Perceptual medical image fusion with internal generative mechanism

Jiansheng Qian, Rong Bao, Wei Shen, Junfeng Hu, Lu Tang✉ and Zhifang Xia

Medical image fusion is the process of integrating two medical images with a visual enhanced single fused image, to attain a resultant image richer in information to aid medical practitioners in better diagnosis. A perceptual medical image fusion method is proposed by employing Internal Generative Mechanism. First, source images are divided into a predicted layer and a detail layer with a Bayesian prediction model. Then, the detail layer is merged with the energy of Tchebichef moments for blocks while the predicted layer is fused using the averaging strategy as activity level measurement. The fused image is finally obtained by merging coefficients in both fused layers. Experimental results prove that the proposed fusion algorithm is superior to the previously developed methods.

Introduction: With the prompt advance of the computer technology, there are various medical images such as computed tomography (CT), magnetic resonance imaging (MRI), and ultrasonography imaging, single photon emission CT (SPECT) and positron emission CT (PET). Signal medical imaging modality is not enough to provide doctors with the ample information to diagnose disease. Medical image fusion offers an important approach by integrating complimentary features of different imaging modalities into one fused image. It is really helpful to doctors in clinical diagnoses and treatment planning [1].

Various medical image fusion algorithms have been proposed during the past few years. These approaches were developed based on multi-scale transform (MST) [2], pulse coupled neural network (PCNN) [3] and sparse representation [4]. The MST-based methods produce very possibly lead to loss of information, pixel distortion [5]. The PCNN-based methods have many parameters, which are difficult to be reliably determined. The SR-based methods suffer from the limited ability in detail preservation and high sensitivity to misregistration [6].

Recently, researches on brain theory and neuroscience, the Bayesian brain theory [7] and the free-energy principle [8, 9] reported that the brain works with an internal generative mechanism (IGM) for visual perception and understanding. Inspired by the IGM theory about the active inference procedure [10–12], we adopt a Bayesian prediction-based linear autoregressive (AR) model in our method, the source images can be decomposed into a predicted layer and a detail layer with different physical features. The predicted layer represents the approximation component of the source images. The averaging strategy is applied to the this layer. On the other hand, the detail layer reflects the texture or edges information. Tchebichef moment is very efficient in shape representation, it can effectively capture edge features [13]. Therefore, the energy of Tchebichef moments for blocks is utilised to fuse this layer.

Proposed medical image fusion scheme: We suppose a medical image to be decomposed with the size of $M \times N$, i_{xy} is the intensity of the pixel at the xth row and yth column. The mutual information $M(i_{xy}, i_{lk})$ between the pixels i_{xy} and i_{lk} is used as the AR coefficient, and a linear AR model is created to predict the value of the pixel i_{xy},

$$i'_{xy} = \sum_{i_{lk} \in W^K_{i_{xy}}} C_{lk} i_{lk} + \varepsilon \tag{1}$$

where i'_{xy} is the predicted value of pixel i_{xy}, $W^K_{i_{xy}}$ is the image block with the size of $K \times K$ which surrounds the pixel i_{xy}, ε is the additive Gaussian noise term

$$C_{lk} = \frac{M(i_{xy}, i_{lk})}{\sum_{X_{lk} \in W_{i_{xy}}} M(i_{xy}, i_{lk})}$$

being the normalised coefficient of i_{lk}. The whole predicted part of the source image I is denoted by (1)

$$I^P = \left\{ i'_{xy} | 1 \le x \le M, 1 \le y \le N \right\} \tag{2}$$

The detail part of I can be obtained by subtraction

$$I^D = I - I^P \tag{3}$$

Next, for the detail layer fusion, the theory of Tchebichef moments is adopted here to detect the saliency. Tchebichef moments are defined directly on image coordinates, the nth order, N-point Tchebichef kernel is defined as

$$t_n(x; N) = n! \sum_{k=0}^{n} (-1)^{n-k} \binom{N-1-k}{n-k} \binom{n+k}{n} \binom{x}{k} \tag{4}$$

weighted Tchebichef kernel is usually used

$$\tilde{t}_n(x; N) = \sqrt{\frac{\omega(x; N)}{\rho(n; N))}} t_n(x; N) \tag{5}$$

where

$$\omega(x; N) = \frac{1}{N+1}, \quad \rho(n; N) = \frac{(2n)!}{N+1} \binom{N+n+1}{2n+1}$$

are the weight and norm, respectively. The $(m+n)$th-order Tchebichef moments of an $M \times N$ image $f(x,y)$ is defined as

$$T_{mn} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \tilde{t}_m(x; M) \tilde{t}_n(y; N) f(x, y) \tag{6}$$

where $m \in \{0, 1, 2, \ldots, M-1\}$, $n \in \{0, 1, 2, \ldots, N-1\}$. In this Letter, we choose the energy of Tchebichef moments for blocks to detect the saliency of the detail layers of source images. Specifically, the image in detail layer is denoted by $I^D(x, y)$, $x \in \{1, 2, 3, \ldots, M\}$, $y \in \{1, 2, 3, \ldots, N\}$, $I^D(x, y)$ is divided into blocks with equal-size $D \times D$. The image block set is denoted by $\{B^{I^D}_{ij}\}$, $i \in \{1, 2, 3, \ldots, P\}$, $j \in \{1, 2, 3, \ldots, Q\}$, where $P = \lfloor M/D \rfloor$, $Q = \lfloor N/D \rfloor$, $\lfloor \cdot \rfloor$ is the floor operator. The Tchebichef moments of the image blocks in $\{B^{I^D}_{ij}\}$ are computed and denoted by $\{T_{ij}\}$ according to (6). The energy of Tchebichef moment values is computed and denoted by

$$E_{ij} = \sum_{p=0}^{m} \sum_{q=0}^{n} (T_{pq})^2 - (T_{00})^2 \tag{7}$$

where $m, n \in \{0, 1, 2, \ldots, (D-1)\}$, because the zeroth-order moment is the DC component of the image, $(T_{00})^2$ is removed.

We suppose that A and B are two medical images with different modalities, F is the fused image. The two source images A and B are represented in detail layers $\{A^D, B^D\}$ and the predicted layers $\{A^P, B^P\}$ by using IGM. F^D and F^P are the fused images of detail layers and predicted layers, respectively. The energy of Tchebichef moments for blocks $\{E^A_{ij}, E^B_{ij}t\}$ of $\{A^D, B^D\}$ is calculated by (7). The fusion result of detail layers is reconstructed by

$$F^D = \begin{cases} \max\left(A_D(x, y), B^D(x, y)\right) & E^A_{ij} = E^B_{ij} \\ A_D(x, y) & E^A_{ij} > E^B_{ij} \\ B^D(x, y) & E^A_{ij} < E^B_{ij} \end{cases} \tag{8}$$

The predicted layer represents the primary visual information of the source image. When source images are obtained with different modalities, as the grayscale values of the same position may vary significantly. Thus, merge A^P and B^P to obtain F^P with the averaging strategy as activity level measurement.

Finally, having F^D and F^P, the fused image F is reconstructed by

$$F = F^D + F^P \tag{9}$$

Experimental results and analysis: In our experiments, all preregistered source images have the same size of 256 × 256 pixels, with 256-level grey scale. Four pairs of medical images are shown in Fig. 1, which are divided into group a (CT and MRI), group b (MRI-T2 and PET), group c (B ultrasound and SPECT), group d (MRI-T2 and SPECT). The proposed method is compared with following five medical fusion algorithms, including guided filtering (GFF) [14], Laplacian pyramid-based using sparse representation (LP-SR) [4], NSCT-based using PCNN (NSCT-PCNN) [3], nuclear norm minimisation (NNM) [15]. Fig. 2 displays these four medical images fusion examples. The first one is CT and MRI image fusion, it is clear that the image fused by our algorithm not only preserves edge information but also improves the detail information. In contrast to the fused images in group b, the images fused by our method obtain better results. Edge, details and texture information are clearly visible, while the metabolic information is presented explicitly. In group c example, the proposed method obtains

better performance, other methods cannot fuse this type of medical images well. Compared with the fused images of each algorithm in group d example, the proposed fusion algorithm preserves detailed information and the metabolites edge information.

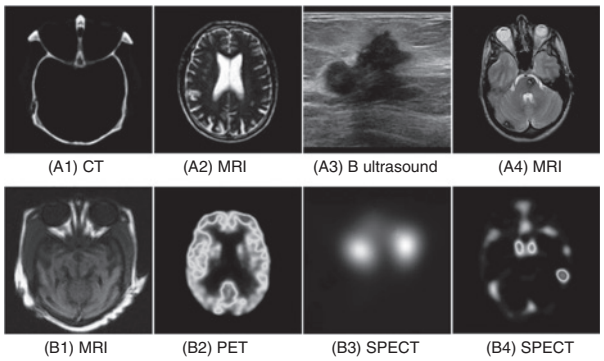

Fig. 1 *Source multimodal medical images in different groups*

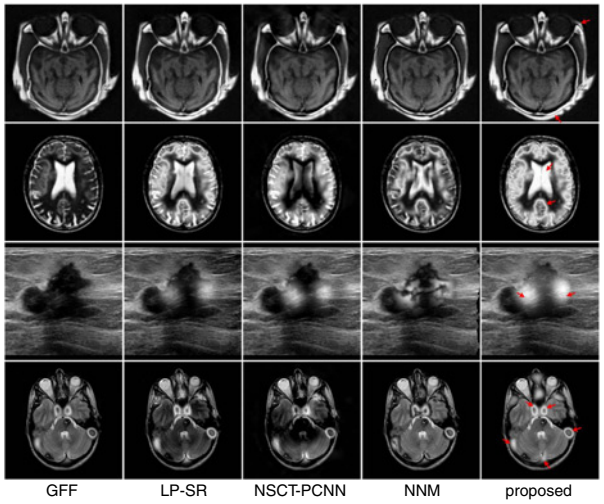

Fig. 2 *Four fusion examples in our experiments*

To evaluate the performance of the proposed fusion algorithm, four popular objective metrics, namely, mutual information MI [16], standard deviation (SD) [4], feature based metrics Q_G [17] and structure-based metrics Q_Y [18], are adopted to evaluate fusion performance. Table 1 reports the objective assessment results of different fusion methods on four groups medical images, where the best result is marked in boldface. It can be seen that the proposed method has considerable advantages over other methods of fusion tasks on each metric.

Table 1: Objective assessment of different fusion methods

Images	Metric	GFF	LP-SR	NSCT-PCNN	NNM	Proposed
Group a	MI	3.43	3.13	2.20	4.83	**5.78**
	SD	53.56	59.94	55.61	55.83	**59.96**
	Q_G	0.81	0.79	0.50	0.83	**0.91**
	Q_Y	0.89	0.85	0.72	0.90	**0.97**
Group b	MI	3.72	3.55	3.08	3.83	**4.56**
	SD	54.58	75.56	68.16	63.19	**81.81**
	Q_G	0.79	0.76	0.42	0.79	**0.82**
	Q_Y	0.52	0.45	0.37	0.48	**0.54**
Group c	MI	6.31	2.79	3.92	6.45	**7.60**
	SD	48.16	43.41	48.56	47.91	**50.72**
	Q_G	0.83	0.80	0.81	0.83	**0.84**
	Q_Y	0.89	0.88	0.91	0.87	**0.98**
Group d	MI	2.35	1.69	1.37	3.09	**3.54**
	SD	49.60	45.60	46.08	51.96	**55.94**
	Q_G	0.86	0.83	0.58	0.88	**0.91**
	Q_Y	0.46	0.40	0.36	0.46	**0.49**

Conclusion: In this Letter, we have presented a medical image fusion method with Internal Generative Mechanism, in which each source image is decomposed into a predicted layer and a detail layer. The energy of Tchebichef moments for blocks is employed to fuse detail parts, and the averaging strategy as activity level measurement is utilised to fuse the predicted layers. The fused image is finally obtained by combining coefficients in both fused layers. Experimental results demonstrate the advantages of the proposed methods over the state-of-the-art fusion methods and it can be used in clinical diagnosis.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (61379143), the Xu Zhou Science and technology Program (BY2013020, KC14SH078, KC15SH019).

Submitted: *23 May 2017* E-first: *1 August 2017*
doi: 10.1049/el.2017.1935
One or more of the Figures in this Letter are available in colour online.

Jiansheng Qian, Wei Shen and Lu Tang (*China University of Mining and Technology, Xuzhou, Jiangsu 221116, People's Republic of China*)

✉ E-mail: tanglu@cumt.edu.cn

Rong Bao (*Xuzhou University of Technology, Xuzhou, Jiangsu 221018, People's Republic of China*)

Junfeng Hu and Lu Tang (*Xuzhou Medical University, Xuzhou, Jiangsu 221004, People's Republic of China*)

✉ E-mail: tanglu@cumt.edu.cn

Zhifang Xia (*State Information Center of P.R. China, Beijing, People's Republic of China*)

References

1 Paulino, A. C., Thorstad, W. L., and Fox, T.: 'Role of fusion in radiotherapy treatment planning', *Semin. Nucl. Med.*, 2003, **33**, (3), pp. 238–243
2 Yang, Y., Song, T., Huang, S., *et al.*: 'Log-Gabor energy based multimodal medical image fusion in NSCT domain', *Comput. Math. Methods Med.*, 2014, **2014**, (2), pp. 1–12
3 Das, S., and Kundu, M.: 'NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency', *Med. Biol. Eng. Comput.*, 2012, **50**, (10), pp. 1105–1114
4 Liu, Y., Liu, S., and Wang, Z.: 'A general framework for image fusion based on multi-scale transform and sparse representation', *Inf. Fusion*, 2015, **24**, pp. 147–164
5 Bhatnagar, G., Wu, Q., and Liu, Z.: 'A new contrast based multimodal medical image fusion framework', *Neurocomputing*, 2015, **2015**, (157), pp. 143–152
6 Liu, Y., Chen, X., Ward, R. K., *et al.*: 'Image fusion with convolutional sparse representation', *Signal Process. Lett.*, 2016, **23**, (12), pp. 1882–1886
7 Knill, D. C., and Pouget, A.: 'The Bayesian brain: the role of uncertainty in neural coding and computation', *Trends Neurosci.*, 2004, **27**, (12), pp. 712–719
8 Gu, K., Zhai, G., Lin, W., *et al.*: 'Visual saliency detection with free energy theory', *Signal Process. Lett.*, 2015, **22**, (10), pp. 1552–1555
9 Gu, K., Zhai, G., Yang, X., *et al.*: 'Using free energy principle for blind image quality assessment', *Trans. Multimed.*, 2015, **17**, (1), pp. 50–63
10 Gu, K., Zhai, G., Yang, X., *et al.*: 'Hybrid no-reference quality metric for singly and multiply distorted images', *Trans. Broadcast.*, 2014, **60**, (3), pp. 555–567
11 Gu, K., Zhai, G., Lin, W., *et al.*: 'Learning a blind quality evaluation engine of screen content images', *Neurocomputing*, 2016, **196**, pp. 140–149
12 Gu, K., Zhai, G., Lin, W., *et al.*: 'No-reference image sharpness assessment in autoregressive parameter space', *Trans. Image Process.*, 2015, **24**, (10), pp. 3218–3231
13 Li, L., Lin, W., Wang, X., *et al.*: 'No-reference image blur assessment based on discrete orthogonal moments', *Trans. Cybern.*, 2016, **46**, (1), pp. 39–50
14 Li, S., Kang, X., and Hu, J.: 'Image fusion with guided filtering', *Trans. Image Process.*, 2013, **22**, (7), pp. 2864–2875
15 Liu, S., Zhang, T., Li, H., *et al.*: 'Medical image fusion based on nuclear norm minimization', *Int. J. Imaging Syst. Technol.*, 2015, **25**, (4), pp. 310–316
16 Qu, G., Zhang, D., and Yan, P.: 'Information measure for performance of image fusion', *Electron. Lett.*, 2002, **38**, (7), pp. 313–315
17 Xydeas, C., and Petrovic, V.: 'Objective image fusion performance measure', *Electron. Lett.*, 2000, **36**, (4), pp. 308–309
18 Yang, C., Zhang, J., Wang, X., *et al.*: 'A novel similarity based quality metric for image fusion', *Inf. Fusion.*, 2008, **9**, (2), pp. 156–160