A NEW FUSION FRAMEWORK FOR MULTIMODAL MEDICAL IMAGE BASED ON GRWT

Dongli Wei[1,2], Guixia Kang[1,2], Chuan Hu[1,2], Beibei Hou[1,2], Ningbo Zhang[1,2], Yuan Tang[1,2]

[1]Key Laboratory of Universal Wireless Communications, Ministry of Education Beijing University of Posts and Telecommunications, Beijing, China

[2]Wuxi BUPT Sensory Technology and Industry Institute CO.LTD, Wuxi, China

ABSTRACT

Hypertension is one of the most important contributors to heart disease and stroke. Multimodality medical image fusion plays an important role in the precise diagnosis, treatment planning and follow-up studies of various diseases. In this paper, we propose an image fusion framework in patients with hypertension, which is based on Frei-Chen operators in generalized reisz-wavelet transform (GRWT) domain. The proposed method is tested on two cases of MRI/CT and MRI/SPECT images. The input medical images are first transformed by GRWT into basic and detail components. Further, they are fused respectively by Frei-Chen operators which extracts ripples, edges, lines and points well. Finally, the fused image is constructed by the inverse GRWT and evaluated by indicators such as average gradient and spatial frequency etc. The visual and quantitative evaluation of the results demonstrated the superior performance of the proposed image fusion method which assists doctors for precise and sufficient diagnosis.

Index Terms— GRWT, multimodal, image fusion, Frei-Chen operators.

1. INTRODUCTION

From the World Health Organization (WHO) statistics [1], hypertension is one of the most important contributors to heart disease and stroke, and they together make up the number one cause of premature death and disability in the world. Researchers estimate that hypertension leads to nearly 9.4 million deaths from cardiovascular disease each year [2]. Additionally, Hypertensive encephalopathy (HE) increases the risk of conditions such as kidney failure and blindness. Therefore, the comprehensive and accurate diagnosis is of great significance for the prevention of cardiovascular and cerebrovascular diseases and the clinical treatment of severe hypertension. To achieve it, multimodality medical images, providing different tissue and structural information, are required such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Single Photon Emission Computed Tomography (SPECT) image [3]. CT image provides bone information, MRI scan offers soft tissue information of human brain, and both are high resolution of structural images, while SPECT image is the functional image supplying metabolic information with low resolution. Compared with Positron Emission Computed Tomography, SPECT imaging can also reflect the metabolic distribution, and the tracer has a wide range of adaptability, and high specificity. Nevertheless, multiple independent images sometimes do not satisfy the clinical needs with highly inconvenient and inefficient and may lead to wrong diagnosis. In this case, some experts proposed integrating multimodal images information into a composite image by fusion, which compensates for the limitations of images. Thereby, a more comprehensive image description of the lesion is obtained which helps in the precise diagnosis[4].

In recent years, various types of image fusion has been carried out. In 1997 Hall and Llians [5] proposed a general introduction to multisensory image fusion. The image fusion techniques depend upon the trade-off between the required spectral consistency and preservation of spatial features. In the initial stage, the properties of wavelet like multi resolution and energy compaction was exploited to fuse medical images. Wavelet decomposition captures limited directional information along horizontal, vertical and diagonal directions and performs poorly in edge and texture [6]. Then Tian used DT-CWT to intensify the image fusion level achieving better edge and shift-invariance properties[7]. The non sub sampling of the Curvelet and the Shearlet transform has also been exploited to preserve edge and contour information[8]. Besides, an algorithm based on artificial neural networks was proposed to do pixel level fusion of multi focus images [9]. Although these methods have achieved good fusion results, they lead to the problem of contrast reduction and high memory requirement. Compressed sensing based fusion methods reduce blocking effect, but suffer invariably from reconstruction error. Despite this, fusion pattern considering structure, spatial-direction or other information in transform domain, has not been explored fully. Jin & Jang [10] put forward a novel fusion method based on GRWT and represented the structural information adaptivity and consistency of GRWT provide accurate representation of low level feature based fusion pattern to fuse various kind of imageries.

In this paper, an image fusion method is proposed based on GRWT [11] and Frei-Chen operators. It aims to develop a model to integrate structural information including ripple, edges, lines and points adaptively and consistently, and achieve a better performance.

2. METHODS

Fig. 1 illustrates an overview of the proposed image fusion framework including decomposition, fusion and reconstruction. The detailed work is described as below.

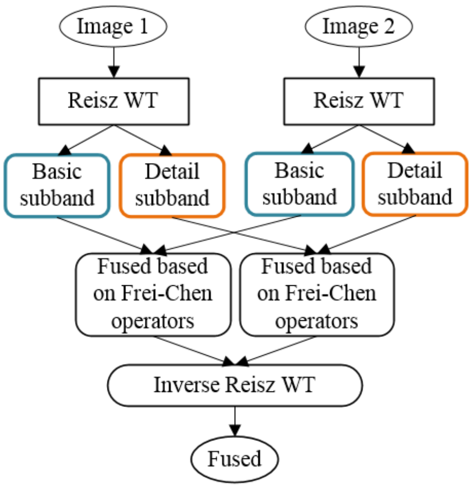

Fig. 1. Methodology of the proposed method.

2.1. Generalized Reisz wavelet transform (GRWT)

The Reisz transformation [11] is considered a natural extension of the Hilbert transform (HT) which is a signal operation from scalar to vector and acts as an all-pass filter. The transfer function is given as

$$\widehat{H}(\omega) = -j\,sign(\omega) = -j\omega / \|\omega\|, \qquad (1)$$

where $H(w)$ is the space-domain version, $\widehat{H}(w)$ denotes the transfer function in frequency-domain and w is the frequency variable. Based on HT, the Reisz WT of a signal $f(w)$ is

$$R[f(\omega)] = -j\frac{\omega}{\|\omega\|} / \widehat{f}(\omega), \qquad (2)$$

where $\widehat{f}(\bullet)$ and $R[\bullet]$ is the Fourier transform (FT) and riesz-transform of the input signal $f(w)$ respectively. The indication of this transformation can be further defined as

$$R[f(x)] = \begin{pmatrix} R_1 f(x) \\ \vdots \\ R_d f(x) \end{pmatrix} = \begin{pmatrix} (h_1 * f)(x) \\ \vdots \\ (h_d * f)(x) \end{pmatrix}, \qquad (3)$$

where the filter h_n is expressed as the frequency response of $\widehat{H}_n(\omega)$, and d is the dimension. The space-domain based equation can be obtained

$$y(x) = R^{-1}\{\|w\|^{-1}\}(x). \qquad (4)$$

2.2. 2D generalized Reisz wavelet transform

Based on 2.1, the 2D-version GRWT [11] can be obtained directly. There are N+1 individual components for N^{th} order Riesz transform with $d = 2$, the subspace $\widehat{V}_{N,2}$ as

$$\widehat{V}_{N,2} = span\{\widehat{H}_{n_1,N-n_1}(\omega)\} \quad (n_1 = 0 \cdots N), \qquad (5)$$

$$\widehat{H}_{n_1,N-n_1}(\omega) = \sqrt{\binom{N}{n_1}}\left(-j\frac{\omega_1}{\|\omega\|}\right)^{n_1}\left(-j\frac{\omega_2}{\|\omega\|}\right)^{N-n_1}. \qquad (6)$$

Meanwhile, through the polar coordinate, we can get $\cos(\theta) = \omega_1 / \sqrt{\omega_1^2 + \omega_2^2}$ and $\sin(\theta) = \omega_2 / \sqrt{\omega_1^2 + \omega_2^2}$, then $z = e^{j\theta}$. The Eq.(6) have a more simple formation expressed as below (2π-periodic radial profile functions [11])

$$H_{p,q}(z) = (-j)^{p+q}\sqrt{\binom{p+q}{q}}\left(\frac{z+z^{-1}}{2}\right)^p\left(\frac{z-z^{-1}}{2j}\right)^q, \qquad (7)$$

where p and q are the orders of the frequency responds of different Reisz components and the sum of them equals to N. The Reisz Wavelet coefficients can be obtained by

$$C_i^{(n_1,\cdots n_d)}[k] = \left\langle f, R^{(n_1,\cdots n_d)}\phi_{i,k}\right\rangle, \qquad (8)$$

where ϕ is the directional analysis based WT, i stands for current decomposition level and k for the location of multi-resolution transformation.

2.3. Frei-Chen operators based Fusion rule

In order to ensure the sensible and efficient information being captured, the criteria followed for sub-band fusion is the key which has an important influence on the quality of the generated fusion image. Thus, it is essential to define the activity metric for the coefficients of the subband. And then the basic and the detail subband are merged respectively.

The Frei-Chen operators have been applied in the feature extraction and edge detection of infrared images which can extract distinctive features such as ripple, edges, lines, points. It is classified into 3 subspaces: edge , line and the average value [12] described in Fig. 2.

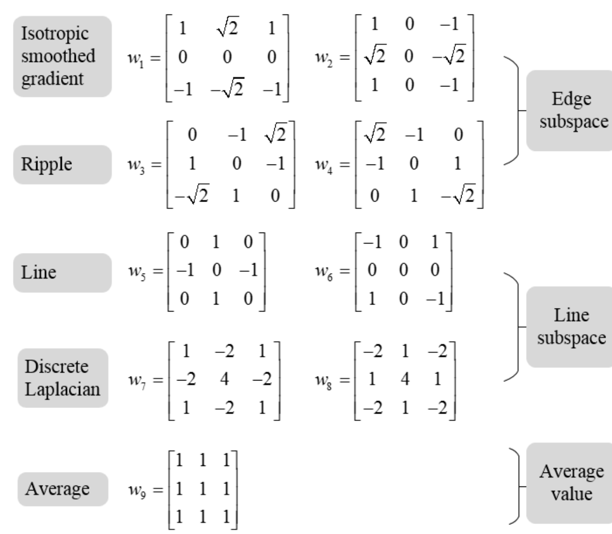

Fig. 2. The set of basis vectors.

The edge subspace is composed of operators w_1-w_4, where w_1 and w_2 are the isotropic average gradient weighting sets, and the ripple value can be obtained by operator w_3 and w_4. Moreover, the line subspace consists of operators w_5-w_8, where w_5 and w_6 make up the line vectors and w_7 and w_8 measure the discrete Laplacian. And the last operator w_9 is added to complete the basis with the average value. Considering that these operators provide all types of basic and detail information, the activity metric of the subband can be defined based on them. The main flow is as follows.

- Square the basic and detail sub-bands separately;
- Convolve the result with the 9 operators;
- Take absolute values for each element in the matrix denoted as F, where BS and DS stand for basic and detail sub-band respectively, m refers to the input images denoted a and b, and the range of n is 1 to 9;

$$F_{mn} = | BS_{m}^{2} * \omega_n \mid or \mid DS_m^2 * \omega_n |. \quad (9)$$

- F is normalized into F', the saliency of sub-band BS_m is denoted as SBS_m, the same as SDS;

$$SBS_m = \sum_{n=1}^{9} F'_{1n}. \quad (10)$$

- Finally, the fusion sub-band is obtained as follow.

$$BS(i,j) = \begin{cases} BS_a(i,j) & SBS_a(i,j) > SBS_b(i,j) \\ BS_b(i,j) & SBS_b(i,j) > SBS_a(i,j) \\ \dfrac{BS_a(i,j)+BS_b(i,j)}{2} & other \end{cases} \quad (11)$$

2.4. Objective quality assessment

A good fusion algorithm should satisfy the requirements that it is reliable and applicable, and that it can extract and retain the saliency information of the input image but does not introduce artifacts or inconsistencies.

In general, the fusion image is evaluated subjectively and objectively. However, the former relies on the observer's vision and professional knowledge. In addition, it is difficult to evaluate one fusion image only through visual analysis in some cases. Hence, it is necessary to establish an evaluation system according to the statistical parameters.

Average Gradient (AG)

The average gradient [13] reflects the clarity of the image, while also reflecting the detail contrast and texture transformation features in the image.

$$AG = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [(\frac{\partial}{\partial x} f(i,j))^2 + (\frac{\partial}{\partial y} f(i,j))^2]^{\frac{1}{2}}, \quad (12)$$

where M and N are the dimensions of the input images. Generally, the larger the AG, the richer the image hierarchy, and the clearer the image is.

Image feature-based metrics: $Q_{AB/F}$

The $Q_{AB/F}$ [14] is used to quantify the similarity between the images. Mathematically, it is computed as follows

$$Q_{AB/F} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [Q_{AF}(i,j) w_A(i,j) + Q_{BF}(i,j) w_B(i,j)]}{\sum_{i=1}^{M} \sum_{j=1}^{N} [w_A(i,j) + w_B(i,j)]}, \quad (13)$$

where A, B are the input images, and F represent and fused image, then $Q_{AF}(i,j)$ and $Q_{BF}(i,j)$ denote the preservation values of edge information at location (i, j), while $w_A(i,j)$ and $w_B(i,j)$ are their weight, respectively. The higher the value, the more edge information is retained from input images.

Peak Signal to Noise Ratio (PSNR)

The PSNR [15] can measure the relationship between the fused image and the input image. A higher PSNR value means better quality of the fused image. It is defined as

$$PSNR = 10 \bullet \lg(\frac{g_{\max}^2}{RMSE^2}), \quad (14)$$

where $g_{\max}$ is the maximum pixel gray level value in the reconstructed image.

Spatial Frequency (SF)

SF [16] depicts the change rate of image intensity which is used to measure the clarity of the image and calculated by row and column frequency

$$SF = \sqrt{RF^2 + CF^2}, \quad (15)$$

$$RF = \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=2}^{N} [I_F(i,j) - I_F(i,j\text{-}1)]^2}, \quad (16)$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{j=1}^{N} \sum_{i=2}^{M} [I_F(i,j) - I_F(i\text{-}1,j)]^2}. \quad (17)$$

A higher spatial frequency value brings in a clearer image.

Entropy (EN)

Entropy [17] can reflect the amount of information in a certain image which is defined as

$$EN = -\sum_{i=0}^{L} P_i \log_2 P_i, \quad (18)$$

where P_i denotes the probability distribution of pixel valuing i, and L is the number of gray levels. The larger the value, the better fusion result is obtained.

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[I_I(i,j) - I_F(i,j) \right]^2}. \quad (19)$$

RMSE [18] is defined as the difference of each pixel value between the input images and the fusion image with the size of $M \times N$. The lower the value, the better the quality.

3. EXPERIMENT RESULTS AND DISCUSSION

Taking the brain images (including 3 modals only: MRI-T2, CT, SPECT-Tc) of hypertensive patients as examples, we study the performance of the proposed image fusion method. All of the images used for practicing are downloaded from the Whole Brain Atlas[19], and grouped into MRI–CT and MRI–SPECT, that is, the inputs of the algorithms. The dimension of MRI and CT is 256*256 and SPECT color image is 256*256*3. More importantly, the corresponding pixels of two input images have been co-aligned perfectly. The original images displayed in Fig.3(a), (b) and (c) are of case one while in Fig.4 (e), (f) and (g) are of case two. Practical programs of fusion are designed by MATLAB R2016b on windows 7 system.

The results of the proposed fusion framework are compared with the DT-CWT [20] and NSCT [21] based methods which are developed by wavelet transform. The fusion image results are displayed in Fig.3 and Fig.4 and the evaluation indicators are shown in Table 1.

For the original images, it can be seen that the source images with different modalities contain complementary information. Both CT images of two cases show no abnormalities, and the prefusion information from the SPECT images can not reflect the specific location of the lesion, whereas double long T2 signals appear on the bilateral occipital lobe from MRI images. Thereby, it may affect the efficiency and accuracy of the diagnosis. However, it is clearly that the long T2 signal appears but no cerebral hemorrhage from MRI-CT fusion image. In addition, there is hypoperfusion in the lesion which may cause infarction from MRI-SPECT fusion image.

Visually, the proposed method preserves the complementary information better and provides spatial consistency of contours, gradients and textures with a higher level of contrast enhancement comparing with the other two methods[20, 21]. Besides, it improves spatial detail without artifacts and distortions. For the evaluation indicators, the best indicator value is highlighted in bold in Table 1, and the proposed method performs better than the other two methods particularly in AG, $Q_{AB/F}$, PSNR and EN.

Multiple individual medical images are inconvenient and inefficient for the doctor, and human visual is limited for the corresponding structure details which may lead to wrong diagnosis. The method of superposing both signals with different degrees of opacity only, does not involve feature fusion and not conducive to the observation and detection of lesions. Through the proposed fusion method with the advantages of high fusion quality and low complexity and easy to be embedded into software, the doctor can obtain a fused image with precise registration and comprehensive information from original images. It helps them determine the cause, locate the lesion and guide the treatment. At present, it can only be used to process 2D images (including PET image) whose pixels have been registered though, that is not a big deal if the original images are from the same hospital. For the 3D fusion, it is feasible to achieve infinite approximation based on 2D image fusion instead of fusion based on voxel level.

4. CONCLUSION

In this paper, an image fusion framework based on Frei-Chen operators in GRWT domain is proposed for multi-modal medical images with hypertension. It is tested on two cases of MRI/CT and MRI/SPECT images comparing with the improved wavelet transform algorithms DT-CWT and NSCT. The visual and quantitative evaluation demonstrate that the proposed method better preserves the image details and improves the image visual effects significantly than the other two algorithms. In conclusion, this study provides an effective medical image fusion algorithm to assist doctors in diagnosing hypertension which may develop into other brain diseases.

5. ACKNOWLEDGMENT

The research is supported by the National Natural Science Foundation of China (No.61471064), and National Science and Technology Major Project of China (No. 2017ZX03001022).

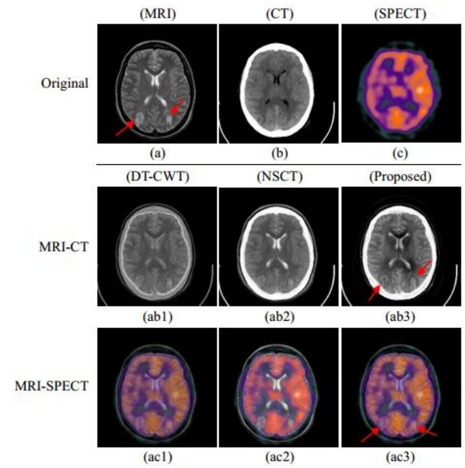

Fig. 3. Performance comparison of DT-CWT[20], NSCT[21] and the proposed method for MRI-CT of case 1.

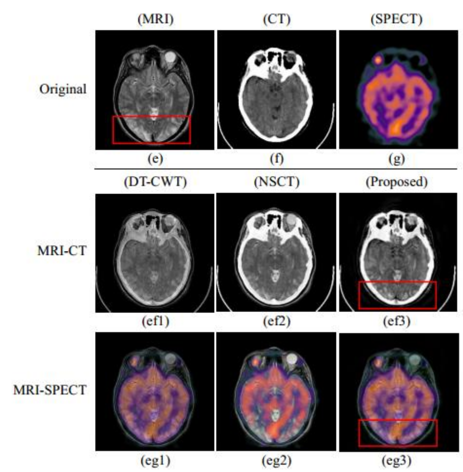

Fig. 4. Performance comparison of DT-CWT[20], NSCT[21] and the proposed method for MRI-SPECT of case 2.

Table 1. The evaluation of indicator results

	Type	Ref.	AG	$Q_{AB/F}$	PSNR	SF	EN	RMSE
1	MRI-CT	[20]	5.18	49.17	16.64	24.40	4.76	**0.15**
		[21]	6.22	64.62	15.74	**27.32**	4.58	0.26
		Our	**6.41**	**65.84**	**17.18**	26.93	**5.15**	0.24
	MRI-SPECT	[20]	3.56	35.53	19.62	13.69	4.61	0.09
		[21]	3.92	41.09	19.89	13.76	4.60	0.11
		Our	**4.13**	**42.77**	**20.30**	**14.36**	**5.00**	**0.08**
2	MRI-CT	[20]	6.39	61.40	16.41	28.80	4.95	**0.15**
		[21]	7.66	80.01	15.53	31.35	4.76	0.26
		Our	**7.86**	**81.44**	**17.07**	**31.77**	**5.33**	0.24
	MRI-SPECT	[20]	3.89	38.31	18.81	14.29	4.74	0.09
		[21]	4.17	43.20	19.45	13.93	4.75	0.10
		Our	**4.45**	**45.59**	**19.69**	**14.92**	**5.20**	**0.08**

6. REFERENCES

[1] World Health Organization. World Health Day 2013: measure your blood pressure, reduce your risk [J]. http://www.who.int/mediaeentre/news/releases/2013/world_health_day_20130403/en, 2015.

[2] World Health Organization. A global brief on hypertension: Silent killer, global health crisis [J]. http://www.who.int/cardiovascular_diseases/publications/global_brief_hypertension/en/, 2016.

[3] Liu X, Mei W, Du H. Multimodality medical image fusion algorithm based on gradient minimization smoothing filter and pulse coupled neural network [J]. Biomedical Signal Processing and Control, 2016, 30: 140-148.

[4] Yerdelen D, Giray S, Tan M, et al. Hypertensive encephalopathy with atypical MRI leukoencephalopathy affecting brain stem and cerebellum [J]. Acta Neurol Belg, 2009, 109(2): 142-145.

[5] Hall D L, Llinas J. An introduction to multisensor data fusion [J]. Proceedings of the IEEE, 1997, 85(1): 6-23.

[6] Miao Q, Shi C, Xu P, et al. A novel algorithm of image fusion using shearlets [J]. Optics Communications, 2011, 284(6): 1540-1547.

[7] Tian Y, Luo J, Zhang W, et al. Multifocus image fusion in q-shift dtcwt domain using various fusion rules [J]. Mathematical Problems in Engineering, 2016, 2016.

[8] Kumar B K S. Image fusion based on pixel significance using cross bilateral filter[J]. Signal, image and video processing, 2015, 9(5): 1193-1204.

[9] Chai Y, Li H F, Guo M Y. Multifocus image fusion scheme based on features of multiscale products and PCNN in lifting stationary wavelet domain[J]. Optics Communications, 2011, 284(5): 1146-1158.

[10] Jin B, Jing Z, Pan H. Multi-modality image fusion via generalized Riesz-wavelet transformation[J]. KSII Transactions on Internet & Information Systems, 2014, 8(11).

[11] Unser M, Van De Ville D. Wavelet steerability and the higher-order Riesz transform [J]. IEEE Transactions on Image Processing, 2010, 19(3): 636-652.

[12] Chen C C. Fast boundary detection: A generalization and a new algorithm [J]. IEEE Transactions on Computers, 1977, 100(10): 988-998.

[13] Bhatnagar G, Liu Z. Multi-sensor fusion based on local activity measure [J]. IEEE Sensors Journal, 2017, 17(22): 7487-7496.

[14] Shahdoosti H R, Mehrabi A. Multimodal image fusion using sparse representation classification in tetrolet domain [J]. Digital Signal Processing, 2018.

[15] Rajasekhar G, Prasad V, Srikanth M V. Multilevel medical image fusion using multi-level local extrema and non-subsampled contourlet transformation [C]//Signal Processing and Communication (ICSPC), 2017 International Conference on. IEEE, 2017: 246-252.

[16] Singh S, Anand R S, Gupta D. CT and MR image information fusion scheme using a cascaded framework in ripplet and NSST domain [J]. IET Image Processing, 2018, 12(5): 696-707.

[17] Shabanzade F, Ghassemian H. Combination of wavelet and contourlet transforms for PET and MRI image fusion[C]//Artificial Intelligence and Signal Processing Conference (AISP), 2017. IEEE, 2017: 178-183.

[18] Du J, Li W, Lu K, et al. An overview of multi-modal medical image fusion [J]. Neurocomputing, 2016, 215: 3-20.

[19] Johnson, K.A., Becker, J.A. 2017. The Whole Brain Altas [Online], http://www.med.harvard.edu/aanlib/.

[20] El-Hoseny H M, Rabaie E S M E, Elrahman W A, et al. Medical image fusion techniques based on combined discrete transform domains[C]//Radio Science Conference (NRSC), 2017 34th National. IEEE, 2017: 471-480.

[21] Bhatnagar G, Wu Q M J, Liu Z. A new contrast based multimodal medical image fusion framework [J]. Neurocomputing, 2015, 157: 143-152.