [2]. NSCT achieves shift-invariance property by using the non-subsampled pyramid filter bank (NSP or NSPFB) and the non-subsampled DFB (NSDFB).

2.1.1 Non-subsampled pyramid filter bank

NSPFB is a shift-invariant filtering structure accounting for the multiscale property of the NSCT. This is achieved by using two-channel non-subsampled 2D filter banks. It has no downsampling or upsampling and hence shift-invariant. Perfect reconstruction is achieved provided the filters satisfy the following identity

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1 \tag{1}$$

where $H_0(z)$ is the lowpass decomposition filter, $H_1(z)$ is the highpass decomposition filter, $G_0(z)$ is the lowpass reconstruction filter, and $G_1(z)$ is the highpass reconstruction filter.

In order to obtain the multiscale decomposition, NSPFB are constructed by iterated non-subsampled filter banks. For the next level, all filters are upsampled by 2 in both dimensions. Therefore, they also satisfy the perfect reconstruction identity. The equivalent filters of a kth level cascading NSPFB are given by

$$H_n^{eq}(z) = \begin{cases} H_1(z^{2^{n-1}}) \prod_{j=0}^{n-2} H_0(z^{2^j}), & 1 \le n < 2^k \\ \prod_{j=0}^{n-1} H_0(z^{2^j}), & n = 2^k \end{cases} \tag{2}$$

where z^j stands for $[z_1^j, z_2^j]$.

2.1.2 Non-subsampled directional filter bank

The NSDFB is constructed by eliminating the downsamplers and upsamplers of the DFB by switching off the

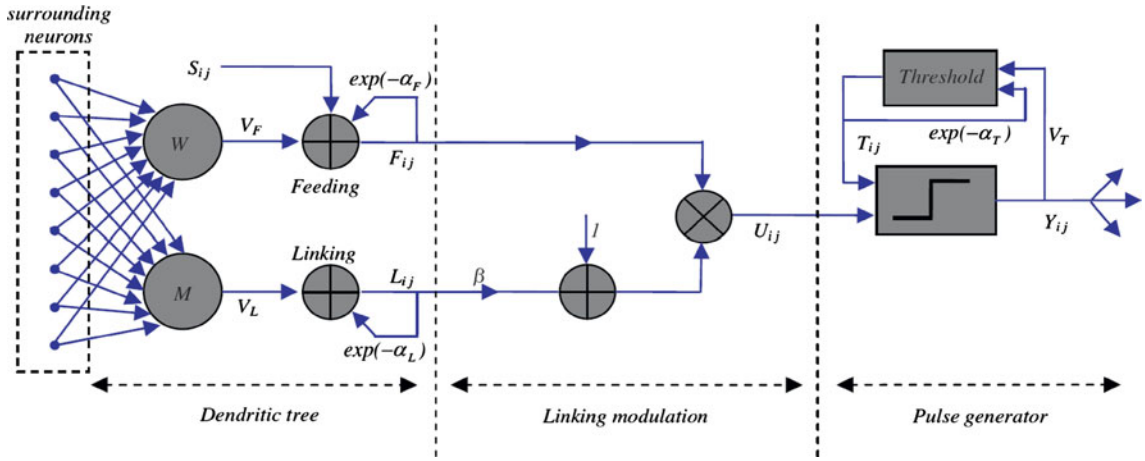

Fig. 1 Structure of PCNN

downsamplers/upsamplers in each two channel filter bank in the DFB tree structure and upsampling the filters accordingly [2]. The outputs of the first level and second level filters are combined to get the four directional frequency decomposition. 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 multidirectional decomposition, the NSDFBs are iterated and to get the next level decomposition all filters are upsampled by a quincunx matrix given by

$$\mathrm{QM} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3}$$

The NSCT is obtained by combining the 2D NSPFB and the NSDFB. The resulting filtering structure approximates the ideal partition of the frequency plane. It must be noted that different from the contourlet expansion the NSCT has a redundancy given by $R = \sum_{j=o}^{j} 2^{l_j}$, where 2^{l_j} is the number of directions at scale j.

2.2 Pulse-coupled neural network

PCNN is a single layered, two-dimensional, laterally connected neural network of pulse coupled neurons. The PCNN neuron's structure is shown in Fig. 1. 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 neuron's 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 [9, 17]:

$$F_{i,j}[n] = e^{-\alpha_F} F_{i,j}[n-1] + V_F \sum_{k,l} w_{i,j,k,l} Y_{i,j}[n-1] + S_{i,j} \tag{4}$$

$$L_{i,j}[n] = e^{-\alpha_L} L_{i,j}[n-1] + V_L \sum_{k,l} m_{i,j,k,l} Y_{i,j}[n-1] \tag{5}$$

$$U_{i,j}[n] = F_{i,j}[n](1 + \beta L_{i,j}[n]) \tag{6}$$

$$Y_{i,j}[n] = \begin{cases} 1, & U_{i,j}[n] > T_{i,j}[n] \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$T_{i,j}[n] = e^{-\alpha_T} T_{i,j}[n-1] + V_T Y_{i,j}[n] \tag{8}$$

In Eq. (4)–(8), 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. (4)–(5). 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, and generally $\alpha_F < \alpha_L$. The linking modulation is given in Eq. (6), 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. (7). $Y_{i,j}[n]$ depends on the internal state and threshold. The dynamic threshold of the neuron is Eq. (8), where V_T and α_T are normalized constant and time constant, respectively.

2.3 Proposed MIF scheme

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 low-frequency subband (LFS) of the image C at the coarsest scale G. $D_{g,h}^C$ represents the high-frequency subband (HFS) of the image C at scale $g, (g = 1, \ldots, G)$ and direction h. (i, j) denotes the spatial location of each coefficient. The method can be easily extended to more than two images.

2.3.1 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:

$$L_G^R(i,j) = \begin{cases} L_G^A(i,j), & |L_G^A(i,j)| \geq |L_G^B(i,j)| \\ L_G^B(i,j), & \text{otherwise} \end{cases} \tag{9}$$

2.3.2 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 NSCT domain directly (i.e., using individual coefficients), modified spatial frequency (MSF) in NSCT domain is considered as the image feature to motivate the PCNN.

Spatial frequency (SF) proposed by Eskicioglu et al. [7] is calculated by row and column frequency. 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). The original SF lacks the directional information present in the image which results in the loss of important fine details of the image. Whereas, MSF incorporates this directional information and this results in an image clarity/activity level measure capable of capturing the fine details present in the image [26]. For an $M \times N$ pixel image, the MSF is defined as

$$\text{MSF} = \sqrt{\text{RF}^2 + \text{CF}^2 + \text{DF}^2} \tag{10}$$

where,

$$\text{RF} = \sqrt{\frac{1}{M(N-1)} \sum_{m=1}^{M} \sum_{n=2}^{N} [f_{m,n} - f_{m,n-1}]^2} \tag{11}$$

$$\text{CF} = \sqrt{\frac{1}{(M-1)N} \sum_{m=2}^{M} \sum_{n=1}^{N} [f_{m,n} - f_{m-1,n}]^2} \tag{12}$$

and,

$$\text{DF} = P + Q \tag{13}$$

where,

$$P = \sqrt{\frac{1}{(M-1)(N-1)} \sum_{m=2}^{M} \sum_{n=2}^{N} [f_{m,n} - f_{m-1,n-1}]^2} \tag{14}$$

and,

$$Q = \sqrt{\frac{1}{(M-1)(N-1)} \sum_{m=2}^{M} \sum_{n=2}^{N} [f_{m-1,n} - f_{m,n-1}]^2} \tag{15}$$

Let, $\text{MSF}_{i,j}^{g,h,C}$ 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:

$$F_{i,j}^{g,h,C}[n] = \text{MSF}_{i,j}^{g,h,C} \tag{16}$$

$$L_{i,j}^{g,h,C}[n] = e^{-\alpha_L} L_{i,j}^{g,h,C}[n-1] + V_L \sum_{k,l} M_{i,j,k,l}^{g,h,C} Y_{i,j,k,l}^{g,h,C}[n-1] \tag{17}$$

$$U_{i,j}^{g,h,C}[n] = F_{i,j}^{g,h,C}[n] * (1 + \beta L_{i,j}^{g,h,C}[n]) \tag{18}$$

$$\theta_{i,j}^{g,h,C}[n] = e^{-\alpha_\theta} \theta_{i,j}^{g,h,C}[n-1] + V_\theta Y_{i,j}^{g,h,C}[n-1] \tag{19}$$

$$Y_{i,j}^{g,h,C}[n] = \begin{cases} 1, & U_{i,j}^{g,h,C}[n] > \theta_{i,j}^{g,h,C}[n] \\ 0, & \text{otherwise} \end{cases} \tag{20}$$

$$T_{i,j}^{g,h,C}[n] = T_{i,j}^{g,h,C}[n-1] + Y_{i,j}^{g,h,C}[n] \tag{21}$$

where, the feeding input $F_{i,j}^{g,h,C}$ is equal to the modified spatial frequency $\text{MSF}_{i,j}^{g,h,C}$. The linking input $L_{i,j}^{g,h,C}$ is equal to the sum of neurons firing times in linking range. $M_{i,j,k,l}$ is the synaptic gain strength and subscripts k and l are the size of linking range in the PCNN. α_L is the decay constant. 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 time.

2.4 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 NSCT to get the LFSs and HFSs.

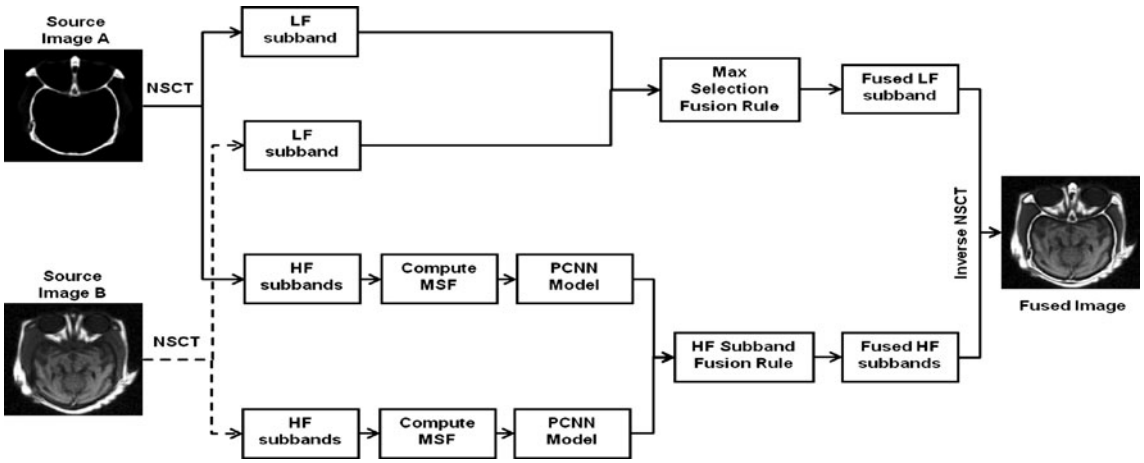

Fig. 2 Block diagram of the proposed MIF method

2. Fused the coefficients of LFSs using the 'max selection' rule described in Sect. 2.3.1, to get the fused LFS.
3. Compute the MSF as described in Sect. 2.3.2, using overlapping window on the coefficients in HFSs.
4. Input MSF of each HFS to motivate the PCNN and generate pulse of neurons with Eqs. (16)–(20) and compute the firing times $T_{i,j}^{g,h,C}[n]$ by Eq. (21).
5. If $n = N$, then iteration stops. Then fuse the coefficients of the HFSs by the following fusion rule:

$$D_{g,h}^{R}(i,j) = \begin{cases} D_{g,h}^{A}(i,j), & T_{i,j}^{g,h,A}[N] \geq T_{i,j}^{g,h,B}[N] \\ D_{g,h}^{B}(i,j), & \text{otherwise} \end{cases} \quad (22)$$

6. Apply inverse NSCT (INSCT) 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. 2.

3 Results

3.1 Experimental setup

We implemented the proposed technique in MATLAB, and experiments were carried out on a PC with 2.66 GHz CPU and 4 GB RAM. The decomposition parameter of NSCT was *levels* = [1, 2, 4] and we used 'pyrexc' and 'vk' as the pyramid filter and orientation filter, respectively. Parameters of PCNN was set as $k \times l = 3 \times 3$, $\alpha_L = 0.06931$, $\alpha_\theta = 0.2$, $\beta = 0.2$, $V_L = 1.0$, $V_\theta = 20$, $W = [0.707\ 1\ 0.707,\ 1\ 0\ 1,\ 0.707\ 1\ 0.707]$ and $N = 200$.

The selected quantitative criterions used in the objective analysis are discussed in the following sections.

3.1.1 Standard deviation (STD)

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

3.1.2 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

$$EN = -\sum_{g=0}^{L-1} p(g) \log_2 p(g) \quad (23)$$

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.

3.1.3 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.

3.1.4 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 [14]:

$$MI = I(x_A; x_F) + I(x_B; x_F) \quad (24)$$

where,

$$I(x_R;x_F)=\sum_{u=1}^{L}\sum_{v=1}^{L}h_{R,F}(u,v)log_2\frac{h_{R,F}(u,v)}{h_R(u)h_F(v)} \quad (25)$$

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.

3.1.5 $Q^{AB/F}$

Xydeas et al. [21] proposed an objective image fusion performance measure $Q^{AB/F}$ as follows:

$$Q^{AB/F}=\frac{\sum_{n=1}^{N}\sum_{m=1}^{M}(Q^{AF}(n,m)w^A(n,m)+Q^{BF}(n,m)w^B(n,m))}{\sum_{n=1}^{N}\sum_{m=1}^{M}(w^A(n,m)+w^B(n,m))} \quad (26)$$

where $Q^{AF}(n,m)=Q_G^C(n,m)Q_\alpha^{AF}(n,m)$. $Q_g^{AF}(n,m)$ and $Q_\alpha^{AF}(n,m)$ are the edge strength and orientation preservation values, respectively. n, m represent the pixel location and N, M are the size of images, respectively. $Q^{BF}(n,m)$ is similarly computed. $w^A(n,m)$ and $w^B(n,m)$ reflect the importance of $Q^{AF}(n,m)$ and $Q^{BF}(n,m)$, respectively. The dynamic range of $Q^{AB/F}$ is [0, 1], and it should be as close to 1 as possible.

3.1.6 Q_0

It is a universal image quality index proposed by Wang et al. [21]. Q_0, between the source image A and the fused image F is defined as follows:

$$Q_0(A,F)=\frac{2\sigma_{af}\cdot 2\bar{a}\bar{f}}{(\sigma_a^2+\sigma_f^2)\cdot(\bar{a}^2+\bar{f}^2)} \quad (27)$$

where σ_{af} represents the covariance between A and F. σ_a, σ_f indicates the standard deviation of A and F; and $\bar{a},\bar{f}$ represent the mean value of A and F, respectively. $Q_0(A,B,F)$ is the average between $Q_0(A,F)$ and $Q_0(B,F)$:

$$Q_0(A,B,F)=\frac{Q_0(A,F)+Q_0(B,F)}{2} \quad (28)$$

Note that $-1\le Q_0\le 1$, and it should be also as close to 1 as possible.

Figure 3, shows five pairs of source medical images of different modalities used in the experiments along with the corresponding fused results obtained by the proposed method. In Fig. 3, $Ci\,(i=1,2,\ldots,5)$ indicates the image combinations: $Ci=(ai,bi,fi)$, ai and bi are the two groups of source images and fi represents the fused results.

The CT image in Fig. 3(a1) shows the bones and the MRI image in Fig. 3(b1) displays the soft tissue information. The T1-weighted MR image in Fig. 3(a2) contains the soft tissues and it also shows 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. 3(b2), but the tissue information is low. In Fig. 3(a3) and 3(b3), CT image demonstrates the calcification and the MR image reveals several focal lesions involving basal ganglia with some surrounding edema, respectively. Both the MR images of Fig. 3(a4) and (b4) show a lesion in the frontal lobe. The CT image in Fig. 3(a5) indicates a medial left occipital infarct involving the left side of the splenium of the corpus callosum and the MR image in Fig. 3(b5) reveals only mild narrowing of the left posterior cerebral artery. For the five source medical images of Fig. 3, the detail quantitative evaluation is given in Table 1. The Table 2 shows the performance comparisons of our proposed method against some of the existing MIF schemes using the images of the image combinations $C1$ and $C5$ as the source images. Fused images for the image combinations $C1$ and $C5$ obtained by the compared methods of Table 2 are shown in Fig. 4. To support our choice of MSF over SF, we also conducted an experiment where all the other configurations of the proposed MIF scheme were kept same, only SF was used instead of MSF (named NSCT_PCNN_SF for convenience). Table 2 and Fig. 4 also contain the quantitative results and the fused images obtained by the method NSCT_PCNN_SF.

4 Discussion

4.1 Subjective analysis and discussion

An expert radiologist was asked to subjectively evaluate the effectiveness of the proposed MIF method. After careful manual inspection of the images of the Fig. 3, the radiologist conformed to the effectiveness of the proposed scheme. He found that the fused images obtained by the proposed MIF scheme were more clear, informative and have higher contrast than the source medical images that is helpful in visualization as well as interpretation. The fused image of image combination $C1$ contains both the bone structure [from Fig. 3(a1)] and the soft tissue information [from Fig. 3(b2)]. Both the lesion and its vascular nature along with the soft tissue information are evident in the fused image [Fig. 3(f2)] of the image combination $C2$. Similarly, the fused images of the other image combinations ($C3$, $C4$ and $C5$) contain information from both the corresponding source images. The resultant fused images

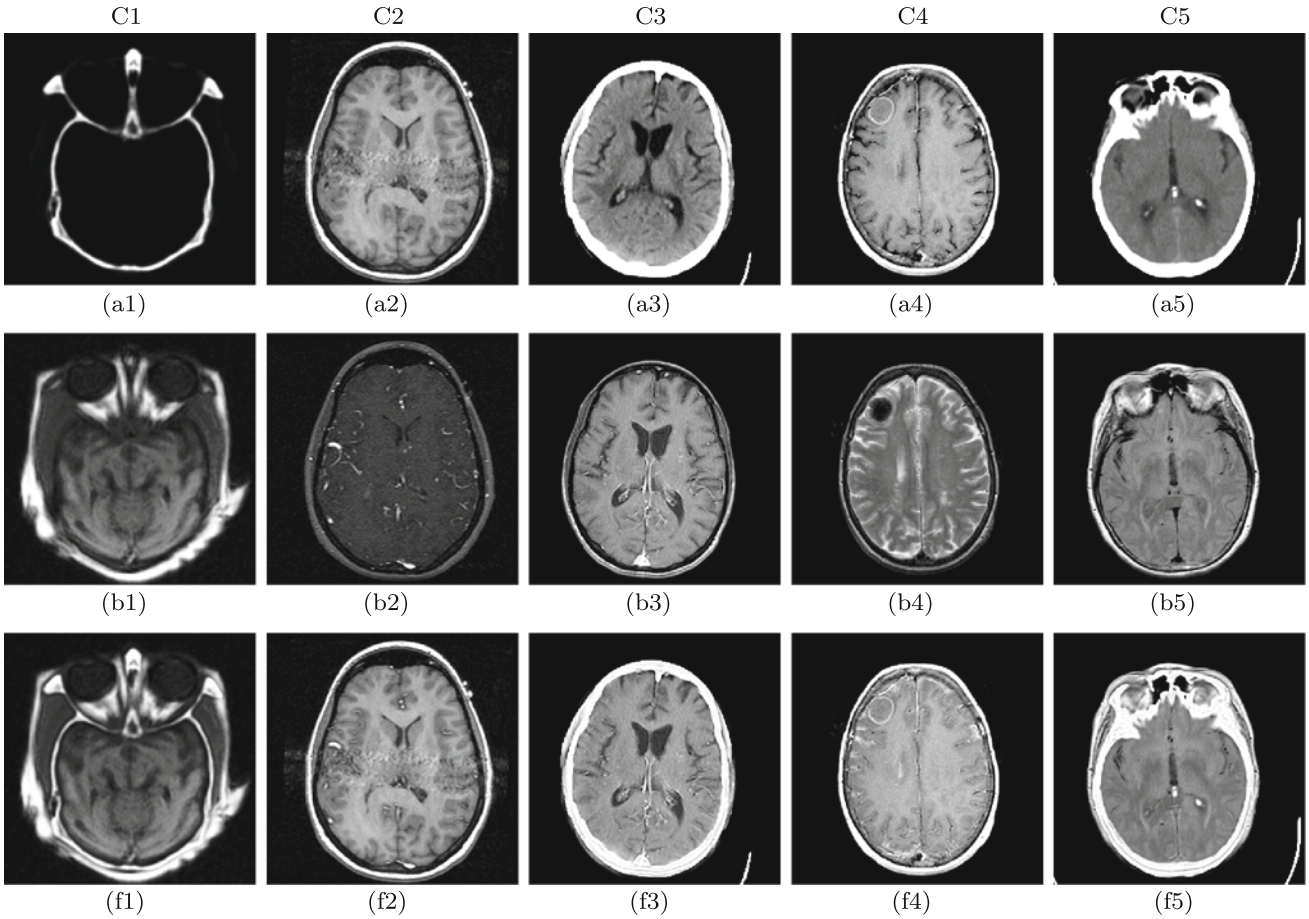

Fig. 3 Source images (*top two rows*) with fusion results (*last row*): (source images are downloaded from http://www.imagefusion.org/; http://www.med.harvard.edu/aanlib/home.html); *a*1, CT; *b*1, MRI; *a*2, T1-weighted MR; *b*2, MRA; *a*3, CT; *b*3, T1-weighted MR-GAD; *a*4, T1-weighted MR; *b*4, T2-weighted MR; *a*5, CT; *b*5, Proton density (PD) weighted MR

Table 1 Performance evaluation of the proposed MIF scheme

Combination	Name	SF	EN	STD	Fused image					
					MI	SF	EN	STD	$Q^{AB/F}$	Q_0
C1	a1	4.4316	1.7126	44.7519	4.8300	**6.9434**	**6.7724**	**65.8646**	0.7771	0.5286
	b1	6.2600	5.6013	58.8283						
C2	a2	7.7005	4.1524	**69.1972**	5.0067	**7.8946**	**6.0659**	68.9896	0.6699	0.6646
	b2	6.4901	4.3310	25.5812						
C3	a3	6.0280	3.3019	79.2907	3.1200	**6.8315**	**4.5234**	**82.3317**	0.5180	0.8990
	b3	7.2990	3.4385	61.7932						
C4	a4	6.9383	3.3046	77.1245	3.4700	**6.9678**	**4.0450**	**79.5945**	0.5410	0.8883
	b4	6.5795	3.2856	52.6946						
C5	a5	4.8089	2.9001	79.8634	3.0593	6.3261	**4.3645**	**83.7037**	0.5338	0.8796
	b5	**6.8405**	3.6014	61.9829						

of Fig. 4 obtained by the compared methods of Table 2 were also shown to the radiologist. The resultant fused images obtained by NSCT_PCNN_SF are visually very much similar to the fused images obtained by the proposed technique [as can be seen from the fused images of Figs. 3(f1), (f5) and 4(a), (g)]. But during the quantitative analysis, we have found that the fused images obtained by the proposed MIF scheme have higher quantitative results than the method of NSCT_PCNN_SF. All the compared methods of Fig. 4 except the schemes of [24] and

Table 2 Performance comparisons using *C*1 and *C*5

Scheme	Combination	MI	SF	EN	STD	$Q^{AB/F}$	Q_0
Scheme [19]	*c*1	2.6817	5.0373	6.2781	29.7318	0.6859	0.4762
	*c*5	2.6651	5.9269	4.2015	55.6347	0.3163	0.7626
Scheme [16]	*c*1	**5.4036**	5.3194	5.8783	33.7291	0.7527	0.5212
	*c*5	**3.1076**	5.9834	4.1933	55.1152	0.4958	0.8578
Scheme [15]	*c*1	2.0575	5.6108	4.9822	33.6529	0.4019	0.4527
	*c*5	2.8041	6.2651	4.1899	56.2076	0.4850	0.7529
Scheme [22]	*c*1	2.5295	6.5575	6.3877	53.8200	0.4537	0.4976
	*c*5	2.4406	6.2136	4.2635	56.5361	0.4371	0.4981
Scheme [24]	*c*1	2.7148	6.6591	6.7295	57.9787	0.5219	0.5071
	*c*5	2.6217	6.1865	4.3216	78.4728	0.4210	0.6150
Scheme NSCT_PCNN_SF	*c*1	4.7477	6.9326	6.7704	65.8304	0.7754	0.5272
	*c*5	2.9788	6.2938	4.3528	81.9448	0.5007	0.8751
Proposed scheme	*c*1	4.8300	**6.9434**	**6.7724**	**65.8646**	**0.7771**	**0.5286**
	*c*5	3.0593	**6.3261**	**4.3645**	**83.7037**	**0.5338**	**0.8796**

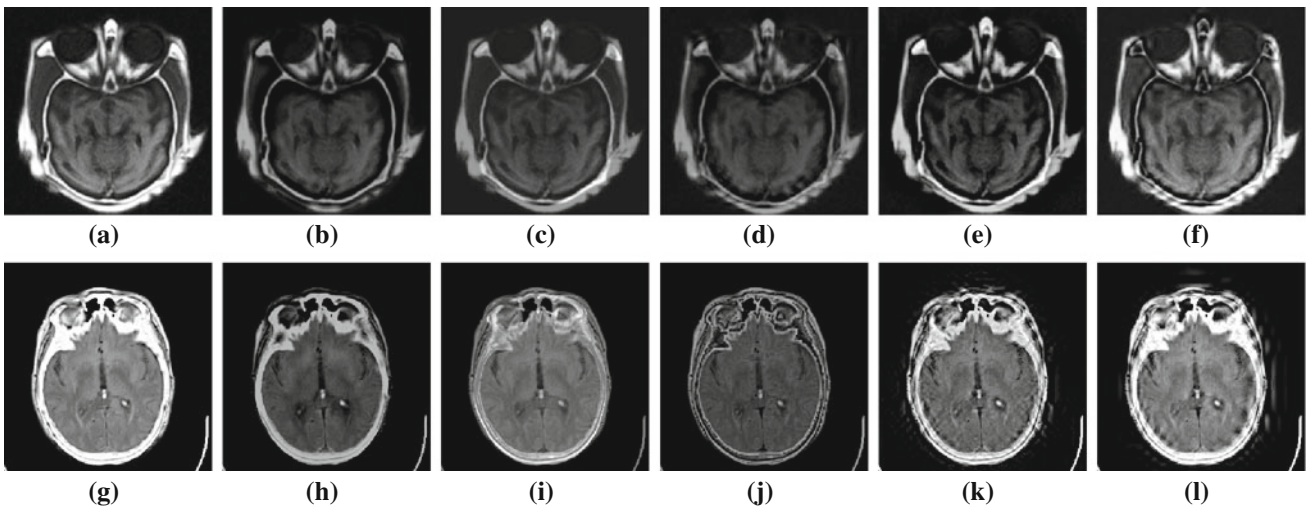

Fig. 4 Fusion results on image combinations *C*1 and *C*5: **a**, **g** method NSCT_PCNN_SF, **b**, **h** method of [19], **c**, **i** method of [16], **d**, **j** method of [15], **e**, **k** method of [22] and **f**, **l** method of [24]

NSCT_PCNN_SF suffer from the problem of contrast reduction. It is clear from the images of Fig. 4 that the methods of [15, 19, 22] [Fig. 4(b) (h), (d) (j) and (e) (k)] have lost large amount of image details. As can be easily seen from the images of Fig. 4(d)(j) and (f)(l), the methods of [15, 24] suffer from the problems of blocking effect (as evident from the lower portions of the images) and contain unwanted image degradations. We can clearly see from the resultant images given in Figs. 3 and 4 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. Therefore, it is clear from the subjective analysis of the fused images that the proposed MIF method is very effective in fusing multi-modality medical images and superior than many state-of-the-art MIF techniques.

4.2 Objective analysis and discussion

In Table 1, columns 3–5 show the spatial frequencies, entropies and standard deviations of the source medical images, and columns 6–11 give the values of the different quantitative measures of the fused images obtained by the proposed MIF technique. The 'bold' values indicate the highest values in the Table 1 for that quantitative measure. The higher values of *SF* for the image combinations *C*1–*C*4 indicate that the fused images obtained by our proposed method have more activity and clarity level than the source images. Only the proton density weighted MR image

[Fig. 3(b5)] of image combination *C*5 has the higher value of SF than the fused image. The reason behind it may be that the CT image [Fig. 3(a5)] of the image combination *C*5 contains a thick whitish outer-boundary which become predominant in the fused result. Similarly the higher 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 the Table 1 that the standard deviation's values of the resultant images for 4 out of 5 source image combinations are higher than their corresponding source images, which indicates that the fused images obtained by our proposed MIF method have higher contrast than the corresponding source images. Only in case of image combination *C*2 the *STD* value of one of the source image Fig. 3(a2) (T1-weighted MR) is greater than the fused image. It may be because of the fact that the other source image Fig. 3(b2) (MRA) of the image combination *C*2 has very low contrast (indicated by low *STD* value) causing the fused image to have a lower STD value (lower by very small amount). Therefore, it is clear from Table 1 that the fused images obtained by the proposed MIF method are more clear, informative and have higher contrast which is helpful in visualization and interpretation.

In Table 2, the 'bold' values indicate the highest values. It is clear from the Table 2 that the proposed MIF technique has all the highest quantitative results except for mutual information (MI). The method of [16] has the highest value for the MI measure. It may be because of the fact that the method of [16] is based on m-PCNN in the spatial (pixel) domain. It preserves the information from both the source images better than our proposed method. But since our method is based on modified spatial frequency motivated PCNN in NSCT domain, hence it is superior in capturing the fine details of the source images into the fused image. The highest value of SF indicates that the fused image obtained by our proposed method has more activity and clarity level than the source images. 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 source images. It is also clear from the Table 2 that the fused image obtained by NSCT_PCNN_SF has lower quantitative results than the results obtained by the proposed MIF technique. For the other image combinations used in the experiments we have got similar kind of results.

Medical images of different modalities contain large amount of edges and directional features, which are quite often very subtle in nature. Through MIF we try to combine these complementary as well as contrasting features from source medical images into one fused image. Most of the existing state-of-the-art MIF techniques suffer from various problems of image degradations like contrast reduction, blocking effects and loss of image details etc., and most of these schemes are modality and task specific. This shortcoming is a big limit to the automatic process and the generalization for MIF techniques. The original spatial frequency (SF) lacks the directional information present in the image, which results in the loss of important fine details of the image. Whereas, modified spatial frequency (MSF) incorporates this directional information, and this results in an image clarity/activity level measure capable of capturing the fine details present in the image [26].

Keeping these above-mentioned issues in mind, in our proposed method we have used the shift invariance, multiscale and multi-directional properties of NSCT along with the modified spatial frequency motivated PCNN in such a way that can capture the subtle differences as well as the fine details present in the source medical images into the fused image without reducing the contrast. The use of modified spatial frequency along with the use of different fusion rules for different subbands produces fused images with higher spatial resolution and less unwanted degradations with less difference to the source images. Therefore, it is obvious from the results and comparisons given above that the fused images obtained by the proposed MIF method are more clear, informative and have higher contrast which is very helpful for the clinicians in their diagnosis and treatment.

Acknowledgments We would like to thank the editor, associate editor and the anonymous reviewers for their invaluable suggestions. We are grateful to Dr. Pradip Kumar Das (Medicare Images, Asansol-4, West Bengal) for the subjective evaluation of the fused images. 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. Barra V, Boire JY (2001) A general framework for the fusion of anatomical and functional medical images. NeuroImage 13(3):410–424
2. da Cunha A, Zhou J, Do M (2006) The nonsubsampled contourlet transform: theory, design, and applications. IEEE Trans Image Process 15(10):3089–3101
3. Das S, Chowdhury M, Kundu MK (2011) Medical image fusion based on ripplet transform type-I. Prog Electromagn Res B 30:355–370
4. Deepika MM, Vaithyanathan V (2012) An efficient method to improve the spatial property of medical images. J Theor Appl Inf Technol 35(2):141–148
5. Deng H, Ma Y (2009) Image fusion based on steerable pyramid and PCNN. In: Proceedings of 2nd international conference of applications of digital information and web technologies, pp 569–573
6. Eckhorn R, Reitboeck HJ, Arndt M, Dicke P (1990) Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex. Neural Comput 2(3):293–307
7. Eskicioglu A, Fisher P (1995) Image quality measures and their performance. IEEE Trans Commun 43(12):2959–2965

8. Feng K, Zhang X, Li X (2011) A novel method of medical image fusion based on bidimensional empirical mode decomposition. J Converg Inf Technol 6(12):84–91
9. Johnson J, Padgett M (1999) PCNN models and applications. IEEE Trans Neural Netw 10(3):480–498
10. Li H, Manjunath BS, Mitra SK (1995) Multi-sensor image fusion using the wavelet transform. CVGIP Graph Model Image Process 57(3):235–245
11. Li M, Cai W, Tan Z (2006) A region-based multi-sensor image fusion scheme using pulse-coupled neural network. Pattern Recogn Lett 27(16):1948–1956
12. Li S, Yang B (2010) Hybrid multiresolution method for multisensor multimodal image fusion. IEEE Sens J 10(9):1519–1526
13. Li S, Yang B, Hu J (2011) Performance comparison of different multi-resolution transforms for image fusion. Inf Fusion 12(2): 74–84
14. Qu GH, Zhang DL, Yan PF (2002) Information measure for performance of image fusion. Electron Lett 38(7):313–315
15. Tian H, Fu YN, Wang PG (2010) Image fusion algorithm based on regional variance and multi-wavelet bases. In: Proceedings of 2nd international conference of future computer and communication, vol 2, pp 792–795
16. Wang Z, Ma Y (2008) Medical image fusion using m-PCNN. Inf Fusion 9(2):176–185
17. Wang Z, Ma Y, Cheng F, Yang L (2010) Review of pulse-coupled neural networks. Image Vision Comput 28(1):5–13
18. Wang Z, Ma Y, Gu J (2010) Multi-focus image fusion using PCNN. Pattern Recogn 43(6):2003–2016
19. Xiao-Bo Q, Jing-Wen Y, Hong-Zhi X, Zi-Qian Z (2008) Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain. Acta Autom Sin 34(12):1508–1514
20. Xin G, Zou B, Li J, Liang Y (2011) Multi-focus image fusion based on the nonsubsampled contourlet transform and dual-layer PCNN model. Inf Technol J 10(6):1138–1149
21. Xydeas CS, Petrovic V (2000) Objective image fusion performance measure. Electron Lett 36(4):308–309
22. Yang L, Guo BL, Ni W (2008) Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. Neurocomputing 72(1-3):203–211
23. Yang S, Wang M, Lu Y, Qi W, Jiao L (2009) Fusion of multiparametric SAR images based on SW-nonsubsampled contourlet and PCNN. Signal Process 89(12):2596–2608
24. Yang Y, Park DS, Huang S, Rao N (2010) Medical image fusion via an effective wavelet-based approach. EURASIP J Adv Signal Process 44:1–13
25. Yonghong J (1998) Fusion of landsat TM and SAR images based on principal component analysis. Remote Sens Technol Appl 13(3):46–49
26. Zheng Y, Essock EA, Hansen BC, Haun AM (2007) A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. Inf Fusion 8(2):177–192