Medical image fusion based on pulse coupled neural networks and multi-feature fuzzy clustering

Xiaoqing Luo, Xiaojun Wu

School of IoT Engineering, Jiangnan University, Wuxi, China
Email: xiaojun_wu_jnu@163.com, xiaoqing_luo_jnu@163.com

ABSTRACT

Medical image fusion plays an important role in clinical applications such as image-guided surgery, image-guided radiotherapy, noninvasive diagnosis, and treatment planning. In order to retain useful information and get more reliable results, a novel medical image fusion algorithm based on pulse coupled neural networks (PCNN) and multi-feature fuzzy clustering is proposed, which makes use of the multi-feature of image and combines the advantages of the local entropy and variance of local entropy based PCNN. The results of experiments indicate that the proposed image fusion method can better preserve the image details and robustness and significantly improve the image visual effect than the other fusion methods with less information distortion.

Keywords: PCNN; Multi-Feature; Medical Image; Image Fusion; Local Entropy

1. INTRODUCTION

Image fusion is a kind of information fusion, which refers to the synergistic combination of different sources of sensory image into a single image. The fused image can reflect multiple properties of source images, which makes it more suitable for the purpose of human visual perception and computer processing tasks [1,2]. Recently, multimodal medical image fusion has attracted more attentions due to the increasing demands of clinical applications. Availability of numerous kinds of biomedical sensors makes the medical image fusion a practical topic for researchers and scientists in this field. X-ray, ultrasound, magnetic resonance imaging (MRI) and computed tomography (CT) are a few examples of biomedical sensors. These sensors are used for extracting clinical information, which are generally complementary in nature [1]. Thus, in order to extract more information, medical image fusion combines these contrasting and complimentary features into one fused image. Medical image fusion

not only helps in diagnosing diseases, but it also reduces the storage cost by reducing storage to a single fused image instead of multiple-source images [3].

Image fusion can occur at three different levels: pixel level, feature level, and decision level respectively. Generally, pyramid and wavelet transforms are commonly used for multi-resolution image fusion, such as gradient pyramid [4], morphological pyramid [5], DWT pyramid [6], contrast pyramid [7] etc. In the past few years, several researchers have proposed different image fusion methods. Wei proposed a multi-focus image fusion method based on the selection of image blocks from source images by using the energy of image Laplacian [8]. PCNN model derives from the research on cat visual cortex. Its mode of information processing is much closer to the mode of human visual processing. And then PCNN has the flexible structure which can be changed according to different tasks. Additionally, the existing PCNN methods also show PCNN has the higher performance. Su and Wu proposed a image fusion method based on multi-feature fuzzy clustering, which belong to feature level fusion [9]. In this paper, we focus on the image fusion based on the combination of multifeature extraction and PCNN.

The principle of the new image fusion method can be illustrated by **Figure 1**, where A and B are input medical images, F represents the fused image. One of the source images A is divided into windows. The features of image is defined as $f_{num\times6}$, where *num* is the number of windows and 6 is the number of element of feature vector. We use FCM to cluster the feature vectors of image to obtain regions. A and B are decomposed by wavelet transform. l_A and l_B are the low frequency sub-images, y_A and y_B are the high frequency sub-images. The segmentation result is mapped to y_A and y_B. The fusion of l_A and l_B are based on pixel and get by PCNN. The fusion of y_A and y_B are based on region and get by Cauchy similarity of corresponding region. l_F is the low frequency sub-image and y_F is the high frequency sub-image. F can be obtained by inverse wavelet transform.
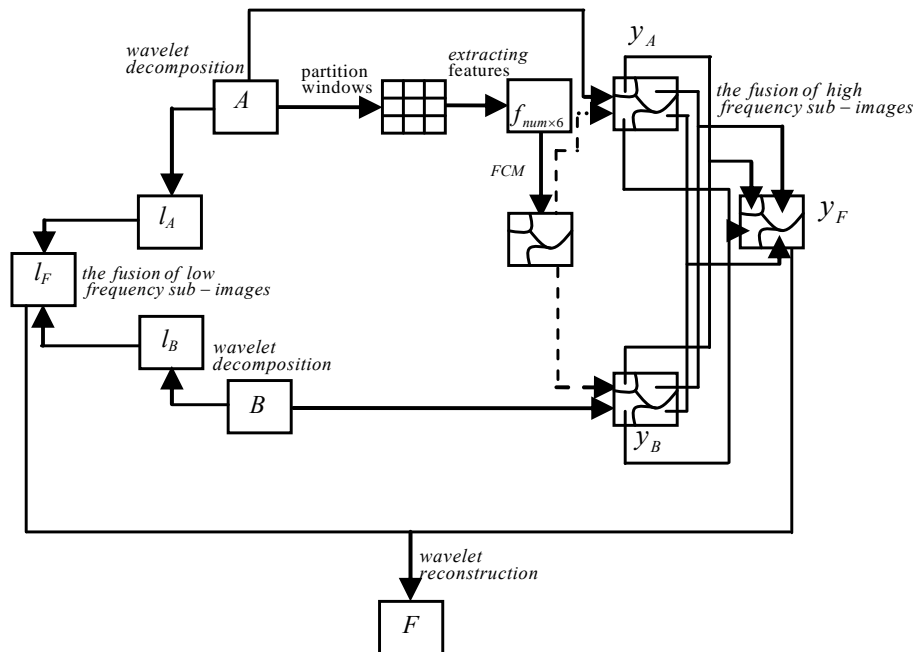
Figure 1. Scheme of the proposed image fusion method.

The proposed method designs more reasonable fusion rules using the global information and multi-feature of image. Compared to weighted average, maximum or region-energy-based rules, the proposed method can obtain better fusion result. The experimental results show that the new algorithm can provide better performance in medical image fusion.

The remainder of the paper is organized as follows: The fusion rule of low frequency sub-images is discussed in Section 2. Section 3 shows the fusion processing of high frequency sub-images. Experimental results and performance evaluations are given in Sections 4. Finally, conclusions are drawn in Section 5.

2. THE FUSION RULE OF LOW FREQUENCY SUB-IMAGES

The medical images are transformed using wavelet. The fusion of low frequency sub-images is performed using average weighted fusion rule. Let the low frequency subband of two medical images be l_A and l_B respectively, the fusion result of low frequency subband is l_F. i and j are the position of pixels. The fusion rule is given as follows:

$$l_F(i,j) = w_1 l_A(i,j) + w_2 l_B(i,j)$$
$$w_1 + w_2 = 1 \tag{1}$$

It is important that define the weight (w_1 and w_2). In view of the global coupled property and pulse synchronization characteristics of the PCNN [10,11], we use a simplified model and discrete form to determine weights.

In the simplified PCNN model, the PCNN neuron

consists of three parts: the dendritic tree, the linking modulation, and the pulse generator [8]. The role of the dendritic tree is to receive the inputs from two kinds of receptive fields. Depending on the type of the receptive field, it is subdivided into two channels (the linking and the feeding). The linking channel receives local stimulus from the output of surrounding neurons, while the feeding channel, besides local stimulus, still receives external stimulus. In the following expressions, the indexes i and j refer to the pixel location in the image, k and l refer to the dislocation in a symmetric neighborhood around one pixel, and n denotes the current iteration(discrete time step). Here n varies from 1 to N (N is the total number of iterations)

$$\begin{cases} F_{ij}(n) = S_{ij}(n) \\ L_{ij}(n) = \exp(-\alpha_L)L_{ij}(n-1) + V_L \sum_{k,l} m_{ij,kl} Y_{ij}(n-1) \\ U_{ij}(n) = F_{ij}(n) \times (1 + \beta L_{ij}(n)) \\ Y_{ij}(n) = \begin{cases} 1 & U_{ij}(n) > T_{ij}(n) \\ 0 & U_{ij}(n) \le T_{ij}(n) \end{cases} \\ T_{ij}(n) = \exp(-\alpha_T)T_{ij}(n-1) + V_T Y_{ij}(n) \end{cases} \tag{2}$$

The dendritic tree is given by $F_{ij}(n)$ and $L_{ij}(n)$. The two main components F and L are called feeding and linking, respectively. $m_{ij,kl}$ are the synaptic weight coefficients and S is the external stimulus. V_L is normalizing constants. α_L is the time constants. The linking modulation is given in the 3rd row of **Eq.2**, where $U_{ij}(n)$ is the internal state of the neuron. β is the

linking parameter and the pulse generator determines the firing events in the model in the 4-th row of **Eq.2**. $Y_{ij}(n)$ depends on the internal state and threshold. The dynamic threshold of the neuron is the 5-th row of **Eq.2**, where V_T and α_T are normalized constant and time constant, respectively.

Suppose that the size of each low frequency sub-image is $M \times N$. The corresponding PCNN is then designed for them, and the size of each PCNN is $M \times N$. For each neuron, the feeding input is the intensity of the pixel corresponding to it. The number of neurons in the network is equal to the number of pixels in the input image. There exists a one-to-one correspondence between the image pixels and network neurons. Each pixel is connected to a unique neuron, and each neuron is connected to surrounding neurons.

β is the linking parameter. The pixel with significant feature should give strong response. Therefore, the setting of β should be related to the features of the corresponding pixels of the images. In this paper, local entropy(L_en) and the variance of local entropy of each of pixel (VL_en)are considered as the information of pixel [12]. We define a sliding window , the size of which is $m \times n$. ML_en is defined as the mean of local entropy of each of pixel. ML_en and VL_en are calculated with the sliding window and VL_en is regarded as the distribution feature of local entropy of the center pixel in the sliding window. L_en, ML_en and VL_en can be described as follows:

$$L\_en(i,j) = -\sum_{\omega,x} p_x \ln p_x \qquad (3)$$

$$\begin{aligned} &ML\_en(i,j) \\ &= \frac{1}{mn} \sum_{i'=-[m/2]}^{[m/2]} \sum_{j'=-[n/2]}^{[n/2]} L\_en(i+i')(j+j') \end{aligned} \qquad (4)$$

$$\begin{aligned} &VL\_en(i,j) \\ &= \frac{1}{mn-1} \sum_{i'=-[m/2]}^{[m/2]} \sum_{j'=-[n/2]}^{[n/2]} \big(L\_en(i+i')(j+j') \\ &\quad -ML\_en(i,j) \big)^2 \end{aligned} \qquad (5)$$

$$\beta(i,j) = L\_en(i,j) \times VL\_en(i,j) \qquad (6)$$

L_en can to some extend reflect the amount of information locally. VL_en shows the discrete degree of information distribution in local region, that is , the significant of image feature. **Eq.6** defines the linking parameter of each of pixel, which fully reflects the information of pixel. And β is not sensitive to noise because VL_en of noise is small due to the confusion of noise.

The fusion steps of low frequency sub-images using PCNN are described as follows:

1) Let $l_A(i,j)$ and $l_B(i,j)$ separately denote low frequency sub-images of the original medical images A and B. $l_A(i,j)$ and $l_B(i,j)$ are then normalized between 0 and 1, which are denoted as $l'_A(i,j)$ and $l'_B(i,j)$ respectively. $l'_A(i,j)$ and $l'_B(i,j)$ are treated as the feeding input to stimulate a corresponding PCNN.

2) Calculate the L_en and VL_en of $l'_A(i,j)$ and $l'_B(i,j)$ respectively, then, $\beta(i,j)$ can be obtained.

3) If $n = N$, then iteration stops. $T(i,j)$ is denoted as the total pulse value of the (i,j) neuron of the low frequency sub-image during n iterations. The fused weight can be defined as follows:

$$w_1 = T_A(i,j) / \big(T_A(i,j) + T_B(i,j) \big), w_2 = 1 - w_1 \qquad (7)$$

According to **Eq.1**, we can get the fusion result of low frequency sub-images.

3. THE FUSION RULE OF HIGH FREQUENCY SUB-IMAGES

The fusion rule of high frequency sub-images is performed based on multi-feature fuzzy clustering. First, the source images are divided into windows. Multi-feature is extracted from window and the fuzzy C-means (FCM) clustering algorithm [13] is used to segment the image. Second, average weighted fusion method is used to each of region. The weighting factors are constructed based on the local energy and the fuzzy similarity measure defined by Cauchy function [9] .The wavelet coefficients of the fused image are acquired by the weighting factors.

3.1. Multi-Feature of Images

To extract texture feature, the input image A is first segmented into $wm \times wn$ small windows. Let num be the number of windows in the output image. A feature vector is then extracted from each window, each feature vector, $f_w \big(f_w \in R^6, 1 \le w \le num \big)$, consists of 6 features. Three of them are the texture features, which are denoted by a wavelet transform applied to windows of the image A. The other two features are computed similarly from the LH and HH bands. The fourth feature is the gray value of image denoted by the mean value of window in the spatial domain. The other two are the x-coordinate and y-coordinate on the top left corner of the window, which denotes the spatial location. Finally, we obtain the feature vector f_w, and normalize f_w to $\hat{f}_w$.

3.2. Image Segmentation Based on FCM

Image segmentation is an essential step in image fusion. The well known method of fuzzy clustering is FCM Algorithm, which is an unsupervised fuzzy clustering.

FCM adopts the iterative algorithm to optimize objective function based by least square algorithm.

FCM algorithm is used to cluster the feature vectors $\hat{f}_w$, we obtain $\mu_{c,w}$ ($1 \leq c \leq c'$, $1 \leq w \leq num$), c' is the number of clusters, and $\mu_{c,w} \in [0,1]$ denotes the membership of the feature vector $\hat{f}_w$ in the c-th cluster.

After FCM clustering, the clusters are obtained as follows:

If $\mu_{ci,w} > \mu_{cl,w}$ ($cl = 1,2\cdots,c'$, $ci = 1,2,\cdots,c', cl \neq ci$) then the vector $\hat{f}_w$ belongs to the ci-th cluster. So the set of feature vectors is partitioned into c' groups $\{F_1,\cdots,F_{c'}\}$, and consequently, the image is segmented into c' regions $\{Rg_1,\cdots,Rg_{c'}\}$ with Rg_c being the region corresponding to the feature set F_c. The number of clusters is chosen by cross-validation technique.

3.3. Cauchy Fuzzy Similarity

The wavelet coefficients of the c-th region in R level and P band is defined as $y_c^{R,p}$. Let x denote the mean of window and U denote the mean set of windows ($x \in U$). The source image A and B are decomposed by wavelet, the memberships of fuzzy region A and B in a band are given by

$$\mu_{\tilde{A}}(x) = \frac{1}{1 + \left(\dfrac{\|\bar{x} - \bar{\mu}\|^\alpha}{d_a}\right)}, \mu_{\tilde{B}}(x) = \frac{1}{1 + \left(\dfrac{\|\bar{x} - \bar{v}\|^\alpha}{d_b}\right)} \quad (8)$$

where α determine the kurtosis of Cauchy function, $\bar{\mu}$ and $\bar{v}$ are the regional mean of

$y_{c,A}^{R,p}$ and $y_{c,B}^{R,p}$,

respectively. d_a and d_b are the mean distance between the centre of clustering.

$\bar{y}_{c,A}^{R,p}$ and $\bar{y}_{c,B}^{R,p}$

are the regional mean value.

$$d_a = \frac{2}{c'(c'-1)} \sum_{c=1}^{c'} \sum_{c+1}^{c'} \left\| \bar{y}_{c,A}^{R,p} - \bar{y}_{c+1,A}^{R,p} \right\|$$
$$d_b = \frac{2}{c'(c'-1)} \sum_{c=1}^{c'} \sum_{c+1}^{c'} \left\| \bar{y}_{c,B}^{R,p} - \bar{y}_{c+1,B}^{R,p} \right\| \quad (9)$$

The Cauchy fuzzy similarity is $S(\tilde{A}, \tilde{B})$ of fuzzy set A and B, that is, regional matching degree $m_{c,AB}^{R,p}$.

$$S(\tilde{A}, \tilde{B}) = \frac{(d_a + d_b)^\alpha}{(d_a + d_b)^\alpha + \|\bar{\mu} - \bar{v}\|^\alpha} \quad (10)$$

Let $E_c^{R,p}$ be the regional information energy. If w_num is the window number of region c, $E_c^{R,p}$ can be defined as follows:

$$E_c^{R,p} = \sum_{wj=1}^{w\_num} \left| \bar{y}_{c,wj}^{R,p} \right|^2 \quad (11)$$

3.4. Decision Method

The decision method of image fusion directly affects the quality of fusion result. Let $y_{c,F}^R(n/p)$ is the wavelet coefficients of the c-th region of fused image in R level and P band.

$$y_{c,F}^R(i, j/p) = \omega_A^{R,p} y_{c,A}^R(i, j/p) + \omega_B^{R,p} y_{c,B}^R(i, j/p)$$
$$\omega_A^{R,p} + \omega_B^{R,p} = 1 \quad (12)$$

where T is the threshold of $m_{c,AB}^{R,p}$ and $\omega^{R,p}$ is the weighted factor.

$\omega_A^{R,p}$ is defined as follows:

$$\omega_A^{R,P} = \begin{cases} 1 & \text{if } m_{c,AB}^{R,P} \leq T \text{ and } E_{c,A}^{R,P} > E_{c,B}^{R,P} \\ 0 & \text{if } m_{c,AB}^{R,P} \leq T \text{ and } E_{c,A}^{R,P} \leq E_{c,B}^{R,P} \\ \dfrac{1}{2} + \dfrac{1}{2}\left(\dfrac{1 - m_{c,AB}^{R,P}}{1-T}\right) & \text{if } m_{c,AB}^{R,P} > T \text{ and } E_{c,A}^{R,P} > E_{c,B}^{R,P} \\ \dfrac{1}{2} - \dfrac{1}{2}\left(\dfrac{1 - m_{c,AB}^{R,P}}{1-T}\right) & \text{if } m_{c,AB}^{R,P} > T \text{ and } E_{c,A}^{R,P} \leq E_{c,B}^{R,P} \end{cases} \quad (13)$$

4. THE FUSION PROCEDURE AND EXPERIMENTAL RESULTS

4.1. The Fusion Procedure

The proposed method can be summarized as follows:

1) Select one of medical images to perform multi-feature extraction as discussed in Section 3.1.

2) We use FCM to segment image based on the results of multi-feature extraction and the segmentation results are mapped into other medical image to be fused.

3) Medical images are decomposed by wavelet transform into low frequency and high frequency bands respectively.

4) The fusion rule of low frequency sub-images adopts average weighted method based on pixel level. The weighted factor is obtained by PCNN. Then we calculate the L_en and VL_en for each pixel, which is used to define β.

5) The fusion rule of high frequency sub-images is the average weighted method based on region. The weighted of region can be got by the regional similarity and region energy.

6) The fused image can be got by taking the inverse wavelet transform of the fusion result of low frequency sub-images and high frequency sub-images.

4.2. Experimental Results

Medical imaging is taking on an increasingly critical role in healthcare, which has enabled radiologists to quickly acquire images of the human body and its internal struc-

tures with effective resolution and realism. These images are often known as multimodality medical images such as computed tomography (CT), ultrasound, positron emission tomography (PET), nuclear magnetic resonance (NMR) etc.. In our experiments, we evaluate the performance of our proposed method by CT and MR images, which come from the web of http://www. imagefusion.org/. The excellent views of bones and other dense structures are given by CT images whereas the excellent views of soft tissues are given perfectly by MR images. In certain situations, both soft and dense tissues are needed to be visualized simultaneously for better diagnosis. Hence, we need medical image fusion of various kinds for better treatment planning.

It is important to perform statistical assessment of the quality of different fusion techniques. Our method is compared to image fusion based on multi-feature fuzzy clustering [13] and image fusion based on DWT [6] and PCA [14]. Several indices are used for fused image quality assessment, including entropy (E), variance (V), mutual information (MI), spatial frequency (SF) and gradient (G) [15].

The size of original CT and MR images is 256×256, with 256-level gray scale. When wavelet transform is exploited, three decomposition levels and the wavelet basis "$sym4$" are used. For the proposed methods, the image block size is 4×4 pixels and $c' = 50$ and T = 0.6 in the experiment. The parameters setting are given as follows:

$$\alpha_L = 1, \alpha_\theta = 0.2, V_L = 1, V_\theta = 20, N = 250$$

$$W = \begin{bmatrix} 0.707 & 1 & 0.707 \\ 1 & 0 & 1 \\ 0.707 & 1 & 0.707 \end{bmatrix}$$

The visual comparison for fused images according to different fusion algorithms are shown in **Figures 2** and **3** respectively. From the two figures, it is clear that the proposed algorithm improves better detail information than the existing algorithms. The PCA algorithm gives baseline results because it is not a method based on multi-scale decomposition. Among multi-resolution based algorithms, our method based on multi-feature fuzzy clustering and PCNN has the best visual perception. The fusion method based on DWT lacks the ability that extracts image feature result in the degradation of fusion performance. Statistically, the overall performance and comparison among existing and proposed algorithms using quality metrics are depicted in **Tables 1** and **2** respectively. From **Table 1**, the E, V and MI of PCA are the best and the SF and G of our method are the best. Comparing **Figure 2(c)** with **Figure 2(a)**, we can find that they are very close that means the main information of **Figure 2(c)** comes from **Figure 2(a)**. In **Table 2**, all metrics except for E are the best for the
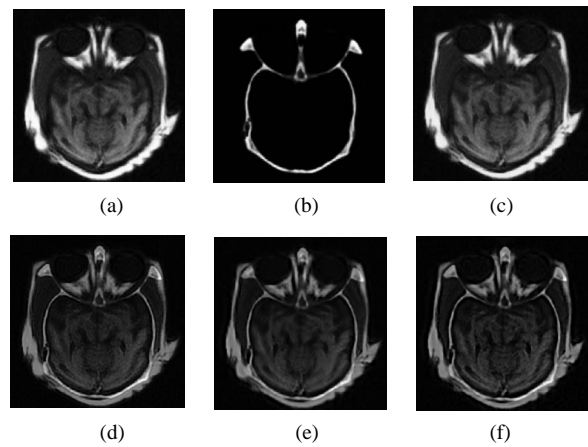

Figure 2. Fusion results for experimental image set 1; (a) heard 1 (MR); (b) heard 2 (CT); (c) fusion with PCA; (d) fusion with DWT; (e) Ref.[13]; (f) proposed method.
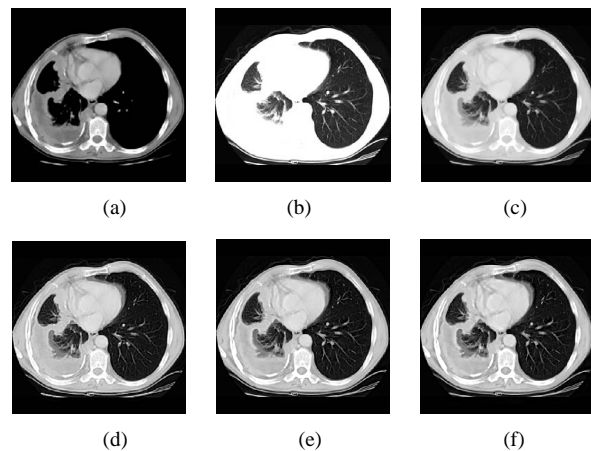

Figure 3. Fusion results for experimental image set 2; (a) head 1 (MR); (b) head 2 (CT); (c) fusion with PCA; (d) fusion with DWT; (e) Ref. [13]; (f) proposed method.

Table 1. The qualities of fusion results for Figure 2.

Images	E	V	MI	SF	G
Figure 2(c)	**4.572**	**56.131**	**6.323**	14.291	7.884
Figure 2(d)	4.198	36.823	3.347	14.221	7.677
Figure 2(e)	4.218	37.094	3.331	14.364	7.926
Figure 2(f)	4.244	38.475	2.721	**14.66**	**8.067**

Table 2. The qualities of fusion results for Figure 3.

Images	E	V	MI	SF	G
Figure 3(c)	4.478	95.538	6.842	25.018	11.657
Figure 3(d)	**4.555**	91.694	4.244	29.705	15.35
Figure 3(e)	4.529	86.608	4.319	31.458	16.056
Figure 3(f)	4.514	**91.939**	**4.335**	**33.355**	**17.072**

proposed method. From **Table 2**, we can easily conclude that the proposed algorithm can preserve high details. The possible reason behind the better performance is the combination of the local entropy and variance of local entropy based PCNN and multi-feature fuzzy clustering, which make more informative detail coefficients are produced when compared to other multi-resolution methods.

5. CONCLUSION

In this paper, a novel multi-modal medical image fusion algorithm based on PCNN and multi-feature fuzzy clustering is proposed. Considering the human visual system characteristics, two different fusion rules are adopted to fuse the low frequency and high frequency sub-images respectively. For the low frequency sub-images, the fusion weight coefficient based on PCNN. The local entropy and the variance of local entropy of each of pixel reflect the local information, which is used to calculate the linking strength. For the high frequency sub-images, the fusion weight based on region is got by multi-feature fuzzy clustering and Cauchy fuzzy similarity. The proposed algorithm can effectively preserve more information in the fused image with improved quality. The experiments are conducted on CT and MR for image fusion, two groups of CT and MR images are considered. The statistical comparisons demonstrate that the proposed algorithm better preserved the image details and significantly improved the image visual effect than the other fusion methods with very less information distortion. The objective statistical assessment findings agree with the visual perception. Although experiments are conducted only on CT and MR images, we believe that the proposed method can be used for other image fusion applications.

6. ACKNOWLEDGEMENTS

REFERENCES

[1] Bhatnagar, G., Wu, J.Q.M. and Liu, Z. (2012) Human visual system inspired multi-modal medical image fusion framework. *Expert Systems with Applications*, in press. doi:10.1016/j.eswa.2012.09.011

[2] Pajares, G. and Cruz, J. (2004) A wavelet-based image fusion tutorial. *Pattern Recognition*, **37**, 1855-1872. doi:10.1016/j.patcog.2004.03.010

[3] Wang, Z.B. and Ma Y.D. (2008) Medical image fusion using m-PCNN. *Information Fusion*, **9**, 176-185. doi:10.1016/j.inffus.2007.04.003

[4] Burt, P.J. (1992) A gradient pyramid basis for pattern-selective image fusion, *Society for Information Displays* (*SID*) *International Symposium Digest of Technical Papers*, **23**, 467-470.

[5] Toet, A. (1989) A morphological pyramidal image decomposition. *Pattern Recognition Letter*, **9**, 255-261. doi:10.1016/0167-8655(89)90004-4

[6] Li, H., Manjunath, B.S. and Mitra, S.K. (1995) Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, **57**, 235-245. doi:10.1006/gmip.1995.1022

[7] Toet, A., van Ruyven, L.J. and Valeton, J.M. (1989) Merging thermal and visual images by a contrast pyramid. *Optical Engineering*, **28**, 789-792. doi:10.1117/12.7977034

[8] Wang, Z.B., Ma, Y.D. and Gu, J.S. (2010) Multi-focus image fusion using PCNN. *Pattern Recognition*, **43**, 2003-2016. doi:10.1016/j.patcog.2010.01.011

[9] Su, D.X. and Wu, X.J. (2006) Image fusion based on multi-feature fuzzy clustering. *Journal of CAD and Graphics*, **18**, 838-843.

[10] Eckhorn, R., Reitboeck, H.J., Arndt, M. and Dicke, P.W. (1990) Feature linking via synchronization among distributed assemblies: Simulation of results from cat cortex. *Neural Computation*, **2**, 293-307. doi:10.1162/neco.1990.2.3.293

[11] Lindblad, T. and Kinser, J.M. (2005) Image processing using pulse-coupled neural networks. 2nd Edition, Springer, New York

[12] Zhang，Y.F. and He, M.Y. (2006) Image fusion algorithm based on local information entropy and its distribution property. *Computer Engineering*, **32**, 22-30.

[13] Nikhil, R.P., Kuhu, P. and James, M.K. (2005) A possibilistic fuzzy c means clustering algorithm. *IEEE Transactions on Fuzzy Systems*, **13**, 517-530. doi:10.1109/TFUZZ.2004.840099

[14] Zhang, Z. and Blum, R.S. (1999) A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proceedings of the IEEE*, **87**, 1315-1326. doi:10.1109/5.775414

[15] Luo, X.Q. and Wu, X.J. (2010) New metric of image fusion based on region similarity. *Optical Engineering*, **49**, 1-13. doi:10.1117/1.3394086