JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Research on the Multi-Focus Image Fusion Method Based on the Lifting Stationary Wavelet Transform

Kaiqun Hu* and Xin Feng*

Abstract

For the disadvantages of multi-scale geometric analysis methods such as loss of definition and complex selection of rules in image fusion, an improved multi-focus image fusion method is proposed. First, the initial fused image is quickly obtained based on the lifting stationary wavelet transform, and a simple normalized cut is performed on the initial fused image to obtain different segmented regions. Then, the original image is subjected to NSCT transformation and the absolute value of the high frequency component coefficient in each segmented region is calculated. At last, the region with the largest absolute value is selected as the post-fusion region, and the fused multi-focus image is obtained by traversing each segment region. Numerical experiments show that the proposed algorithm can not only simplify the selection of fusion rules, but also overcome loss of definition and has validity.

1. Introduction

Due to the limited depth of field of the optical lens, it is very difficult for the visible light imaging system to clearly define all the objects in the scene that have large differences in distance [1]. However, through the multi-focus images fusion methods, we can merge the targets of different focal lengths of the same scene into a new image with all the clear targets for human observation or computer processing [2].

At present, most of the research on multi-focus image fusion technology is based on the pixel layer. The algorithm can be roughly divided into two categories: spatial domain image fusion and transform domain image fusion. For the former, firstly treat the fusion images with some forms of multi-scale geometric transformation (wavelet transform, contourlet transform, shear wave transform, etc.) [3]. A certain fusion rule is used to combine the transformed coefficients to obtain the fused transform coefficients, and finally the inverse transform is performed to obtain the final fused image [4]. The fusion method of multi-scale transformation also has good fusion effect because it is similar to the process of the human visual system from coarse to fine recognition [5]. However, this kind of method has two obvious shortcomings: (1) there are different degrees of loss in the clear degree; (2) the selection

of the rules in the process of fusion is complex, and there is not a unified and effective multi-scale transformation fusion rule. In order to solve the above problems, this paper presents a new multi-focus image fusion method. The experimental results show that this method will not lose sharpness and avoid the complicated selection of fusion rules, which is effective.

2. The Fusion Method Proposed in This Paper

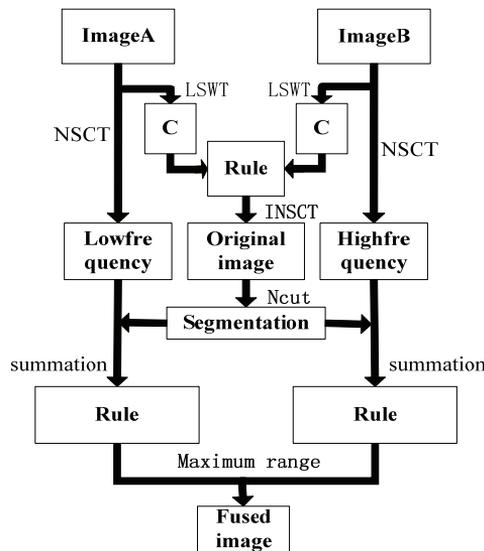The block diagram of the fusion method is shown in Fig. 1.

Fig. 1. Fusion method diagram of this paper.

According to block diagram, the main fusion steps are as follows:

1. According to lifting stationary wavelet transform (LSWT), quickly obtain the fused initial image, prepared for next step.
2. The fused initial image gets smoothing, and then normalized cut image segmentation is performed.
3. The source images are respectively nonsubsampled contourlet transform (NSCT) decomposed to obtain a coefficient matrix of the same size as the source image.
4. Find the sum of the absolute values of the high-frequency coefficients in each divided regions and select the region with the larger value as the corresponding region after the fusion.
5. Traverse all the divided areas to get the final fusion image.

2.1 Initial Fusion

LSWT transform is a flexible static wavelet transform method, which can use linear and nonlinear predictors to ensure the reversibility of the transform. Different from the classical wavelet transform, LSWT has the invariance of translation and better real-time performance. After LSWT decomposition,

the size of the sub-band image is the same as the size of the source image, which can better determine the relationship between the sub-band images and is very conducive to the implementation of the fusion algorithm [6]. Therefore, this paper selects LSWT for initial fusion.

2.2 Normalized Cut

Normalized cut [6] is a normalized form, which gets the highest similarity inside the sub-graph and the smallest similarity among the subgraphs. The normalized cut algorithm divides the initial fused image into different regions.

First, let $w(i, j)$ be the similarity between the pixel i and the pixel j. The value is mainly related to the brightness value, relative distance, texture and edge between two pixels. In this paper, only the pixel brightness and relative distance are considered. The smaller the brightness distance of the two pixels or the closer the spatial distance, the larger the value of $w(i, j)$ is. The weight is defined as:

$$w(i,j) = e^{\frac{-\|F_i - F_j\|_2^2}{\sigma_F^2}} * \begin{cases} e^{\frac{-\|X_i - X_j\|_2^2}{\sigma_X^2}} & \|X_i - X_j\|_2 < r \\ 0 & otherwise \end{cases} \tag{1}$$

where X_i and X_j represent the spatial positions of pixel i and pixel j, respectively, $\|X_i - X_j\|_2^2$ denotes the spatial distance between two pixels, F_i and F_j denote the pixel values of pixel i and pixel j, respectively.

According to literature [7], to minimize N-cut is equivalent to solving general matrix characteristic equation:

$$(D - W) y = \lambda D y \tag{2}$$

where W is the $n \times n$ symmetric matrix defined in Eq. (1), D is the diagonal matrix and $D(i, j) = d_i$. For the solution, using y as the indicator vector, the eigenvector corresponding to the second smallest eigenvalue of Eq. (2) is the solution of the disjoint subsets A and B.

Segmentation steps in this paper are as follows:

1. According to the formula 1, calculate the similarity between any two pixels in the image to be segmented, that is, the weight of the connection between two points, and establish a corresponding undirected graph;
2. Solve the eigenvector corresponding to the second smallest eigenvalue in Eq. (2);
3. According to the appropriate threshold, divide the eigenvector of the second eigenvalue into two parts, and divide the image into two parts A and B.
4. Determine whether it needs iterative segmentation. If, skip to the second step of iteration, or end.

2.3 Fusion Rules

An NSCT decomposition of the source image is performed to obtain the high-frequency detail component and the low-frequency approximation component, respectively. Among them, the high-frequency detail component represents the gradient information of the source image. The larger the absolute value, the stronger the change of the corresponding pixel value in the source image is, and the clearer the pixel value is. Therefore, this article calculates the sum of the absolute value of the high-

frequency part of the NSCT decomposition coefficient corresponding to each of the divided regions in each source image, so as to represent the clarity of the region in the source image. A larger value indicates that the block is clearer. Define the sharpness measurements for the k-th region in source images A and B are:

$$Q_k^A = \sum_{(i,j) \in ROA_k} c_l^A \tag{3}$$

$$Q_k^B = \sum_{(i,j) \in ROB_k} c_l^B \tag{4}$$

where ROA_k and ROB_k represent the k-th region in the source images A and B, respectively.

Then compare the sharpness of the corresponding area of the two images, and the larger the resolution, the more clearly the area will be clearer than the other image in its corresponding source image. Then directly select the clear area as the fusion of the corresponding area.

$$RoF_k = \begin{cases} RoA_k & Q_k^A \geq Q_k^B \\ RoA_k & Q_k^A < Q_k^B \end{cases} \tag{5}$$

where RoF_k represents the k-th region of the fused image F. Traverse all the divided areas to get the final fusion image.

3. Experimental Results

In order to verify the effectiveness of the proposed algorithm in multi-focus image fusion, we proceed simulation experiment with two groups of fully-focused multi-focus images by MATLAB software. All source images were 512×512 in size and 256 in gray scale. For comparison, the improved static wavelet transform (LSWT), the NSCT fusion method and the Shearlet fusion method are adopted. LSWT fusion method adopts the second decomposition A(l=2) taking the average of the simple fusion algorithm in the low frequency part, and taking large fusion rules of the modulus in the high frequency part. The NSCT transform uses 4-level decomposition, the directional decomposition filter adopts the "dmaxflat" filter, the scale decomposition filter adopts the "maxflat" filter, and the order of decomposition is 2, 3, 3, 4; the fusion rules is same as LSWT. Shearlet method employs parameters in the literature [8].
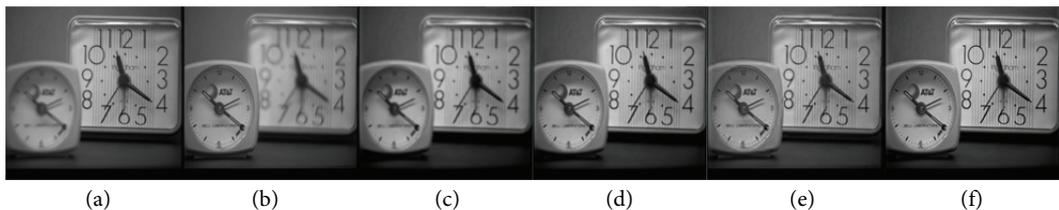

| (a) | (b) | (c) | (d) | (e) | (f) |

Fig. 2. The contrast of fusion result for multi-focus images (group 1). (a) Original image 1, (b) original image 2, (c) LSWT method, (d) NSCT method, (e) Shearlet method, and (f) the proposed method.

The first group of multi-focus image fusion results are shown in Fig. 2, in which Fig. 2(a) is a right clear original timepiece image and Fig. 2(b) is a left clear original timepiece image. Fig. 2(c)–(f) show the fusion results of multi-focus images by LSWT, NSCT, Shearlet and the proposed fusion method respectively. Due to the translational invariance, LSWT is better to restrain the distortion in image fusion, but because of the simple rule of selection, the clarity of the image to be fused cannot be well preserved. From the result, it can be seen that the writing on the dial is very fuzzy, but with high real-time.

The results of NSCT fusion have more edge information than LSWT, which is mainly due to the fact that NSCT transform is more powerful than wavelet in image edge representation. The Shearlet transform method, due to the unrestricted direction of the shearing operation and the ability to analyze the image in more directions, takes advantage of its fast coefficient degradation and the ability to approximate image edges at any scale, resulting in clearer detail of the edges and contrast. Based on the fusion of LSWT transform, the method of this paper selects a clear region of the source image through a simple normalized cut and introduces NSCT transform. The algorithm is simple and has almost no loss of sharpness. Therefore, this method has the highest definition and best edge detail information.
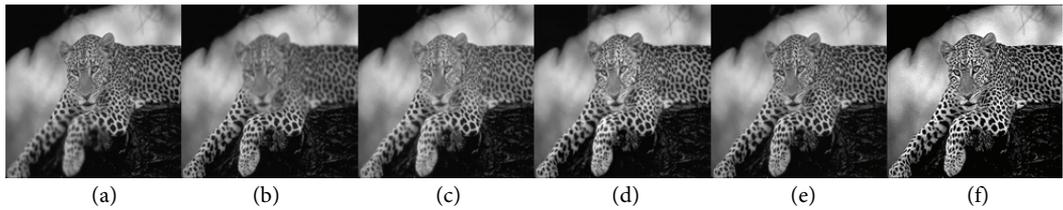

| (a) | (b) | (c) | (d) | (e) | (f) |

Fig. 3. The contrast of fusion result for multi-focus images (group 2). (a) Original image 1, (b) original image 2, (c) LSWT method, (d) NSCT method, (e) Shearlet method, and (f) the proposed method.

The second group of multi-focus image fusion results are shown in Fig. 3, where Fig. 3(a) and (b) are the original images. Fig. 3(c)–(f) show the fusion results of multi-focus images by LSWT, NSCT, Shearlet and the proposed fusion method, respectively. From the results of the fusion, several methods can well eliminate the focus difference of the source image, and obtain a clearer image with a foreground and a background. However, by careful comparison, we can find that the proposed algorithm has the highest fusion performance. The LSWT method is less sharp and has more obvious spurious information than the source image, especially in richly textured areas. In NSCT, due to the increase of translational invariance, false information is well suppressed. Shearlet method compared with the method of this paper, the apparent low contrast, and the goal of this method is not clear after the fusion. And the algorithm in this paper is also carried out in the NSCT domain with translational invariance, which can restrain false noise well. So comprehensively, this method has certain advantages both in definition and in phantom suppression.

Fig. 4 is a comparison of local images obtained by the fusion of all the methods. It can be clearly seen that the fusion result of the method not only preserves the edge characteristics but also overcomes the edge oscillation and has the best effect.

In order to objectively evaluate the fusion image quality, this paper selects information entropy (IE), mutual information (MI), peak signal-to-noise ratio (PSNR), edge information retention ($Q_{AB/F}$) and execution time T (s) as evaluation indicators.

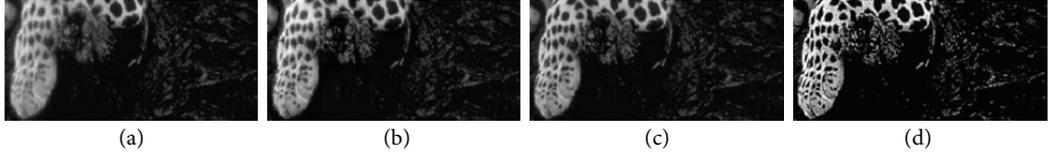|			
(a)	(b)	(c)	(d)

Fig. 4. The contrast of fusion result for local zoom images: (a) LSWT method, (b) NSCT method, (c) Shearlet method, and (d) the proposed method.

Image information entropy is an important indicator to measure whether the image information is rich, defined as:

$$H(F) = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{6}$$

where L is the grayscale of the image F and p_i is the distribution probability of the grayscale.

Mutual information defines the correlation between two images statistically. The mutual information between the reference image A and the fused image F is:

$$MI^{AF} = \sum_{i=1}^{L} \sum_{j=1}^{L} h_{AF}(i,j) \log_2 \frac{h_{AF}(i,j)}{h_A(i) h_F(j)} \tag{7}$$

where, $h_{AF}(i,j)$ is a normalized joint histogram of images A and , $h_A(i)$ and $h_F(j)$ are normalized histograms of the two images, and L is the number of gray levels. Similarly, the mutual information between the image B and the fused image F is denoted by MI^{BF}. The result of image fusion is evaluated by the sum of mutual information of each source image and fusion image, that is:

$$MI = MI^{AF} + MI^{BF} \tag{8}$$

The peak signal-to-noise ratio is mainly used to measure the vividness of images, which is defined as:

$$PSNR = 10 \times \lg \frac{(M \times N)^2}{\sum_{m=1}^{M} \sum_{n=1}^{N} [R(m,n) - F(m,n)]^2} \tag{9}$$

The higher the PSNR value, the better the fusion effect is.

Edge information retention ($Q_{AB/F}$), proposed by Xydeas and Petrovic, is used to evaluate the amount of edge information passed from the source image to the fused image.

It can be seen from Table 1 that the proposed method is the best on the four indicators of IE, MI, PSNR, and $Q^{AB/F}$ [9]. And it consumes only a little more running time than the LSWT method, which is much lower than the other two fusion methods of multi-scale geometric analysis. It proves that the method in this paper also has certain advantages in computational complexity.

Table 1. Quantitative comparison of the fusion results with different methods

	Method	IE	MI	PSNR	$Q^{AB/F}$	T/s
Group 1	LSWT	5.7632	3.6581	4.6526	0.6238	25
	NSCT	5.8526	3.5927	4.7629	0.6765	657
	Shearlet	5.8322	3.3245	4.6737	0.6863	401
	The proposed	6.0627	3.3639	4.8936	0.7023	104
Group 2	LSWT	4.6720	3.0576	4.5235	0.6023	31
	NSCT	4.7331	3.0227	4.6598	0.6365	697
	Shearlet	4.7236	2.9126	4.5226	0.6322	485
	The proposed	4.8597	2.8201	4.7289	0.6897	168

4. Conclusion

In this paper, aiming at the characteristics of the focused image, an improved fusion method is proposed. Firstly, Lifting Stationary Wavelet Transform is used to obtain the initial fused image quickly. After the simple normalized cut of the initial fused image, the sum of the absolute values of the coefficients of the NSCT transform is selected as the fusion rule. The experimental results show that this method will not lose the sharpness and avoid the complex selection of the fusion rules, which is effective.

Acknowledgement

References

[1] B. Liu, W. Liu, and J. Ma "Multi-focus image fusion based on three channel nonseparable symmetrical wavelets," *Chinese Journal of Scientific Instrument*, vol. 33, no. 5, pp. 1110-1116, 2012.

[2] G. Wang, M. Z. Ma, Y. L. Zhao, Q. Ma, L. Xiao, and A. He, "Algorithm for image fusion in the curvelet transform domain," *Chinese Journal of Scientific Instrument*, vol. 29, no. 9, pp. 1841-1845, 2008.

[3] S. Li, B. Yang, and J. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Information Fusion*, vol. 12, no. 2, pp. 74-84, 2011.

[4] L. Demanet and L. Ying, "Wave atoms and sparsity of oscillatory patterns," *Applied and Computational Harmonic Analysis*, vol. 23, no. 3, pp. 368-387, 2007.

[5] H. Zheng, C. Zheng, X. Yan, and H. Chen "Visible and infrared image fusion algorithm based on shearlet transform," *Chinese Journal of Scientific Instrument*, vol. 33, no. 7, pp. 1613-1619, 2012.

[6] W. Zhang, Q. Huang, and Z. Wang, "The fusion of remote sensing images based on lifting wavelet transformation," *Computer and Information Science*, vol. 2, no. 1, pp. 69-75, 2009.

[7]   T. Cour, F. Benezit, and J. Shi, "Spectral segmentation with multiscale graph decomposition," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, 2005, pp. 1124-1131.

[8]   Q. Miao, C. Shi, P. Xu, M. Yang, and Y. Shi, "Multi-focus image fusion algorithm based on shearlets," *Chinese Optics Letters*, vol. 9, no. 4, article no. 041001, 2011.

[9]   G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313-315, 2002.

Kaiqun Hu https://orcid.org/0000-0002-5590-2584

She received her Ph.D. degree from China Agriculture University in 2011. Now she is a lecturer in Chongqing Technology and Business University. Her main research direction is the technology of computer control and image processing.

Xin Feng https://orcid.org/0000-0001-8793-3775

He received Ph.D. degrees in School of Computer Science and Engineering from Lanzhou University of Technology in 2012. Since March 2013, he has been teaching in the School of Chongqing Technology and Business University.