CT and MR Image Fusion Scheme in Nonsubsampled Contourlet Transform Domain

Padma Ganasala · Vinod Kumar

Abstract Fusion of CT and MR images allows simultaneous visualization of details of bony anatomy provided by CT image and details of soft tissue anatomy provided by MR image. This helps the radiologist for the precise diagnosis of disease and for more effective interventional treatment procedures. This paper aims at designing an effective CT and MR image fusion method. In the proposed method, first source images are decomposed by using nonsubsampled contourlet transform (NSCT) which is a shift-invariant, multiresolution and multidirection image decomposition transform. Maximum entropy of square of the coefficients with in a local window is used for low-frequency sub-band coefficient selection. Maximum weighted sum-modified Laplacian is used for high-frequency sub-bands coefficient selection. Finally fused image is obtained through inverse NSCT. CT and MR images of different cases have been used to test the proposed method and results are compared with those of the other conventional image fusion methods. Both visual analysis and quantitative evaluation of experimental results shows the superiority of proposed method as compared to other methods.

Keywords Image fusion · Nonsubsampled Contourlet Transform · X-ray computed tomography · Magnetic resonance imaging · Image visualization

Introduction

With the advent of modern technology, there is a there is a tremendous improvement in the capabilities of medical imaging systems and an increase in the number of imaging

P. Ganasala (✉) · V. Kumar
Department of Electrical Engineering, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India 247667
e-mail: padma_417@yahoo.co.in

modalities in clinical use. Each medical imaging modality gives the specific information about the human body and the same is not available with other imaging modality. Medical imaging modalities can be classified into two types, i.e. functional imaging modalities and anatomical imaging modalities. Functional imaging modalities like positron emission tomography (PET), single positron emission computed tomography (SPECT), and functional magnetic resonance imaging (FMRI) gives the metabolic information or physiologic information of the human body. Whereas anatomical imaging modalities like X-ray computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound imaging (US) gives mainly structural information of the human body. CT images provides electron density map required for accurate radiation dose estimation and superior cortical bone contrast; however, it is limited in soft tissue contrast. MRI provides excellent soft tissue contrast which permits better visualization of tumors or tissue abnormalities in different parts of the body. But MRI has lack of signal from cortical bone and has image intensity values that have no relation to electron density. For the precise diagnosis of disease and for more effective interventional treatment procedures, radiologists need the information from two or more imaging modalities [1]. Through image fusion, it is possible to integrate and present the information from two or more imaging modalities in a more effective way. Image fusion finds applications in Oncology, Neurology, Cardiology and others [2, 3]. Fusion of CT and MR images is used to improve lesion delineation for Radiation therapy planning, prostate seed implant quality analysis [1], and planning the correct surgical procedure in computer-assisted navigated neurosurgery of temporal bone tumors [4], and orbital tumors [5]. This paper aims at designing an efficient CT and MR image fusion method.

Image fusion is process of integrating information from two or more images into a single composite image which is more suitable for human visual perception and further computer processing tasks [6]. Image fusion process must retain

both redundant and complementary information present in the source images and it should not introduce any artefacts into the fused image. Depending on the merging stage, image fusion can be classified into three categories, i.e. pixel level, feature level and decision level as shown in Fig. 1. Pixel level image fusion directly combines the pixel data of source images to obtain fused image. It needs perfect registration of source images to sub pixel accuracy. Feature level fusion involves the extraction of representative features present in the source (e.g. by using segmentation) and then combining those features into a single feature vector by using neural networks, clustering algorithm or template methods [7]. Decision level is a high level fusion method in which first source images are processed individually for information extraction and then extracted information is combined (e.g. by using voting method). Compared to others, pixel level image fusion is more computationally efficient. Pixel level image fusion can be done in spatial domain or in transform domain (multiscale decomposition-based image fusion). Few spatial domain pixel level image fusion techniques are simple averaging of source images and principle component analysis (PCA) method. Transform domain pixel level image fusion methods are more efficient compared to spatial domain because it is possible to analyze the images at different resolutions through multiscale decomposition (MSD) transformation. Features which are sensitive to human visual system (HVS) are present at different resolutions or scales of images.

The work presented in this paper is transform domain/ multiscale decomposition-based image fusion method. Wide variety of MSD-based image fusion techniques have been proposed by various researchers. Key steps in MSD-based image fusion methods are as follows. First step is

decomposing the source images with a MSD transform into low- and high-frequency sub-bands at different resolutions and spatial orientations. Second step is combining the coefficients of different sub-bands of source images by using a fusion rule. Third step is taking the inverse MSD of composite coefficient to obtain the fused image. Mainly the quality of fused image depends on the two factors: MSD transform used for decomposition and the fusion rule used to combine the coefficients. Initially Toet, [6] and Toet et al. [8] introduced different pyramid schemes for multi sensor image fusion. Pyramid scheme failed to provide any spatial orientation selectivity in the decomposition processes. Hence, it often causes blocking artefacts in the fused image. Next, discrete wavelet transform (DWT) has been introduced for image fusion by Munjunath et al. [9]. An enhancement of DWT called Dual–Tree complex wavelet Transform which is shift-invariant was proposed by N.G. Kingsbury has been used for image fusion [10]. DWT can provide only limited directional information, i.e. horizontal, vertical and diagonal information. It cannot capture the smoothness along the contours of the images and hence often causes artefacts along the edges. Next advanced MSD transforms like curvelet [11–13], ripplet [14, 15], bandelet transform [16], shearlet transform [17], and contourlet transform [18, 19] which can provide more directional information have been used for image fusion. But these transforms lack shift-invariance and causes pseudo-Gibbs phenomena around the singularities. Shift-invariant version of Contourlet transform called nonsubsampled contourlet transform (NSCT) which is proposed by da Cunha et al. [20] has been used for image fusion in different applications. NSCT gives better performance for medical image fusion due to flexible multiscale, multidirection, and shift-invariant image decomposition [21, 22]. In the proposed method, NSCT is used for multiscale decomposition of source images.

Other important factor which influences the fusion quality is the fusion rule used for combining the coefficients of different sub-bands. Since low- and high-frequency sub-bands carries different information of source images, different fusion rules are used for combing LF sub-band and HF sub-bands. LF sub-band is smoothed version of original image and it represents the outline of the image. HF sub-bands represents details like edges and contours of original image. The basic fusion rule used for LF sub-band is averaging the source image sub-bands. It has a serious drawback, i.e. reduction in contrast and hence possible cancellation of few patterns in source images. Most commonly used fusion rule for high-frequency sub-band is selecting the source image coefficient having absolute maximum value. This scheme is sensitive to noise. These simple schemes which combine the coefficients based on single coefficient value may not retain the important information present in the source images because image features sensitive to human visual system are not completely defined by single pixel or coefficient. Hence window- or region-based activity level measurement at
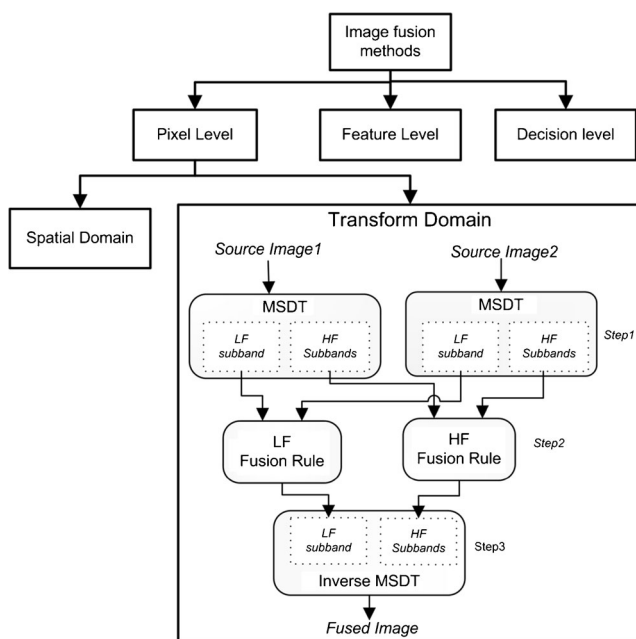

Fig. 1 Classification of image fusion methods

each coefficient is done and this information is used for coefficient combination. The coefficient combinations schemes used are (1) choose a maximum scheme in which source image coefficient having maximum activity level measurement is selected at each location, and (2) weighted averaging in which coefficient weights are calculated based on its activity level measurement value, and (3) a hybrid scheme which includes both of the above two schemes based on match measure value at that location. Comprehensive overview of possible fusion rules are given in references [23, 24].

Statistical parameters and texture features are used as activity level measurement parameters. For low-frequency coefficient fusion, features like energy [22], visibility [25], weighted energy, entropy [26], and spatial frequency [27], etc. are used as activity measurements by researchers. For high-frequency coefficient fusion features like contrast [28], gradient, variance [25], sum-modified laplacian (SML) [22], energy of gradient, and energy of laplacian, etc. are used as activity level measurement parameters by researchers. In the proposed method, new activity level measurement parameter is used for low-frequency sub-band fusion. Weighted sum of modified laplacian is used for high-frequency sub-band fusion.

The remaining paper is organised as follows: "Nonsubsampled Contourlet Transform" gives brief overview of NSCT, "Proposed Fusion Scheme" describes the proposed fusion rule, "Experimental Results and Comparative Analysis" presents the experimental results and comparative analysis and finally "Conclusions" gives the conclusions.

Nonsubsampled Contourlet Transform

NSCT is a shift-invariant, multiscale, and multidirection image decomposition transform. Its design is based on nonsubsampled pyramid structure (NSP) and nonsubsampled directional filter bank (NSDFB). NSP ensures the multiscale feature and NSDFB ensures multi direction feature of NSCT. Shift-invariance of NSCT is obtained by eliminating upsamplers and downsamplers both in NSP and NSDFB. The NSCT is constructed by combining the NSP and the NSDFB as shown in Fig. 2a.

Nonsubsampled Pyramid

The basic building block of NSP is a two-channel filter bank without upsamplers and downsamplers and its ideal frequency response is as shown in Fig. 2b. $H_k(Z)$ (k=0,1) are the first-stage analysis filters and $G_k(Z)$ (k=0,1) are the synthesis filters. Each stage of NSP produces one low-pass filtered image (y_0) and one bandpass-filtered image (y_1). To get the subsequent stages of decomposition, low-frequency sub-band is filtered iteratively. The two-stage decomposition structure of NSP is shown in Fig. 2a. The filters for subsequent stages are obtained by upsampling the filters of the previous stage. This gives the multiscale property without the need for additional filter design. First-stage low-pass and bandpass filters are denoted as $H_0(Z)$ and $H_1(Z)$ respectively and second stage low-pass and bandpass filters are $H_0(Z^2)$ and $H_1(Z^2)$, respectively.

Nonsubsampled Directional Filter Bank

The bandpass images from NSP structure are fed to NSDFB for directional decomposition. The basic building block of NSDFB is a two-channel fan filter bank and its ideal frequency response is as shown in Fig. 2c. $U_k(Z)$ (k=0,1) are the analysis filters and $V_k(Z)$ (k=0,1) are the synthesis filters. To get more directional decomposition, this two-channel filter

Fig. 2 Nonsubsampled contourlet transform. **a** NSFB structure that implements the NSCT. **b** Two-channel NSP filter bank. **c** Two-channel NSDFB. **d** Four-channel analysis NSDFB structure. **e** Frequency portioning obtained with NSCT
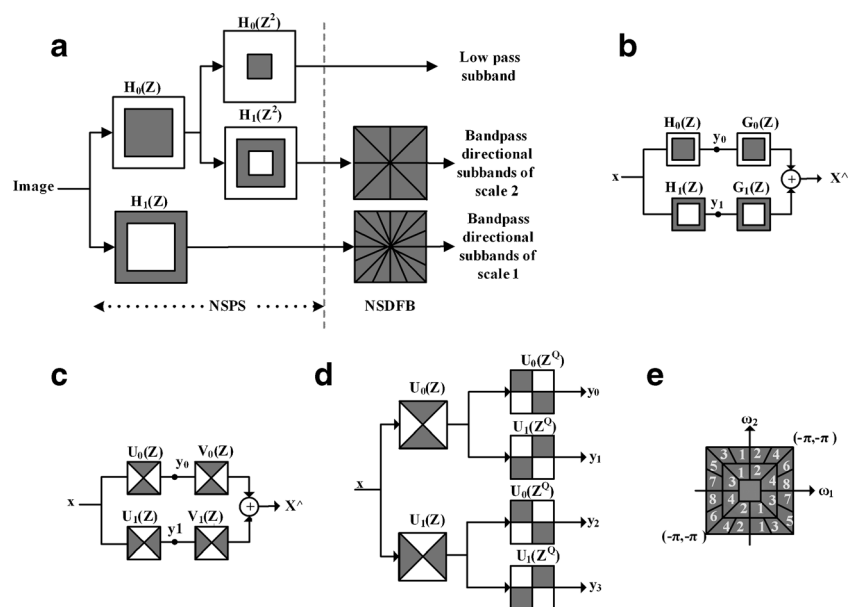
bank is iterated like tree structure after upsampling all filters by a quincunx matrix given by

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad (1)$$

The second stage synthesis filters are denoted as $U_k(Z^Q)$ ($k=0,1$) and these have checker-board frequency support. The two-stage analysis NSDFB structure which gives four directional sub-bands (y_k, $k=0,1,2,3$) is shown in Fig. 2d. The resulting structure divides the 2D frequency plane into directional wedges. The L stage NSDFB produces 2^L directional sub-bands. The NSCT is flexible in allowing any number of directions in each scale. The frequency partitioning with eight and four directional sub-bands decomposition in scales 1 and 2 is shown in Fig. 2e in which w_1 and w_2 represents the frequency in two dimensions.

Proposed Fusion Scheme

Different steps involved in the proposed image fusion method are represented in the Fig. 3. The first step is to decompose the source image into different resolutions and different directions by using NSCT. As more number of decomposition levels introduces artefacts into the fused image, source images are decomposed into two levels with eight and four directional sub-bands in the first and second decomposition level respectively. Then low-frequency sub-band and directional sub-bands (high-frequency sub-bands) are combined by using different fusion rules as discussed in the following subsections. Finall,y fused image is obtained by taking the inverse NSCT of the composite fused MSD coefficients.

Low-Frequency Sub-Band Fusion

Low-frequency sub-band is smoothed version of original image. It represents the outline of the image. As the number of decomposition levels are restricted to two in this work, most of the signal energy and few details of the original image still present in the low-frequency sub-band of the image. Hence, it is important fuse the low-frequency sub-band in such a way to retain both the detailed information as well as approximate information present in it. In the proposed method, activity measure used for low-frequency sub-band fusion is entropy of square of the coefficients within a 3×3 window. It is given by the following equation.

$$a_L^A(m,n) = \sum_{i=-1}^{i=1}\sum_{j=-1}^{j=1} C_L^{A\,2}(m+i,n+j)\log\left(C_L^{A\,2}(m+i,n+j)\right)/9 \qquad (2)$$

Where $C_L^A(m,n)$ is the low-frequency sub-band coefficient of source image A at location (m,n). Similarly for the source image B, activity of low-frequency coefficient $C_L^B(m,n)$ at location (m,n) is given by

$$a_L^B(m,n) = \sum_{i=-1}^{i=1}\sum_{j=-1}^{j=1} C_L^{B\,2}(m+i,n+j)\log\left(C_L^{B\,2}(m+i,n+j)\right)/9 \qquad (3)$$
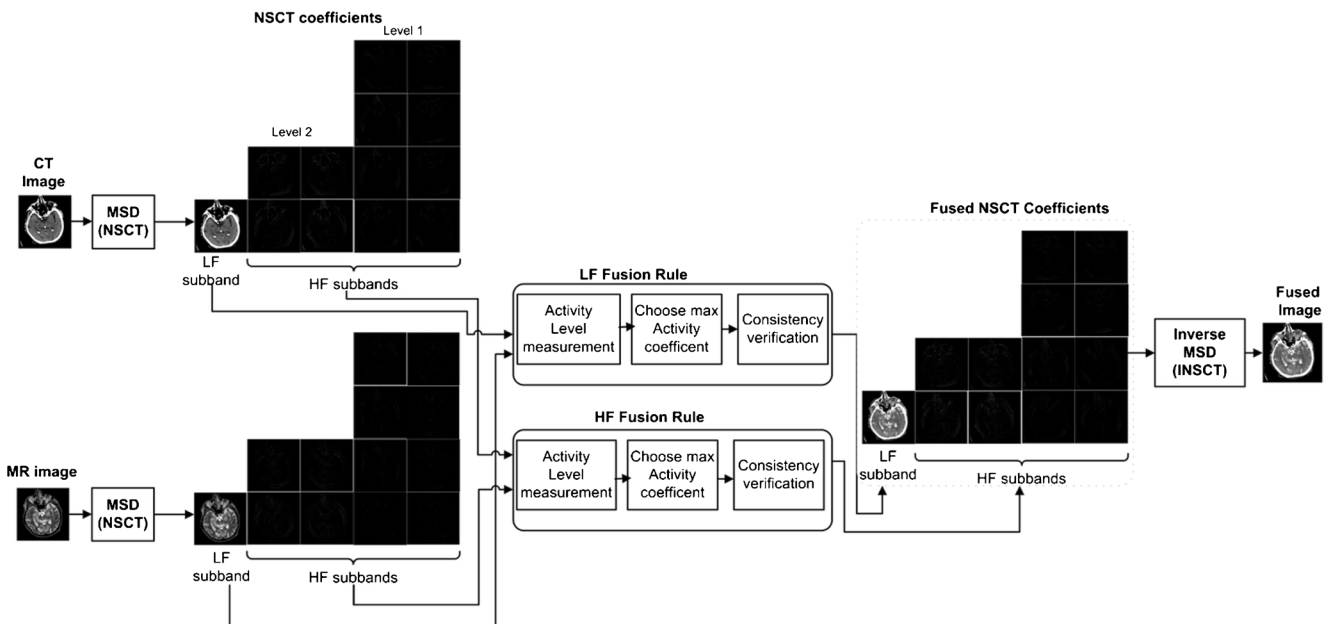

Fig. 3 Block diagram of the proposed image fusion method

Initial fusion decision map is obtained by choose max combination scheme, i.e. selecting the coefficient having the maximum activity measure as follows:

$$d_i(m,n) = \begin{cases} 1 & if\ a_L^A(m,n) \geq a_L^B(m,n) \\ 0 & if\ a_L^A(m,n) < a_L^B(m,n) \end{cases} \tag{4}$$

Which implies that if $d_i(m,n)=1$ then image A coefficient is to be selected at (m, n) location and if $d_i(m,n)=0$ then image B coefficient is to be selected at (m, n) location. Then final fusion decision map (d_f) is obtained through consistency verification in a 3×3 window by using majority filtering operation. That is, in each 3×3 window, if more number of coefficients are from image A, whereas centre coefficient is from B, then centre coefficient is also made to come from image A. Otherwise it is kept as it is and vice versa. This verification is done at each coefficient. This is to make the neighbouring coefficients in the composite MSD belong to the same source image in order to overcome the effect due to noise and guarantee the homogeneity of the fused image. Then, fused low-frequency sub-band coefficients ($C_L^F(m,n)$) are

calculated by using final fusion decision map (d_f) as follows.

$$C_L^F(m,n) = \begin{cases} C_L^A(m,n) & if\ d_f(m,n) = 1 \\ C_L^B(m,n) & if\ d_f(m,n) = 0 \end{cases} \tag{5}$$

High-Frequency Sub-Band Fusion Rule

High-frequency sub-bands represent the detailed component of the source images such as edges, contours, and object boundaries. The most commonly used fusion rule for high-frequency sub-band is selecting the coefficient having absolute maximum value. But this scheme is sensitive to noise and also there is possibility to lose some important information as the coefficient selection is based on single coefficient value without considering neighbouring coefficients. Another scheme used is coefficient selection based on activity level measurement value. In the proposed method, weighted sum-modified laplacian (WSML) is used as activity level measurement parameter for high-frequency sub-band coefficients. The complete expression for WSML is as follows:

Modified Laplacian of $f(x,y)$ is

$$ML_f(x,y) = |2f(x,y)-f(x-1,y)-f(x+1,y)| + |2f(x,y)-f(x,y-1)-f(x,y+1)| \tag{6}$$

WSML of $f(x,y)$ is

$$WSML[f(x,y)] = \sum_{i=-1}^{1}\sum_{j=-1}^{1} w(i+1,j+1).ML_f(x+i,y+j) \tag{7}$$

where w is the weight matrix. In the proposed method, city block distance weight matrix is used. That is

$$w = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{8}$$

WSML is calculated at each high-frequency sub-band coefficient of image A and B as their activity measure.

$$a_{d,k}^A(m,n) = WSML\left[C_{d,k}^A(m,n)\right] \tag{9}$$

Where $C_{d,k}^A(m,n)$ is dth level, Kth directional sub-band coefficient of image A at location (m,n).

$$a_{d,k}^B(m,n) = WSML\left[C_{d,k}^B(m,n)\right] \tag{10}$$

where $C_{d,k}^B(m,n)$ is dth level, Kth directional sub-band coefficient of image B at location (m,n). Initial fusion decision map is obtained by the chosen maximum activity coefficient scheme.

$$d_i(m,n) = \begin{cases} 1 & if\ a_{d,k}^A(m,n) \geq a_{d,k}^B(m,n) \\ 0 & if\ a_{d,k}^A(m,n) < a_{d,k}^B(m,n) \end{cases} \tag{11}$$

Then final fusion decision map (d_f) is obtained through consistency verification as discussed in low-frequency fusion rule. Fused high-frequency sub-bands coefficients are calculated from the final fusion decision map.

$$C_{d,k}^F(m,n) = \begin{cases} C_{d,k}^A(m,n) & if\ d_f(m,n) = 1 \\ C_{d,k}^B(m,n) & if\ d_f(m,n) = 0 \end{cases} \tag{12}$$

Experimental Results and Comparative Analysis

The proposed image fusion method has been tested on different cases of CT and MR images. Dataset-1 consists of one pair

of CT and MRI brain images (shown in Fig. 4a, b) which are collected from www.imagefusion.org. The dataset-2 consists of nine pairs of CT and MR brain images corresponding to various pathologies (shown in Fig. 5a, b columns). These images are collected from Harvard university site (http://www.med.harvard.edu/AANLIB/home.html). The proposed image fusion method is compared with other image fusion methods like (1) Pixel averaging method (Pixe_ avg), (2) DWT, (3) CT, and (4) NSCT domain image fusion methods with basic fusion rule, i.e. averaging the low-frequency sub-band and selecting the absolute maximum for high-frequency sub-bands ((2) DWT_avg_max, (3) CT_avg_max, and (4) NSCT_avg_max) [21]. Experimental results are shown in Figs. 4 and 5.

Visual analysis of experimental results reveals that proposed method is retaining both clear bony structure of CT image and soft tissue details of the MR image with good contrast and without introducing any artefacts into the fused image.

Quantitative evaluation of proposed method is done with well-defined image fusion quality metrics like (1) Information Entropy (IE), (2) Overall Cross entropy (OCE) [19], (3) Spatial Frequency (SF), (4) Ratio of spatial frequency error (RSFE) [29], (5) Mutual Information (MI) [30], (6) Cross Correlation coefficient (CC) [19], (7) Xydias and Petrovic metric ($Q^{AB/F}$)[31], (8) Universal Image Quality Index (UIQI) based metrics (a) Q, (b) Q_W, and (c) Q_E [32].

1. IE: Information entropy measures the amount of information present in an image. An image with high information content will have high entropy. Based on the principle of Shannon information theory, the IE of an image is given by the formula.

$$IE = -\sum_{i=0}^{L-1} P_F(i)\log_2 P_F(i) \qquad (13)$$

Where $P_F(i)$ the ratio of the number of the pixels with gray value is equal to i over the total number of pixels in the fused image and L is the maximum gray value of the fused image. It is set to 256 in our case. Larger entropy value implies better fusion quality.

2. Overall cross entropy (OCE): Cross entropy measures the difference between two source images and the fused image. It is given by the formula:

$$OCE(I_A, I_B; I_F) = \frac{CE(I_A, I_F) + CE(I_B, I_F)}{2} \qquad (14)$$

Where I_A, I_B are source images and I_F is the fused image.

$$CE(I_A, I_F) = \sum_{i=0}^{L-1} P_A(i)\log_2 \left| \frac{P_A(i)}{P_F(i)} \right| \qquad (15)$$

$$CE(I_B, I_F) = \sum_{i=0}^{L-1} P_B(i)\log_2 \left| \frac{P_B(i)}{P_F(i)} \right| \qquad (16)$$

Small value of OCE corresponds to good fusion quality.

3. SF: Spatial frequency reflects the activity level and clarity of an image. It is defined as follows:

$$SF = \sqrt{RF^2 + CF^2} \qquad (17)$$

Where RF is row frequency:

$$RF = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=2}^{N}[I_F(i,j) - I_F(i,j-1)]^2} \qquad (18)$$

and CF is column frequency:

$$CF = \sqrt{\frac{1}{MN}\sum_{j=1}^{N}\sum_{i=2}^{M}[I_F(i,j) - I_F(i-1,j)]^2} \qquad (19)$$

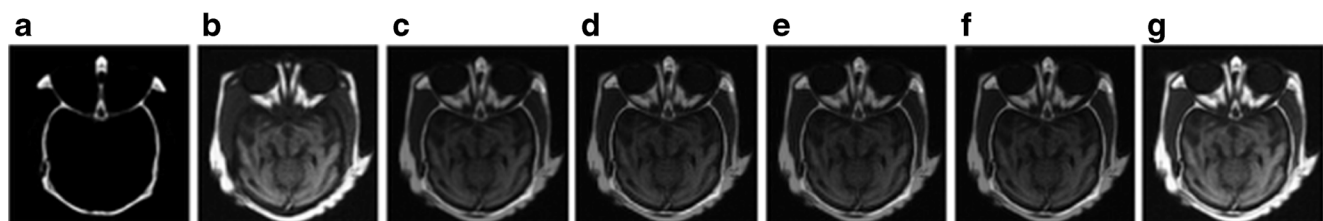Larger spatial frequency value denotes better the fusion quality.

Fig. 4 Comparison of different image fusion methods using brain images (Dataset-1). (*a*) CT image (*b*) MR image (*c*) fused image by Pixel_avg (*d*) fused image by DWT_avg_max (*e*) fused image by CT_avg_max (*f*) fused image by NSCT_avg_max (*g*) fused image by the proposed method
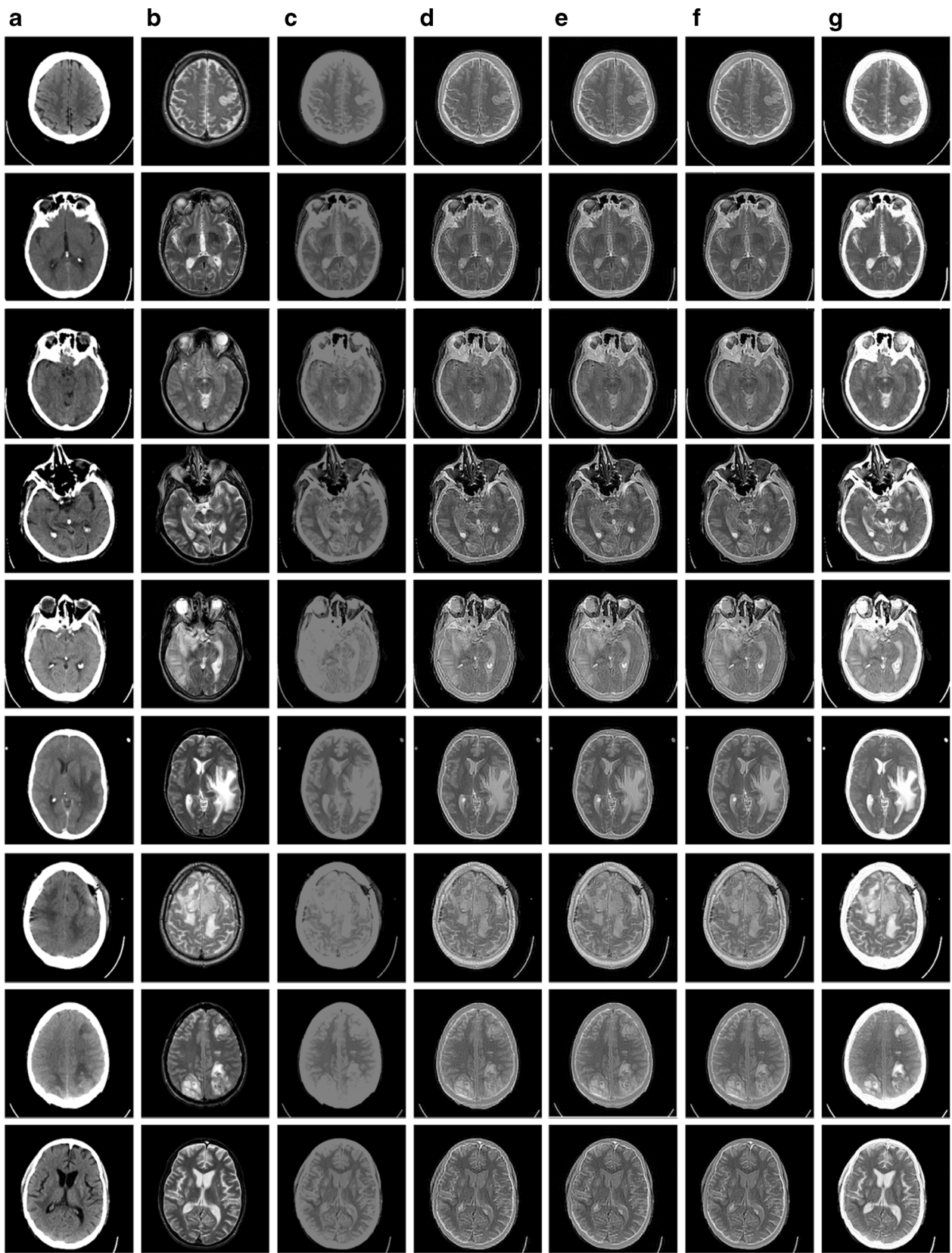
Fig. 5 Comparison of different image fusion methods using brain images of Dataset-2. Columns (*a*) CT images (*b*) MR images (*c*) fused images by Pixel_ avg (*d*) fused images by DWT_avg_max (*e*) fused images by CT_avg_max (*f*) fused images by NSCT_avg_max (*g*) fused images by the proposed method

4. RSFE: Spatial frequency error gives the difference between activity of fused image and ideal fused reference image. It is given by the following formula:

$$\text{RSFE} = \frac{(\text{SF}_F - \text{SF}_R)}{\text{SF}_R} \qquad (20)$$

$$SF_F = \sqrt{RF^2 + CF^2 + MDF^2 + SDF^2} \qquad (21)$$

RF is row frequency and CF is column frequency is as defined above. MDF and SDF are main diagonal and secondary diagonal frequencies which are calculated as follows. All these are basically first order gradients along the four directions.

$$\text{MDF} = \sqrt{w_d \frac{1}{MN} \sum_{i=2}^{M} \sum_{j=2}^{N} \left[I_F(i,j) - I_F(i-1,j-1) \right]^2} \qquad (22)$$

$$\text{SDF} = \sqrt{w_d \frac{1}{MN} \sum_{j=1}^{N-1} \sum_{i=2}^{M} \left[I_F(i,j) - I_F(i-1,j+1) \right]^2} \qquad (23)$$

where $w_d = \frac{1}{\sqrt{2}}$

SF_R is the reference spatial frequency calculated by taking maximum gradients of input images along four directions.

$$\text{Grad}^D(I_R(i,j)) = max\left\{ \text{abs}\left[\text{Grad}^D(I_A(i,j)) \right], \text{abs}\left[\text{Grad}^D(I_B(i,j)) \right] \right\} \qquad (24)$$

for each of four directions, i.e. $D=\{H, V, MD, SD\}$.

An ideal fusion has RSFE equal to zero. Smaller RSFE absolute value corresponds to better fusion quality. Furthermore, RSFE>0 means that distortion or noise is introduced into the fused image and RSFE<0 denotes that some meaningful information is lost in the fused image.

5. MI: Mutual information is the amount of information that one image contains about another. Considering two source images A, B, and fused image F, the amount of information that F contains about A and B can be calculated as

$$I_{FA}(f,a) = \sum_{f,a} P_{FA}(f,a) \log \frac{P_{FA}(f,a)}{P_F(f)P_A(a)} \qquad (25)$$

$$I_{FB}(f,b) = \sum_{f,b} P_{FB}(f,b) \log \frac{P_{FB}(f,b)}{P_F(f)P_B(b)} \qquad (26)$$

Thus the image fusion performance measure Mutual Information (MI) can be defined as

$$\text{MI}(F,A,B) = I_{FA}(f,a) + I_{FB}(f,b) \qquad (27)$$

The larger the MI value, the better the image fusion quality.

6. CC: Correlation coefficient can show similarity in the small structures between the input image and the fused image. A higher value of correlation means that more information is preserved. CC between input image I_A and the fused image I_F is given by the following equation:

$$CC(I_F, I_A) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(I_F(i,j) - \overline{I}_F \right) \left(I_A(i,j) - \overline{I}_A \right)}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(I_F(i,j) - \overline{I}_F \right)^2 \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(I_A(i,j) - \overline{I}_A \right)^2}} \qquad (28)$$

where $\overline{I}_F$ and $\overline{I}_{FA}$ are the mean values of the corresponding images. Similarly, correlation coefficient between image I_B and the fused image I_F can be calculated.

7. Xydeas and Petrovic Metric ($Q^{AB/F}$): Xydeas and Petrovic, proposed an objective performance metric, which measures the relative amount of edge information that is transferred from the source images A and B into the fused image F. This method uses Sobel edge detector to calculate the edge strength $g(n,m)$ and orientation

$\alpha(n,m)$ information for each pixel $p(n,m)$. Thus for an input image A:

$$g_A(n,m) = \sqrt{S_A^x(n,m)^2 + S_A^y(n,m)^2} \qquad (29)$$

$$\propto_A(n,m) = \tan^{-1}\left(\frac{S_A^y(n,m)}{S_A^x(n,m)} \right) \qquad (30)$$

Table 1 Quantitative comparison of five image fusion methods with Dataset-1 image pair (shown in Fig. 4a, b)

Quality metric	Pixel_avg	DWT_avg_max	CT_avg_max	NSCT_avg_max	Proposed method
IE	5.911	6.096	6.199	6.065	*6.756*
OCE	0.445	0.488	0.468	0.496	*0.009*
MI	5.166	3.073	2.748	3.318	*5.616*
SF	10.283	14.282	14.057	13.971	*18.468*
RSFE	−0.5	−0.31	−0.322	−0.325	*−0.102*
CC(A,F)	0.846	0.823	0.826	0.826	*0.909*
CC(B,F)	0.575	0.587	0.585	*0.591*	0.426
$Q^{AB/F}$	0.425	0.515	0.477	0.565	*0.784*
Q	0.628	0.64	0.567	0.665	*0.934*
Q_w	0.631	0.66	0.652	0.669	*0.845*
Q_E	0.584	0.613	0.602	0.627	*0.735*

Where $S_A^x(n,m)$ and $S_A^y(n,m)$ are the output of the horizontal and vertical Sobel templates centred on pixel $p_A(n,m)$ and convolved with the corresponding pixels of image A. The relative strength and orientation values of $G^{AF}(n,m)$ and $A^{AF}(n,m)$ of an input image A with respect to F are formed as

$$G^{AF}(n,m) = \begin{cases} \dfrac{g_F(n,m)}{g_A(n,m)}, & \text{if } g_A(n,m) > g_F(n,m) \\ \dfrac{g_A(n,m)}{g_F(n,m)}, & \text{otherwise} \end{cases} \quad (31)$$

$$A^{AF}(n,m) = \frac{||\propto_A(n,m) - \propto_F(n,m)| - \pi/2|}{\pi/2} \quad (32)$$

The edge strength and orientation preservation values are

$$Q_g^{AF}(n,m) = \frac{\Gamma_g}{1 + e^{K_g\left(G^{AF}(n,m) - \sigma_g\right)}} \quad (33)$$

$$Q_\propto^{AF}(n,m) = \frac{\Gamma_\propto}{1 + e^{K_\propto\left(A^{AF}(n,m) - \sigma_\propto\right)}} \quad (34)$$

Edge information preservation values are then defined as

$$Q^{AF}(n,m) = Q_g^{AF}(n,m) Q_\propto^{AF}(n,m) \quad (35)$$

Then fusion performance metric $Q^{AB/F}$ is obtained as follows:

$$Q^{AB/F} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M} Q^{AF}(n,m)w^A(n,m) + Q^{BF}(n,m)w^B(n,m)}{\sum_{n=1}^{N}\sum_{m=1}^{M}(w^A(n,m) + w^B(n,m))} \quad (36)$$

Where $w^A(n,m) = [g_A(n,m)]^L$ and $w^B(n,m) = [g_B(n,m)]^L$ are the weights and L is a constant. The range of $Q^{AB/F}$ is [0 1]. A value of 0 corresponds to the complete loss of edge

Table 2 Quantitative comparison of five image fusion methods with *Data set-2* image pairs (shown in Fig. 5a and b columns)

Quality metric	Pixel_avg	DWT_avg_max	CT_avg_max	NSCT_avg_max	Proposed method
IE	3.296	4.834	*5.213*	4.726	4.521
OCE	*0.472*	0.544	0.595	0.548	0.59
MI	3.08	2.536	2.473	2.652	*3.113*
SF	14.827	29.117	28.961	28.706	*32.704*
RSFE	−0.6	−0.217	−0.223	−0.229	*−0.116*
CC(A,F)	*0.798*	0.781	0.778	0.784	0.691
CC(B,F)	0.88	0.897	0.898	0.9	*0.928*
$Q^{AB/F}$	0.213	0.466	0.443	0.528	*0.535*
Q	0.692	0.676	0.472	0.744	*0.783*
Q_w	0.471	0.679	0.675	0.696	*0.768*
Q_E	0.371	0.51	0.504	0.531	*0.572*

Values given in Table 2 are the average values of different images pairs of Dataset-2

information, as transferred from A and B into F. A value 1 indicates fusion of A and B to F is with no loss of information.

8. UIQI-based metrics (Q, Q_W, Q_E): Gemma Piella proposed three image fusion quality metrics based on UIQI which was proposed by Wang and Bovik [33]. UIQI quantifies the structural distortion between two images in a local region as:

$$Q_0(A, F/w) = \frac{4\sigma_{af}\bar{a}\bar{f}}{\left(\bar{a}^2 + \bar{f}^2\right)\left(\sigma_a^2 + \sigma_f^2\right)} \quad (37)$$

Where σ_a, σ_f are the variance of image A, F and $\bar{a}, \bar{f}$ are the mean of image A, F.

Image fusion quality metrics based on UIQI are given by

$$Q(A, B, F) = \frac{1}{|W|}\sum_{w \in W}(\lambda_a(w)Q_0(a, f/w) + \lambda_b(w)Q_0(b, f/w)) \quad (38)$$

$$\lambda_a(w) = \frac{s(a/w)}{s(a/w) + s(b/w)}, \lambda_b(w) = 1 - \lambda_a(w) \quad (39)$$

$S(a/w)$ and $S(b/w)$ are local saliency of image A and B respectively in the window w. In this paper variance is considered as local saliency.

$$Q_w(A, B, F) = \sum_{w \in W}c(w)(\lambda_a(w)Q_0(a, f/w) + \lambda_b(w)Q_0(b, f/w)) \quad (40)$$

Where $c(w) = \mathbf{C}(w)/\sum_{w' \in W}\mathbf{C}(w')$ and $\mathbf{C}(w) = max(s(a/w), s(b/w))$

$$c)Q_E(A, B, F) = Q_w(A, B, F)^{1-\alpha}Q_w(A', B', F')^{\alpha}, \alpha \in [0 \quad 1] \quad (41)$$

where $A', B',$ and F' are edge images of A, B, and F, respectively. All the above three image fusion quality measures (Q, Q_W, Q_E) have a dynamic range of $[-1 \quad 1]$. The closer the value to 1, higher is the quality of the fused image.

Tables 1 and 2 show the quantitative evaluation of image fusion methods for Dataset-1 and Dataset-2, respectively. The average value of each quality metric for all images of Dataset-2 is given in Table 2. The best value of each quality metric is highlighted both in Tables 1 and 2. The proposed method

gives the highest value of mutual information, which implies that amount information transferred from source images to the fused images is the maximum in the proposed method. The maximum value for SF and minimum value for absolute value of RSFE quality metrics were given by the proposed method which means that detail information present the source images is better preserved in proposed method than the other methods. The proposed method gives the highest value for $Q^{AB/F}$ compared to the other methods which implies that amount of edge information transferred from source images to the fused image is more in the proposed method. The proposed method also shows the best performance with respect to universal image quality index-based metrics, which implies that structural distortion between source images and fused image is less in the proposed method than other methods. The proposed method gives a comparatively high value for information entropy quality metric. This implies that the amount of information present in the fused image given by proposed method is comparatively high. The minimum overall cross entropy is given by the proposed method in the case of Dataset-1 whereas pixel averaging method is given by the minimum overall cross entropy value for Dataset-2. Correlation coefficient value between two source images and fused images is comparatively high in case of the proposed method, which implies that similarity of fused image with both the source images is comparatively good in the case of the proposed method.

Further, proposed method is compared with two other methods, (1) energy- and contrast-based fusion rule in Contourlet Transform domain (CT-Energy-contrast) [19] and (II) energy- and SML-based fusion rule in Contourlet

Table 3 Quantitative comparison of proposed method with energy- and contrast-based fusion rule in contourlet transform domain [19] by using Dataset-1

Quality metric	CT_Energy_contrast	Proposed method
EN	6.3877	*6.756*
OCE	1.3159	*0.009*
SF	6.5575	*18.468*
CC(A; F)	0.8077	*0.909*
CC(B; F)	*0.5795*	0.426

Table 4 Quantitative comparison of proposed method with Energy and SML-based fusion rule in Contourlet transform domain [22] by using Dataset-1

Quality metric	CT_Energy_SML	Proposed method
Q	0.9291	*0.934*
Q_W	*0.8912*	0.845
Q_E	0.7307	*0.735*

Transform domain (CT-Energy-SML) [22] by using Dataset-1. Comparison with CT-Energy-contrast method is done with respect to entropy (EN), overall cross entropy (OCE), spatial frequency (SF), and cross correlation coefficient (CC(A,F), CC(B,F)) quality metrics as followed in [19]. Table 3 shows the comparison of this method with the proposed method. The proposed method gives high values for EN, SF, and CC compared to CT-Energy-contrast method. This implies that more details are preserved in the proposed method. Low value of OCE is given by the proposed method compared to CT-Energy-contrast method indicates that similarity of fused image with source images is more. Hence, the performance of the proposed method is better than CT-Energy-contrast method. Universal Image Quality Index-based metrics are used for comparison of proposed method with CT-Energy-SML method as followed in [22]. These results are shown in Table 4. The proposed method gives slightly better values for Q and Q_E than the CT-Energy-SML method. Thus, structural distortion is less and the amount of edge information transferred is more in the proposed method. Hence, the performance of proposed method is comparable or slightly better than energy and SML-based fusion rule.

The quantitative evaluation results show the superiority of the proposed method compared to other methods and also these are consistent with the visual analysis results. Hence, the proposed fusion rule in NSCT domain is suitable for CT and MR image fusion.

Conclusions

An efficient CT and MR image fusion scheme in NSCT domain is proposed. A novel window based activity level measurement parameters are used for low- and high-frequency sub-bands fusion. Proposed method is compared with spatial domain averaging method, Discrete Wavelet Transform-based method, Contourlet Transform-based method, NSCT-based method with basic fusion rule, and also with Contourlet Transform-based methods with two different fusion rules. Proposed method has been tested on CT and MR brain images of different cases. Quantitative evaluation results demonstrate that the proposed method is superior to several existing methods compared in this paper. Visual analysis of experimental results reveal that proposed method is retaining the bony structure details present in the CT image and soft tissue details present in the MR image with good contrast.

There is further scope to improve the proposed method by pre-processing the CT image and then fusing the resulting image with the MR image. The proposed method can be extended to the fusion of anatomical and functional medical images which are usually represented as gray scale and colour images, respectively.

References

1. Amdur RJ, Gladstone D, Leopold KA, Harris RD: Prostate seed implant quality assessment using MR and CT image fusion. Int J Radiat Oncol Biol Phys 43:67–72, 1999
2. Joel Leong F, Siegel AH: Clinical applications of fusion imaging. Appl Radiol(32), 2003
3. Constantinos SP, Pattichis MS, Mitheli-Tzanakou E: Medical imaging fusion applications: an overview. Proc 35th Asilomar Conf Signals Syst Comput 2:1263–1267, 2001
4. Nemec SF, Donat MA, Mehrain S, Friedrich K, Krestan C, Matula C, Imhof H, Czerny C: CT–MR image data fusion for computer assisted navigated neurosurgery of temporal bone tumors. Eur J Radiol 62: 192–198, 2007
5. Nemec SF, Peloschek P, Schmook MT, Krestan CR, Hauff W, Matula C, Czerny C: CT-MR image data fusion for computer-assisted navigated surgery of orbital tumors. Eur J Radiol 73: 224–229, 2010
6. Toet A: Image fusion by a ratio of low-pass pyramid. Pattern Recognit Lett 9(A Toet):245–253, 1989
7. Baum KG, Schmidt E, Rafferty K, Krol A, María H: Evaluation of novel genetic algorithm generated schemes for positron emission tomography (PET)/magnetic resonance imaging (MRI) image fusion. J Digit Imaging 24:1031–1043, 2011
8. Toet A, van Ruyven LJ, Valeton JM: Merging thermal and visual images by a contrast pyramid. Opt Eng 28:789–792, 1989
9. Li H, Manjunath BS, Mitra SK: Multisensor image fusion using the wavelet transform. Graph Models 57:235–245, 1995
10. Selesnick IW, Baraniuk RG, Kingsbury NG: The dual-tree complex wavelet transform. IEEE Signal Process Mag 22:123–151, 2005
11. Zhenfeng S, Jun L, Qimin C: Fusion of infrared and visible images based on focus measure operators in the curvelet domain. Appl Opt 51:1910–1921, 2012
12. Candès E, Demanet L, Donoho D, Ying LX: Fast discrete curvelet transforms. Multiscale Model Simul 5:861–899, 2006
13. Ali FE, El-Dokany IM, Saad AA, Abd El-Samie FE: Curvelet fusion of mr and ct images. Prog Electromagn Res C 3:215–224, 2008
14. Xu J, Yang L, Wu D: Ripplet: a new transform for image processing. J Vis Commun Image Represent 21:627–639, 2010
15. Das S, Chowdhury M, Kundu MK: Medical image fusion based on ripplet transform type-I. Prog Electromagn Res B Pier B 30:355–370, 2011
16. Qu X, Yan J, Xie G, Zhu Z, Chen B: A novel image fusion algorithm based on bandelet transform. Chin Opt Lett 5:569–572, 2007
17. Miao Q, Shi C, Xu P, Yang M, Shi Y: A novel algorithm of image fusion using shearlets. Opt Commun 284:1540–1547, 2004
18. Do MN, Vetterli M: The contourlet transform: an efficient directional multiresolution image representation. IEEE Trans Image Process 12: 2091–2106, 2005
19. Yang L, Guo BL, Ni W: Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. Neurocomputing 72:203–211, 2008
20. Cunha AL, Zhou JP, Do MN: The non-subsampled contourlet transform: theory, design and applications. IEEE Trans Image Process 15: 3089–3101, 2006

21. Li S, Yang B, Hu J: Performance comparison of different multiresolution transforms for image fusion. Inf Fusion 12:74–84, 2011
22. Lu H, Zhang L, Serikawa S: Maximum local energy: an effective approach for multisensor image fusion in beyond wavelet transform domain. Comput Math Appl 64:996–1003, 2012
23. Pajares G, de la Cruz JM: A wavelet-based image fusion tutorial. Pattern Recognit 37:1855–1872, 2004
24. Piella G: A general framework for multiresolution image fusion: from pixels to regions. Inf Fusion 4:259–280, 2003
25. Yang Y, Park DS, Huang S, Rao N: Medical image fusion via an effectivewavelet-based approach. EURASIP J Adv Signal Process 2010:44, 2010
26. Lewis JJ, O'Callaghan RJ, Nikolov SG, Bull DR, Canagarajah N: Pixel and region-based image fusion with complex wavelets. Inf Fusion 8:119–130, 2007
27. Li H, Chai Y, Li Z: Multi-focus image fusion based on nonsubsampled contourlet transform and focused regions detection. Optik 124:40–51, 2013
28. Chai Y, Li H, Zhang X: Multifocus image fusion based on features contrast of multiscale products in nonsubsampled contourlet transform domain. Optik 123:569–581, 2012
29. Zheng Y, Essock EA, Hansen BC, Haun AM: A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. Inf Fusion 8:177–192, 2007
30. Qu G, Zhang D, Yan P: Information measure for performance of image fusion. Electron Lett 38:313–315, 2002
31. Petrović V, Xydeas C: Objective image fusion performance measure. Electron Lett 36:308–309, 2000
32. Piella G, Heijmans H: A new quality metric for image fusion. ICIP 3: 173–176, 2003
33. Wang Z, Bovik AC: A universal image quality index. IEEE Signal Process Lett 9:81–84, 2002