Medical Image Fusion Based on Sparse Representation and Guided Filtering

Sa Huang¹, Guangyu Chu¹, Yifan Fei², Xiaoli Zhang² and Hailiang Wang³*

¹ Department of Radiology, The Second Hospital, Jilin University Changchun, Jilin 130041, China.
² Key Laboratory of Symbol Computation and Knowledge Engineer of Ministry of Education, Jilin University, Changchun 130012, China
³ Department of Neurosurgery, The Second Hospital, Jilin University Changchun, Jilin 130041, China
* Corresponding author, Email: wanghailiang1977@sina.com

Abstract. In this paper, we proposed a medical image fusion algorithm based on sparse representation and guided filtering. One of attractive features in the algorithm is that it can preserve the structural information of structural image and color information of functional image. We use a sparse representation in low frequency, initialize the dictionary by using DCT transform, and train the dictionary with each input source image as a training example. It not only ensures time complexity, but also ensures that the low-frequency fusion rules are adaptive. At high frequencies, we use the method of guided filtering to extract structural information from high-frequency images, and use the injection method to fuse high-frequency sub-bands to ensure the validity and richness of structural information. Experimental results show that the proposed fusion algorithm is superior to comparative algorithms in terms of subjective and objective evaluation methods.

1. Introduction
With the rapid development of medical imaging technology, various algorithms for image fusion have been proposed. This paper mainly studies the fusion of functional images (such as MRI and CT) and functional images (such as PET, SPECT and CBF). The main purpose is to integrate the complementary information of multi-source images and reduce redundant information to obtain high-quality fused images.

Until now, many methods have been proposed, among which some are dedicated to medical image fusion [1-16]. This paper aims at designing a pixel-level fusion algorithm. Multiscale image fusion accounts for a large percentage of the existing algorithms. There are a plenty of multiscale image presentation models which have been utilized in image fusion. Pyramid family (including Laplacian, contrast, ratio pyramid and the like) are widely used in early stage. Then wavelets are brought into this field [17, 18]. Recently, several geometric extensions of wavelets are introduced into the image fusion, such as Shearlet Transform [3, 4], Curvelet transform [16] and Contourlet transform [19]. In recent years, the theory of compressive sensing (CS) is brought into image fusion. In [20], the authors proposed a new fusion approach for low spatial resolution multispectral and the high spatial resolution Panchromatic (Pan) image. In this paper, we proposed a new multiscale fusion algorithm for structural and functional images.
2. The Proposed Image Fusion Algorithm

In this paper, a hybrid algorithm is proposed to fuse medical images. In order to preserve the color information of functional images, we first map the functional image to the HSV color space, and then multi-scale decomposition of the V channel using the multi-scale analysis tool NSCT is performed on structural image and V channel of functional image to obtain their low frequency sub-bands and the high frequency sub-bands, respectively. For the low frequency sub-band, we use the sparse representation as the fusion rule. For the high frequency sub-band, we propose a fusion rule based on the guided filter. After that, the fusion result of the v channel is obtained by using the inverse NSCT transform. Finally, the image is transformed from the HSV space to the RGB space to obtain the final fused image.

2.1. Sparse Representation

Sparse representation is a useful tool for signal analysis, and overcomplete sparse representation is to map signals through an overcomplete dictionary. Sparse representation needs to construct of an overcomplete dictionary $D$. An original signal $X$ can be described by the overcomplete dictionary $D$ by the sparse coefficient $\alpha$.

For a signal $X \in R^n$, where $n$ is the dimension of the signal, if the overcomplete dictionary is represented as $D \in R^{nk \times k}$, $k > n$, the original signal can be represented as $X \approx Da (\alpha \in R^k)$. Here, $\alpha$ is an unknown sparse coefficient, then the related problem can be transformed into the following problem

$$\min_{D,\alpha} \sum_{n=0}^{K} \| \alpha_m \|_0 \quad s.t. \| X_m - Da_m \|_2 < \varepsilon, m \in \{1,2,\ldots,K\}$$

where $X_m$ denotes the $m_{th}$ instance, $D$ is overcomplete dictionary, $\alpha_m$ is the sparse representation coefficient of the signal $X_m$, and $\varepsilon$ is the error rate. The right side of the whole expression shows that the sparse coefficient should be as consistent as possible with the original signal, and the left side shows that the sparse coefficient should be as sparse as possible. Eq. (1) can be solved by Orthognal Matching Pursuit (OMP). The basic idea is to perform multiple iteration calculations.

2.2. Fusion Rule for Low Frequency Sub-bands

Let $L_A$ and $L_B$ denote the low frequency sub-bands of source images $A$ and $B$ obtained by NSCT transform, respectively. The fusion rule for low frequency sub-bands are defined as follows:

1. Using the sliding window of $mn$ size, the matrices $L_A$ and $L_B$ are respectively segmented with a step size of 1, and the block matrices $V_A$ and $V_B$ are obtained, both of which are $n^2 \times m$.
2. Combine $V_A$ and $V_B$ to construct train dataset $S$, where the size is $n^2 \times 2m$.
3. Construct an overcomplete dictionary through the K-SVD algorithm. The specific process is as follows:

(a) Using DCT transform, initialize the dictionary $D \in R^{nk \times k}$.
(b) The train set is $S=\{S_i\}, i=1,\ldots,K$.
(c) The overcomplete dictionary learning process can be expressed by the following optimization formula

$$\min_{D,\alpha} \{\| S - DX \|_2^2 \} \quad s.t. \forall i, \| x_i \|_0 < T_0$$

where $D$ is the initialized dictionary, $X = \{X_i\}_{i=1}^{K}$ is the set of sparse coefficients, and $T_0$ represents the upper bound of the non-zero component of the sparse matrix.
(d) Use the OMP algorithm to obtain the sparse matrix $X$ under the initial dictionary.
(e) Fixed the sparse matrix $X$, we update each atom with singular value decomposition to minimize the approximation error value.
(f) Repeat steps 4 and 5 until the number of iterations reaches the predetermined number of...
iterations.

(g) Obtain the most matrix sparse $X$ and the optimal overcomplete dictionary $D$.

4. Using the obtained over-complete dictionary $D$, the sparse coefficients $X_A$ and $X_B$ of the low-frequency sub-bands $P_A$ and $P_B$ are reconstructed by OMP.

5. Obtain the fused low-frequency sparse coefficient matrix $L_F$:

$$
X_F^i = \begin{cases} 
X_A^i, & \|X_A^i\| > \|X_B^i\| \\
X_B^i, & \text{otherwise}
\end{cases}
$$

(3)

6. Transform $F$ to vector by $L_F = DX_F$, and then obtain the fused low frequency $L_F$.

2.3. Fusion Rule for High Sub-bands

Because the high frequency sub-band contains a wealth of spatial details, we propose a high frequency sub-band fusion rule based on guided filtering:

$$
H_F^i = H_A^i + \alpha(H_B^i - A^i_u), \ i = 1, ..., K
$$

(4)

where $H_F^i$ represents the i-th high-frequency fusion sub-band, $H_A^i$ represents the i-th high-frequency sub-band obtained by the NSCT transform of the functional image, and $H_B^i$ represents the ith high sub-band of the structural image obtained by NSCT transformation. $A^i_u$ represents the intermediate image processed by the guided filter. In order to get $A^i_u$, we use the guided filter to process $H_B^i$ with the $H_A^i$ as the guide image.

2.4. The Framework of the Proposed Fusion Algorithm

Assuming there are two source images A and B. Let A denote the functional image and B donate the structural image. They are well denoised and registered. The flow chart of the fusion algorithm can be summarized as follows.

STEP 1. The functional image A is transformed from RGB space to HSV space, namely $A \rightarrow \{H, S, V\}$.

STEP 2. B and the V channel of A are decomposed into low-frequency and high frequency sub-bands using NSCT, namely $B \rightarrow \{L_B, H_B^1, ..., H_B^K\}$ and $V \rightarrow \{L_V, H_V^1, ..., H_V^K\}$.

STEP 3. low-frequency sub-bands are fused using guided filtering $\{L_B, L_V\} \rightarrow L_F$, and high-frequency sub-bands are fused using Sparse Representation $\{H_B^i, H_V^i\} \rightarrow H_F^i$.

STEP 4. fused sub-bands are combined using inverse NSCT, $\{L_B, H_B^1, ..., H_B^K\} \rightarrow F$.

STEP 5. The V channel of A is replaced by F, and the final fused image $\overline{F}$ is obtained by transforming from HSV space to RGB space, $\{H, S, F\} \rightarrow \overline{F}$.

3. Experimental Settings and Results

The proposed fusion algorithm is compared with five popular methods: Laplacian Pyramid based algorithm (LP), Discrete wavelet transform based algorithm (DWT), Nonsubsampled contourlet transform based algorithm (NSCT), Guided filtering based fusion algorithm (GFF), and LP-SR based algorithm. In the fusion algorithms LP, DWT, NSCT, source images are decomposed into four layers. The coefficients in approximation layers are combined by averaging, and the ones in detail layers are combined by selecting maximum absolute values.

3.1. Experimental Results and Analysis

In this subsection, we list the experimental results on the test data set. Fig.1 shows four examples of source images in the test set.
Figure 1. Four pairs of source images (in each pair, the first one is structural image, and the second one is functional image)

Because of the limitation of space, we only list one group of test results as examples in Fig.4.

As can be seen from the Fig.2, the fused image obtained by the LP method lacks some important spatial information, resulting in weaker detail display capability. Although the DWT method retains spectral information, it has significant texture loss. The LP-SR method retains the fineness of detail, but there is spectral distortion. The NSCT method also has excellent display capabilities, but it has severe spectral distortion. The algorithm in this paper does not have the defects of the above algorithm. The objective evaluation results of the six fused images are shown in Table 1.

It can be seen from Table 1 that the algorithms proposed in this paper are superior to other algorithms in CC, UIQI, RMSE, SSIM, ERGAS, and the performance of SAM is slightly insufficient.
Table 1. Objective evaluating results of the fusion algorithms in Fig.2

|     | CC    | UIQI  | RMSE | SAM | SSIM   | ERGAS |
|-----|-------|-------|------|-----|--------|-------|
| LP  | 0.7416| 0.6542| 43.2521|     | 1.2403 | 0.6694 | 5.5364 |
| DWT | 0.7758| 0.7792| 44.1029| 2.2840| 0.6541 | 6.3244 |
| NSCT| 0.6695| 0.5515| 50.6315| 4.1803| 0.5782 | 10.3950|
| GFF | 0.7490| 0.6168| 44.1440| 2.8370| 0.6699 | 7.2591 |
| LP-SR| 0.7537| 0.5568| 45.9652| 3.4181| 0.6571 | 14.5370|
| Proposed | **0.8128** | **0.8335** | **42.5671** | 1.9811 | **0.7022** | **5.3663** |

In order to check the performance of the proposed algorithm on various medical images, we list the averaged objective evaluation results in Table 2.

Table 2. Objective evaluating results of the fusion algorithms on image set “mountain”

|     | CC    | UIQI  | RMSE  | SAM | SSIM  | ERGAS  |
|-----|-------|-------|-------|-----|-------|--------|
| LP  | 0.8424| 0.8018| 16.6915|     | 2.0050 | 0.7921 | 5.5267 |
| DWT | 0.8549| 0.7466| 19.8842| 2.6854| 0.7703 | 7.3807 |
| NSCT| 0.8566| 0.7692| 18.2991| 3.0581| 0.7828 | 7.0226 |
| GFF | 0.8400| 0.7794| 16.6028| 3.2120| 0.7965 | 5.8823 |
| LP-SR| 0.8524| 0.8446| 16.8771| 2.5080| 0.7608 | 5.2104 |
| Proposed | **0.8769** | **0.8648** | **15.8039** | 2.3566| **0.8105** | **4.8242** |

By analysing the objective evaluation results, we found that the proposed algorithm works the best except on the measure SAM.

4. Conclusions
This paper proposes a medical image fusion algorithm based on sparse representation and guided filtering. We use a sparse representation in low frequency, initialize the dictionary by using DCT transform, and train the dictionary with each input source image as a training example. It not only ensures low-frequency time complexity, but also ensures that the low-frequency fusion rules are adaptive. At high frequencies, we use the method of guided filtering to extract spatial information from high-frequency images, and use the injection method to fuse high-frequency sub-bands to ensure the validity and richness of spatial information. We have adjusted the parameters of the algorithm and evaluated it in both the subjective and objective ways. We found that this algorithm has good spectral information retention compared to other algorithms.

5. References
[1] Wang, Q Z, S Li, H Qin, and A M Hao 2015 Robust multi-modal medical image fusion via anisotropic heat diffusion guided low-rank structural analysis. Information fusion. 26 pp 103-121.
[2] Bhatnagar, G, Q M J Wu, and Z Liu 2015 A new contrast based multimodal medical image fusion framework. Neurocomputing. 157 pp 143-152.
[3] Wang, L, B Li, and L F Tian 2014 EGGDD: An explicit dependency model for multi-modal medical image fusion in shift-invariant shearlet transform domain. Information fusion. 19 pp 29-37.
[4] Wang, L, B Li, and L F Tian 2014 Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients. Information fusion. 19 pp 20-28.
[5] James, A P and B V Dasarathy 2014 Medical image fusion: A survey of the state of the art. Information fusion. 19 pp 4-19.

[6] Xu, Z P 2014 Medical image fusion using multi-level local extrema. Information fusion. 19 pp 38-48.

[7] Das, S and M K Kundu 2013 A Neuro-Fuzzy Approach for Medical Image Fusion. Ieee T Bio-Med Eng. 60 pp 3347-3353.

[8] Bhatnagar, G, Q M J Wu, and Z Liu 2013 Human visual system inspired multi-modal medical image fusion framework. Expert Syst Appl. 40 pp 1708-1720.

[9] Li, S T, H T Yin, and L Y Fang 2012 Group-Sparse Representation With Dictionary Learning for Medical Image Denoising and Fusion. Ieee T Bio-Med Eng. 59 pp 3450-3459.

[10] Das, S and M K Kundu 2012 NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency. Medical & Biological Engineering & Computing. 50 pp 1105-1114.

[11] Yang, L, B L Guo, and W Ni 2008 Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. Neurocomputing. 72 pp 203-211.

[12] Wang, Z B and Y Ma 2008 Medical image fusion using m-PCNN. Information fusion. 9 pp 176-185.

[13] Qu, G H, D L Zhang, and P F Yan 2001 Medical image fusion by wavelet transform modulus maxima. Optics Express. 9 pp 184-190.

[14] Suh, J W, O K Kwon, D Scheinost, A J Sinusas, G W Cline, and X Papademetris 2012 CT-PET weighted image fusion for separately scanned whole body rat. Med Phys. 39 pp 533-542.

[15] Daneshvar, S and H Ghassemian 2010 MRI and PET image fusion by combining IHS and retina-inspired models. Information fusion. 11 pp 114-123.

[16] Ali, F, I El-Dokany, A Saad, and F E-S Abd El-Samie 2008 Curvelet fusion of MR and CT images. Progress In Electromagnetics Research C. 3 pp 215-224.

[17] Li, H, B Manjunath, and S K Mitra 1995 Multisensor image fusion using the wavelet transform. Graphical models and image processing. 57 pp 235-245.

[18] Pajares, G and J M De La Cruz 2004 A wavelet-based image fusion tutorial. Pattern Recognition. 37 pp 1855-1872.

[19] Bhatnagar, G, Q Wu, and Z Liu 2013 Directive contrast based multimodal medical image fusion in NSCT domain. IEEE Transactions on Multimedia. 15 pp 1014-1024.

[20] Harikumar, V, P P Gajjar, M V Joshi, and M S Raval 2014 Multiresolution Image Fusion: Use of Compressive Sensing and Graph Cuts. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 7 pp 1771-1780.