Block-Matching Based Multimodal Medical Image Fusion via PCNN with SML

Hu Shaohai[1,2], Yang Dongsheng[1,2], Liu Shuaiqi[3,4], Ma Xiaole[1,2]
1. Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
2. Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China
3. College of Electronic and Information Engineering, Hebei University, Baoding 071002, China
4. Key Laboratory of Digital Medical Engineering of Hebei Province, Baoding 071002, China
Email: dsyang@bjtu.edu.cn, shhu@bjtu.edu.cn, shdkj-1918@163.com, 15120342@bjtu.edu.cn

***Abstract*—Multimodal medical image fusion has become a powerful tool for medical diagnoses. To further enhance its effect, we adopt a new approach similar with the block-matching and 3D filtering algorithm (BM3D) for image de-noising. The algorithm first block the input images and group the blocks into 3D arrays by the block-matching step before multi-scale transform, and then use a 3D transform which consist of a 2D multi-scale transform (e.g. non-subsampled shearlet transform, NSST) and a 1D transform (e.g. DCT)to obtain the coefficients. An averaging fusion rule is used on the obtained low frequency component; for high-frequency components, the algorithm first calculate the Sum of Modified Laplacian (SML), and then use them to calculate ignition maps by Pulse Coupled Neural Network (PCNN) and decide the fused image's high-frequency coefficients by the maps. The final fused image obtained after the inverse transform by using an aggregation process. Experimental result shows the proposed algorithm has better results than some state-of-the-art methods on both subjective and objective evaluation.**

Keywords—Image fusion; Block matching; BM3D; NSST; PCNN; SML

I. INTRODUCTION

Image fusion refers to the process that obtaining a new single synthesized image from two or more images, which is obtained from the same or different sensors in the same scene [1]. As a powerful tool for modern medical diagnoses, multimodal medical image fusion can merge the medical imaging information from different imaging equipment (e.g. Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Positron Emission Computed Tomography (PET)) into a fused image. It can help clinicians make more accurate analysis, so as to formulate more appropriate treatments.

Recently, multi-scale image fusion has become the mainstream of image fusion technology. Following the extensive application of Wavelet [2], Contourlet [3], Shearlet [4] and a number of new methods which are better than wavelet is emerging. Compared with others, the Shearlet transform has a relatively simple mathematical structure and a better performance in direction representation. The non-subsampled Shearlet transform (NSST) [5] used in this paper makes the Shearlet has translation invariance and stronger direction selectivity. Besides, the effectiveness and efficiency of fusion is ensured, since the removing of the sampling process.

Typical common transform domain methods usually use relatively fixed fusion framework, which perform multi-scale transform first, and then use fusion rules process transform coefficients, finally use inverse transform to obtain fused image. In such structure, process of the input images only happens after transform, which easily lead to the loss of specific characteristics in spatial domain(e.g. edge contours, etc.), and since the absence of these, it usually cause distortion, textures and blockiness in fused image.

Compared to existing frameworks, the proposed one use a new structure similar with the BM3D denoising algorithm [6], which is a combine structure of block-matching and 3D transform. The purpose is to achieve better results by using spatial domain information to assist transform domain fusion. We first introduce the IHS conversion [7] step for color medical images, and for the obtained components: Intensity (I), Hue (H) and Saturation (S), only the intensity join the subsequent fusion operations. Then we add the blocking-matching process in the spatial domain, and adopt a 3D transform which is the combination of a 2D and a 1D transform. For obtained transform coefficients, we use the averaging fusion rule in low-frequency. As for high-frequency rule, we first calculate the SML [8] of corresponding image blocks, then use them as the link intensity values of PCNN [9] to obtain the ignition maps, and then decide the fused coefficients by the maps. The final fused image is obtained from a series of fused 3D image block groups after the inverse transform by using an aggregation process.

II. FUSION FRAMEWORK

The proposed fusion framework can be described as Fig.1. If the input multimodal medical image is a gray-level image like MRI, we directly process the block-matching step. Otherwise, a conversion from RGB to IHS color space is appended for pseudo-colored images, then the obtained intensity component is used for the subsequent process. Then blocking the image by a fixed sliding window; setting a fixed searching area base on the current reference block and

matching them by similarities; grouping the chosen blocks of each reference block into a 3D array; transforming the 3D arrays by 3D transform.

After the transform, the coefficients can be divided into high- and low-frequency components. We use an averaging fusion rule in low-frequency. For high-frequency components, we calculate the SML value of the high-frequency portion, and then calculate the PCNN link intensity to obtain the ignition maps, the fused high-frequency coefficients are decided by choosing the max value of the maps. A series of image blocks formed as 3D arrays can be obtained by inverse transform; the gray-level fused image is obtained by utilizing an aggregation approach which recalculate the pixel values of overlapping pixels; through an IHS inverse conversion, the results of the color medical image fusion is obtained.

As mentioned in the introduction, the idea of BM3D algorithm in image denoising is referenced in the proposed fusion framework. The detailed procedures of these structures have been boxed by blue dashed lines in Fig.1. The added procedure before and after the transform domain process are both the used part of BM3D structure. The improvement of such framework is that it not only focus on modifying transforms, but also introducing more spatial features to reduce artificial products.

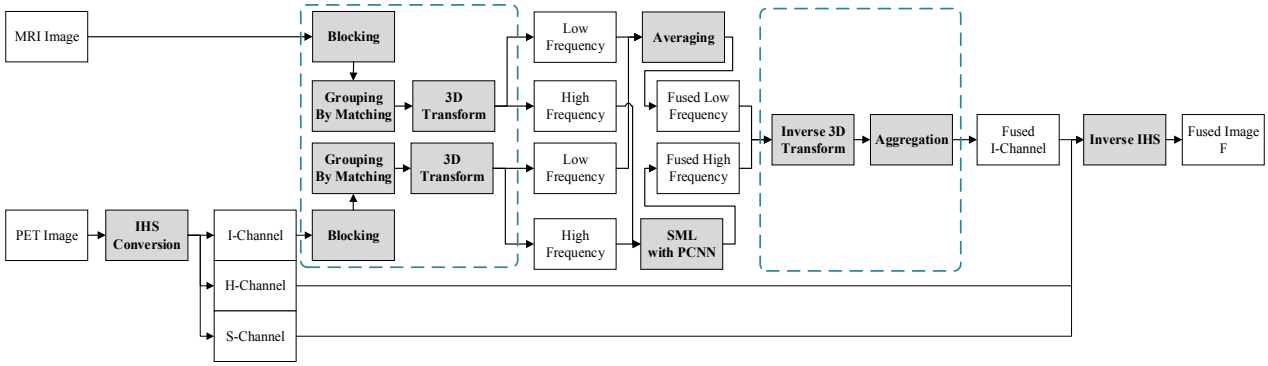


Fig. 1. Block diagram of the proposed image fusion framework.

III. Spatial Domain Processes

The spatial domain process can be divided into IHS conversion, image blocking and grouping by similarity. Here, the IHS conversion mainly targets at the pseudo-colored medical images to extract the intensity components. For image blocking, we use a fixed-size sliding window to segment images. For image grouping, it means stacking similar 2D image blocks together and forming them into 3D arrays. And the measurement of similarity is some kind of distance measure.

A. IHS conversion

IHS color space is a commonly used color coordinate system in multi-spectral image fusion. For color medical image fusion, we need convert an image in RGB color space (i.e. Red, Green and Blue) into IHS model (i.e. Intensity, Hue, Saturation). In order to reduce the spectral information distortion caused by the conversion as small as possible, we use a triangle model in [7] to calculate the IHS conversion. The conversion can be define as:

$I = (R+G+B)/3$;

when B is the minimum, $3H = (G-B)/(I-B), S = 1-B/I$;

when R is the minimum, $3H = (B-R)/(I-R), S = 1-R/I$;

when G is the minimum, $3H = (R-G)/(I-G), S = 1-G/I$;

The inverse conversion can be define as:

when B is the minimum,

$B = I(1-S), G = 3H(I-B)+B, R = 3I-B-G$;

when R is the minimum,

$R = I(1-S), B = 3H(I-R)+R, G = 3I-B-R$;

when G is the minimum,

$G = I(1-S), R = 3H(I-G)+G, B = 3I-R-G$;

B. Block-matching

For 2D image blocking, the sliding window is based on a certain window size and step size. Then we can filter out some blocks within the search area by a pre-selected rule. For each image block after blocking, the algorithm process is illustrated in Fig.2. The white boxed area which is enlarged on the right is the search area that uses the reference block (marked R) as center, and the similar blocks (marked S) are pointed out by the black dash arrow. The block -matching process is as follows:

1) Select the current block as reference block;

2) Draw a fixed search area centering on the reference block;

3) Calculate the distance between the reference block and all the image blocks in the search area (i.e. candidate blocks);

4) List the distance of all the blocks in the region, the least distance means the most similar;

5) Compare the distances with a pre-set threshold, all these blocks lesser than the threshold are defined as similar;

6) Arrange these similar blocks into a 3D array.

To complete entire block-matching process, all blocks need to be traversed as reference blocks.

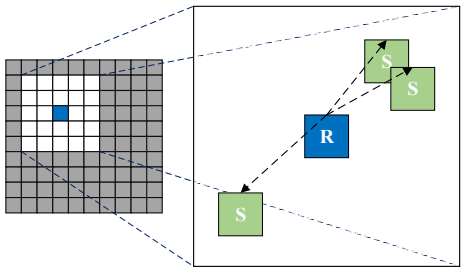


Fig. 2. Schematic diagram of Block-Matching.

C. Similarity measurement

To better reflect the similarity between 2D blocks in the 3D array, a classic way is to calculate a certain distance measure. Here, we use a group-by-similarity method to achieve the classification. This method can be implemented to the intersection-contained classification of mutually similar signal fragments.

Since the process for two input images are exactly the same, we only use the MRI image A as an example. We denote x as a set of 2D spatial coordinate and its value belongs to 2D image domain $X \subset Z^2$. Thus, for any $N \times N$ image block, it can be denoted as A_x, where the subscript x represents the coordinate of the top-left corner of the image block. So A_x is the image block of A which is segmented at location x. Besides, A_S represents a 3D array which consist of many A_x, $x \in \mathbf{S} \subseteq X$. Besides, the distance used herein is ℓ^2-distance, which is obtained through ℓ^2-norm [6].

Since it is unable to obtain the ideal image, the calculation of distance can only use the input images. In order to better solve the results difference of using input image as the estimation of ideal image, we employ a relatively coarse 2D linear pre-filter [6] to preprocess the two original image blocks. This approach relatively reduces the false positives, and the final distance can be calculate as:

$$d(A_{x_R}, A_{x_C}) = \frac{\left\| f_{2D}(A_{x_R}) - f_{2D}(A_{x_C}) \right\|_2^2}{N^2} \qquad (1)$$

where $\|\bullet\|_2$ is the ℓ^2-norm, A_{x_R} and A_{x_C} denote the reference and candidate block respectively, $f_{2D}(\cdots)$ represents the 2D linear filter. The set of all the coordinates x of the similar blocks of the reference block A_{x_R} can be expressed as:

$$\boldsymbol{S}_{x_R} = \{x \in X : d(A_{x_R}, A_{x_C}) \le \tau_{\max}\} \qquad (2)$$

where $\tau_{\max}$ is a fixed threshold that represents the maximum distance of two blocks which are considered as similar, and it is determined by the acceptable value of the ideal difference of the natural images. If the coordinates set of all similar blocks is $\boldsymbol{S}_{x_R}$, then we can use these similar blocks $A_x \in \boldsymbol{A}_S, x \in \boldsymbol{S}_{x_R}$ to form a 3D array of size $N \times N \times N_S$, denoted as $\boldsymbol{A}_{S_{x_R}}$, where N_S is the number of similar blocks.

IV. 3D Transform and NSST

A. 3D Transform

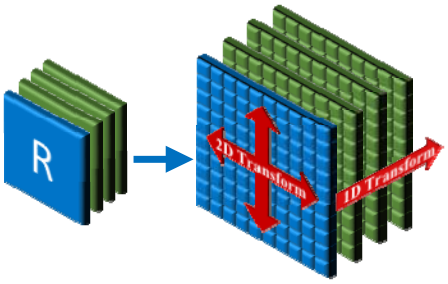


Fig. 3. Schematic diagram of 3D transform.

After block-matching, the coefficients for fusion rules can be obtained by 3D transform. The 3D transform is a combination of a 2D and 1D transforms. In this paper, for each 2D image block in the array, we first adopt the NSST, then use a 1D DCT on each column of this array (i.e. on the third dimension). The effectiveness of such design mainly because the full use of the spatial correlation after transform, and also due to the creation of sparsity on the coefficients by the threshold shrinkage. These processes reduce the uncertainty of fused image, and optimize the fusion result. The 2D and 1D transforms can be denoted as $\boldsymbol{T}_{2D}(\cdots)$ and $\boldsymbol{T}_{1D}(\cdots)$ respectively. The transform can be lively presented by Fig.3, where the red-cross arrow marks the unfolding surfaces of the 2D transform and the one-way arrow indicates the direction of the 1D transform. The 3D transform coefficients of $\boldsymbol{A}_{S_{x_R}}$ can be expressed as:

$$\mathcal{A}_{S_{x_R}}^{3D} = \boldsymbol{T}_{3D}(\boldsymbol{A}_{S_{x_R}}) = \boldsymbol{T}_{1D}(\mathcal{A}_{S_{x_R}}^{2D}) \qquad (3)$$

where $\boldsymbol{T}_{3D}(\cdots)$ represents the entire 3D transform process, $\mathcal{A}_{S_{x_R}}^{2D}$ denotes the set of coefficients after the 2D transform. The decorated letter $\mathcal{A}_x$ means the 2D transform coefficients of the intra-group block A_x. Besides, a set hard threshold filter $f_{ht}(\cdots)$ is used to create sparsity. The filter can be defined as:

$$f_{ht}(\tau, \tau_{ht}) = \begin{cases} \tau, & \text{if } |\tau| > \tau_{ht} \\ 0, & \text{otherwise.} \end{cases} \qquad (4)$$

where τ represents current input coefficient, τ_{ht} is the fixed parameter (i.e. hard-threshold). Then transform coefficients which are provided for fusion rules can be expressed as $\tilde{\mathcal{A}}^{3D}_{S_{xR}} = f_{ht}(\mathcal{A}^{3D}_{S_{xR}}, \tau_{ht})$.

B. Non-subsampled Shearlet transform

For the used 2D transform, we adopt the most effective and efficient transform NSST. NSST is constructed through affine systems with composite dilations, when the dimension n = 2, the affine systems can be defined as follows.

$$\boldsymbol{M}_{DS}(\psi) = \{\psi_{j,l,k}(x) = |\det D|^{\frac{j}{2}} \psi(S^i D^j x - k) : j,l \in Z, k \in Z^2\} \quad (5)$$

where $\psi \in L^2(R^2)$, D, S are both 2×2 invertible matrices and $|\det S| = 1$. D is the dilations matrix, S stands for the shear matrix. If $\sum_{i,j,k} \left|\langle f, \psi_{j,l,k} \rangle\right|^2 = \|f\|^2$ is met with any $f \in L^2(R^2)$, it means $\boldsymbol{M}_{DS}(\psi)$ forms the tight frame, which implies it is compactly supported, and the element of $\boldsymbol{M}_{DS}(\psi)$ are called composite wavelets. When the value of D, S is formed as follows:

$$D = \begin{bmatrix} a & 0 \\ 0 & a^{1/2} \end{bmatrix}, S = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \quad (6)$$

We say the element of $\boldsymbol{M}_{DS}(\psi)$ is shearlets, which means shearlet is a special case of the composite wavelets. Besides, usually we use $a = 4, s = 1$.

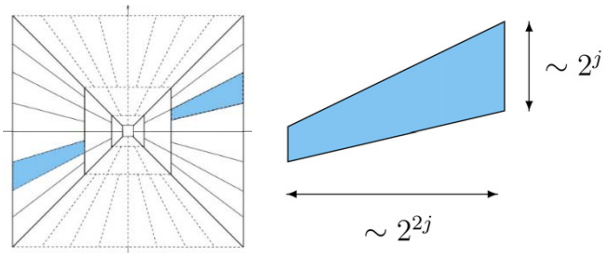


Fig. 4. the support zones of a shearing filter.

The discretization of NSST consists of two phases including multi-scale and multi-direction decomposition. For multi-scale decomposition, NSST adopts the non-subsampled pyramid (NSP). For multi-direction decomposition, it is realized through a modified shearing filter in NSST. The non-subsampled shearing filter map the pseudo-polar grid back to Cartesian grid system, so the entire process can be directly completed through the 2D convolution [5]. The support zones of NSST is a pair of trapeziform zones with the size of $2^{2j} \times 2^j$, which is shown on Fig.4.

V. FUSION RULES AND BLOCK AGGREGATION

In the 3D transform, the purpose of adding the 1D DCT is to utilize correlation and to reduce significant coefficients. It does not destroy the positional distribution of the obtained high- and low- frequency components but only increases the sparsity within the group. So the fusion actually takes place on each 2D surfaces of the array. Hence, the subsequent described words like the frequency components and fusion rules, they still aim at 2D surfaces. Since the strong complementarity of medical images, the corresponding similar block's amount in different images are not necessarily equal. As the image blocks contain detail features (e.g. uniform areas, gradients, etc.) generally more dissimilar with others, they usually have less similar blocks compared to less-feature ones. When this happen, the lesser number is used as the final number of the fused block.

A. High-frequency fusion rule

High-frequency coefficients usually contain detail information and salient features. E.g. MRI has high sensitivity on soft tissue and blood vessels, and they represents detail edges and textures information on imaging. To better retain them, using SML rule for high-frequency components is advantageous. This is because the human visual system is not sensitive to single pixels, but more sensitive to edges and texture information, and SML can properly represent these. The SML can be defined as follow:

$$\begin{aligned} \nabla^2_{ML}(x,y) = &| 2f(x,y) - f(x-step,y) \\ &- f(x+step,y) | + | 2f(x,y) - \\ &2f(x,y-step) - f(x,y+step) | \end{aligned} \quad (7)$$

where $step$ represents the variable spacing between pixels, usually $step = 1$.

$$SML(i,j) = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} \nabla^2_{ML} f(i,j) \quad (8)$$

where N is the window size, $f(i,j)$ denotes the coefficient at the position (i,j) .

PCNN is a kind of simplified model with synchronized pulsed, it composed of plenty of neurons, each neuron consists of three parts: receptive field, modulation field and pulse generator. When PCNN used in digital image processing, we set each pixel in the image corresponds to a neuron. The following formula is the simplified model:

$$\begin{cases} F_{i,j}(n) = I_{i,j}(n) \\ L_{i,j}(n) = e^{-\alpha_L} L_{i,j}(n-1) + V_L \sum_{p,q} W_{ij,pq} Y_{pq} \\ U_{i,j}(n) = F_{i,j}(n) \cdot (1 + \beta \cdot L_{i,j}(n)) \\ \Theta_{i,j}(n) = e^{-\alpha_\Theta} \Theta_{i,j}(n-1) + V_\Theta Y_{i,j}(n) \\ Y_{i,j}(n) = \begin{cases} 1 & U_{i,j}(n) > \Theta_{i,j}(n) \\ 0 & otherwise \end{cases} \end{cases} \quad (9)$$

where n is the number of iterations, $F_{i,j}, L_{i,j}$ are the feeding and linking input of the neuron (i,j); $I_{i,j}$ is an external stimulus; $\beta$is the linking parameter between neurons; $U_{i,j}$ is the internal activity; $\Theta_{i,j}$ is the dynamic threshold; $Y_{i,j}$ is the output of the PCNN; W is the synaptic weight coefficients; V_L is a normalizing constant; V_Θ is the normalized constant; α_L and α_Θ are the time constants of linking input and dynamic threshold; If $U_{i,j} > \Theta_{i,j}$, the neuron generates a pulse, we call it an ignition. After ignition, the output of PCNN is the ignition maps consist of the total number of ignition. Thus, the fusion based on SML and PCNN can be described as follow:

$$\mathcal{F}_x(i,j) = \begin{cases} \mathcal{A}_x(i,j) & O_A \geq O_B \\ \mathcal{B}_x(i,j) & O_A < O_B \end{cases} \tag{10}$$

where $\mathcal{A}_x, \mathcal{B}_x$ are the SML value of the corresponding high-frequency coefficients, $\mathcal{F}_x$ is the fused coefficient. O_A, O_B denote the ignition maps obtained by the two channels.

B. Low-frequency fusion rule

The low-frequency coefficient reflects the background information of the source image. Therefore, it may not be able to obtain the salient feature, by using a high-pass requirement in the approximation coefficients. Hence, the most common fusion rule in low-frequency is the averaging operation:

$$\mathcal{F}_x(i,j) = \frac{\mathcal{A}_x(i,j) + \mathcal{B}_x(i,j)}{2} \tag{11}$$

where $\mathcal{A}_x$ and $\mathcal{B}_x$ denote the approximation coefficients of the input image block respectively.

C. Aggregation of blocks

A series of after-fused 3D arrays $\boldsymbol{F}_{S_{x_R}}$ can be obtained from the fused coefficients $\mathcal{F}^{3D}_{S_{x_R}}$ by the inverse 3D transform. When we restore the image blocks into 2D surfaces, because of the similar blocks selection, there will be overlap pixels between blocks. To solve the overlapping problem, an averaging operation is needed to calculate the mean value of overlap pixels located on every single position, we call it the aggregation of blocks. For the coefficient of each image block's pixel values ω_{x_R}, it can be defined as:

$$\omega_{x_R} = \begin{cases} \dfrac{1}{n_{x_R}}, & \text{if } n_{x_R} \geq 1. \\ 1, & \text{otherwise} \end{cases} \tag{12}$$

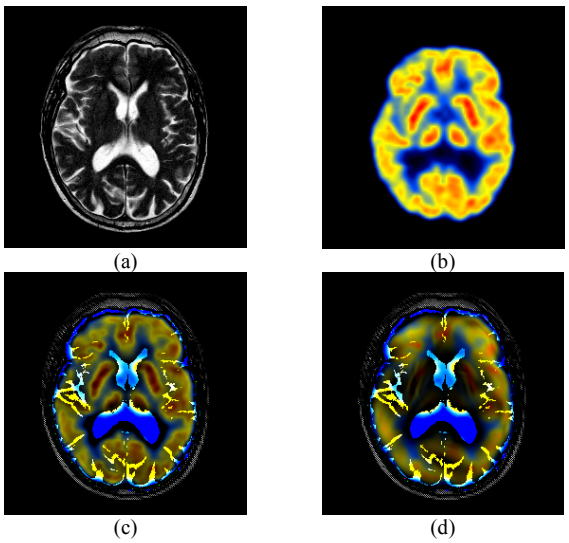
where n_{x_R} is the number of retained non-zero coefficients in $\boldsymbol{F}_{S_{x_R}}$. The final fused image $F = \sum_{x_M \in X} \sum_{x_m \in S_{x_M}} \omega_{x_M} F_{x_m}, \forall x \in X.$ where x_M is the coordinate of an unspecific reference block, x_m denotes the position of a similar block which is included in the x_M locate group.

VI. Experiments

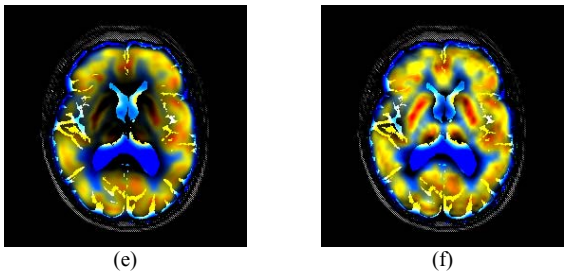
In order to provide an effective evaluation of the effectiveness and reliability of the proposed algorithm, we used a set of human brain's MRI and PET images as an example, and carried out the comparative experiments among the existing DWT, NSCT, NSST and the proposed algorithms. We comparatively analyze the relative advantages of the proposed algorithm through the subjective and objective results.

In the experiment, the size of the blocks are pixels, and the sliding window's step is 8 pixels. For DWT, we use a 3 levels db2 wavelet. The decomposition level of NSCT is 4 levels, and there are respectively 2, 8, 16, 16 directional sub bands in each level. For NSST, we use 3 levels decomposition, the number of directional sub bands are 10, 10 and 18. For the NSST in proposed algorithm, due to the blocking process, the decomposition level is 2. In addition, the experiment is implemented on an Intel Core i5 2.27 GHz with 4GB RAM. The simulation software is Matlab 2014a.

As for the evaluation index system, the experiment use Average Gradient [11], Normalized Mutual Information (MI) [12], Edge Based Similarity Measure ($Q^{AB/F}$) [13] and Structural Similarity (SSIM) [14] as the objective evaluation criteria. The larger the value of these four indexes are, the clearer the image is, and the higher the fusion quality is.



(a) (b) (c) (d)



(e) (f)

Fig. 5. the results of the different algorithms.

Fix.5. shows the fusion results of MRI and PET images of a human brain affected with Alzheimer. (a) is the MRI image, by which, normal and pathological soft tissue can be better visualized; (b) is the PET image, which can provide better information on blood flow and flood activity, and show the functional eloquent brain areas such as motor or speech; (c) is the result of DWT-MAX; (d) is the result based on NSCT-MAX; (e) is the result of NSST-MAX; (f) is the result of the proposed algorithm. Compare with others, the proposed algorithm has a better performance on edge details. Because of the clearer edges and more abundant textures, some salient features on its result remain intact. The location of the diseased tissue in the fused image still has a clear structure, texture and color characteristics, which can further facilitate the diagnosis and localization of the disease.

The objective criteria can be seen on Table 2. The proposed algorithm has relatively good performance on all of these four evaluation indexes, especially on MI and $Q^{AB/F}$. This shows that the proposed's result is more similar with both input images on the edges structure.

TABLE I. THE INDEXES OF DIFFERENT ALGORITHMS

Fusion method	Objective Criteria			
	AVG	***MI***	$Q^{AB/F}$	***SSIM***
DWT	6.9843	1.5424	0.2420	0.56586
NSCT	6.5165	1.3968	0.1934	**0.6058**
NSST	6.5281	1.4216	0.2287	0.58572
The Proposed	**7.0389**	**1.5738**	**0.2733**	0.59207

VII. CONCLUSION

In this paper, we present a new multimodal medical image fusion framework based on block-matching, PCNN and SML. Compared with existing methods, by using the blocking-matching in spatial domain and the optimal selection of high-frequency coefficients through PCNN and SML, the proposed method enhance the fusion result by utilizing more spatial correlation. Experimental results demonstrate that the proposed algorithm outperforms the existing ones in terms of qualitative and quantitative evaluations. Since the computational complexity of classification, how to improve efficiency will be one of the main research directions.

ACKNOWLEDGMENTS

Thanks to the Natural Science Foundation of China: No.61401308 and No.61572063, and the Fundamental Research Funds for the Central Universities No. 2016YJS039.

REFERENCES

[1] M. B. A. Haghighat, A. Aghagolzadeh, H. Seyedarabi. Multi-focus image fusion for visual sensor networks in DCT domain [J]. Computers & Electrical Engineering. 2011, 37(5): 789-797.

[2] Guihong Qu, Dali Zhang, and Pingfan Yan. Medical image fusion by wavelet transform modulus maxima [J]. Optics Express. 2001, 9(4): 184-190

[3] M. N. Do, M. Vetterli. The Contourlet transform: an efficient directional multi-resolution image representation. [J]. IEEE Transactions on Image Processing. 2006 14(12): 2091-2106.

[4] Guo K, Labate D. Optimally sparse multidimensional representation using Shearlets [J]. SIAM Journal on Mathematical Analysis. 2007 39(1): 298-318.

[5] Gao Guorong, Xu Luping, Feng Dongzhu. Multi-focus image fusion based on non-subsampled shearlet transform [J]. IET Image Processing, 2013, 7(6): 633 - 639

[6] K.Dabov, A.Foi, V.Katkovnik, and K. Egiazarian. Image denosing by sparse 3D transform-domain collaborative filtering [J]. IEEE Transactions on Image Processing, 2007 16(8): 2080 - 2095.

[7] Smith A R. Color Gamut Transform Pairs [J]. Computer Graphics, 1978, 12(3):12-19.

[8] Huang Wei, Jing Zhongliang. Evaluation of focus measures in multi-focus image fusion [J]. Pattern Recognition Letters. 2007, 28(4):493-500.

[9] Johnson J L, Padgett M L. PCNN models and applications [J].IEEE Transactions on Neural Networks, 1999, 10(3): 480-498.

[10] Liu S, Zhu Z, Li H, et al. Multi-focus image fusion using self-similarity and depth information in nonsubsampled shearlet transform domain [J]. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2016 9 (1): 347-360

[11] Qu G H,Zhang D L,Yan P F.Information measure for performance of image fusion [J].Electronic Letters, 2002,38(7) : 313-315.

[12] Xydeas C S,Petrovi V.Objective image fusion performance measure [J].Electronics Letters, 2000,36 (4):308-309.

[13] Wang Z,Bovik A C Sheik H R,et al. Image Quality Assessment: From error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.