International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.7, No.2 (2014), pp.399-410
http://dx.doi.org/10.14257/ijsip.2014.7.2.37

Multi-focus Image Fusion Based on Sparse Features

Yongxin Zhang[1,2], Li Chen[1], Zhihua Zhao [1] and Jian Jia [3]

[1]*School of Information Science and Technology, Northwest University, Xi'an, 710127, China*
[2]*Luoyang Normal University, Luoyang, 471022, China*
[3]*Department of Mathematics, Northwest University, Xi'an, 710127, China*
tabo126@126.com, chenli@nwu.edu.cn

Abstract

In order to effectively extract the focused regions from the source images and inhibit the blocking artifacts of the fused image, a novel adaptive block-based image fusion scheme based on sparse features is proposed. The source images are decomposed into principal and sparse matrices by a newly developed robust principal component analysis (RPCA) decomposition. The problem of multi-focus image fusion is transformed into a problem of choosing the sparse features of the sparse matrices to form a feature space. An optimal subdivision of blocks of the sparse matrices is obtained by using a quad tree structure to inhibit the blocking artifacts. The focused regions of the source images are detected by the local sparse feature of the blocks and integrated to construct the resulting fused image. Experimental results show that the proposed scheme can significantly inhibit the blocking artifacts and improve the fusion quality compared to the other existing fusion methods in terms of some objective evaluation indexes, such as structural similarity, mutual information and the edge information transferred from the source images to the fused image.

Keywords: *image fusion; block-based fusion; robust principal component analysis; quad tree structure*

1. Introduction

Image fusion aims to produce a single sharper image by combining a set of images captured from the same scene with different focus points [1]. In general, the image fusion methods can be categorized into two groups: spatial domain fusion and transform domain fusion [2]. This paper particularly focuses on the spatial domain methods.

The spatial domain fusion methods partition the source images into blocks or regions by using their region homogeneity, and detect the focused blocks or regions by using their local spatial features [3], such as energy of image gradient (EOG) [4] and spatial frequency (SF) [5]. Then, the focused blocks or regions are integrated to construct final fused image. The spatial domain methods are easy to implement and have low time consumption [1]. However, if the size of the blocks is too small, the blocks selection is too sensitive to noise and is subject to incorrect selection from the corresponding source images. Or else, if the size of the blocks is too large, the in-focus and out-of-focus pixels are partitioned in the same blocks, which are selected to build the final fused image. Accordingly, the blocking artifacts are produced and may compromise the quality of the final fused image. For eliminating the blocking artifacts, researchers have developed many improved schemes. Fedorov *et al.*, [6] have selected the best focus by titling source images with overlapping neighborhoods and improved the visual quality of the fused image. But this method is afflicted by temporal and

ISSN: 2005-4254 IJSIP

geometric distortions between images. Aslantas *et al.*, [7] have selected the optimal block-size by using differential evolution algorithm and enhanced the self-adaptation of the fusion method. But this method requires longer computational time. Wu *et al.*, [8] have selected the focused patches from the source images by using a belief propagation algorithm. But the algorithm is complicated and time-consuming. Goshtasby *et al.*, [9] have detected the focused blocks by computing the weight sum of the blocks. The iterative procedure is time-consuming. De *et al.*, [10] have determined the optimal block-size by using quad tree structure and effectively solved the problem of determining of block-size. These schemes all achieve better performance than the traditional methods and significantly inhibit the blocking artifacts. But they cannot eliminate the blocking artifacts completely.

In this paper, a novel image fusion scheme different from the methods mentioned above is proposed. This method is implemented in the robust principal component analysis (RPCA) decomposition domain. RPCA is a very important method of low-rank matrix recovery (LRMR) [11]. Wan *et al.*, [12] have investigated the potential application of RPCA in image fusion and achieved a consistently good fusion result. However, their method requires longer computational time for the application of sliding window techniques. Different from Wan's method, the main contribution of this paper is that a new method based on sparse features is proposed for image fusion. The principal and sparse matrices of the source images are obtained by using RPCA decomposition. The optimal partition of the source images is obtained by using the region homogeneity of the temporary sparse matrix. The focused blocks are detected by computing the salient feature of the corresponding sparse matrix blocks. The proposed method can significantly inhibit the blocking artifacts and better extract the focused image details from the source images.

The rest of the paper is organized as follows. In Section 2, the basic idea of RPCA and quad tree (QT) decomposition will be briefly described, followed by the new method based on RPCA and QT decomposition for image fusion in Section 3. In Section 4, extensive simulations are performed to evaluate the performance of the proposed method. In addition, several experimental results are presented and discussed. Finally, concluding remarks are drawn in Section 5.

2. Related Work

2.1. Robust Principal Component Analysis

RPCA has been proven to be an effective way to recover both low-rank and sparse components exactly from high dimensional data by solving the principal component pursuit [11]. In which, an input data matrix $D \in ¡^{M \times N}$ is subject to low-rank property. In order to recover the low-rank structure of D, D can be decomposed as:

$$D = A + E, \quad rank(A) = \min(M, N) \tag{1}$$

where matrix A is a principal matrix, and E is a sparse matrix. It is obvious that this problem is difficult to solve. Wright *et al.*, [13] have demonstrated that when the sparse matrix E is sufficiently sparse (relative to the rank of A), one can accurately recover the principal matrix A from D by solving the following convex optimization problem [14]:

$$\min_{A,E} \| A \|_* + \lambda \| E \|_1 \quad s.t. \ A + E = D \tag{2}$$

where $||\cdot||_*$ denotes the nuclear norm of matrix A, λ is a positive weighting parameter, and $||\cdot||_1$ denotes the l_1 norm of the matrix E.

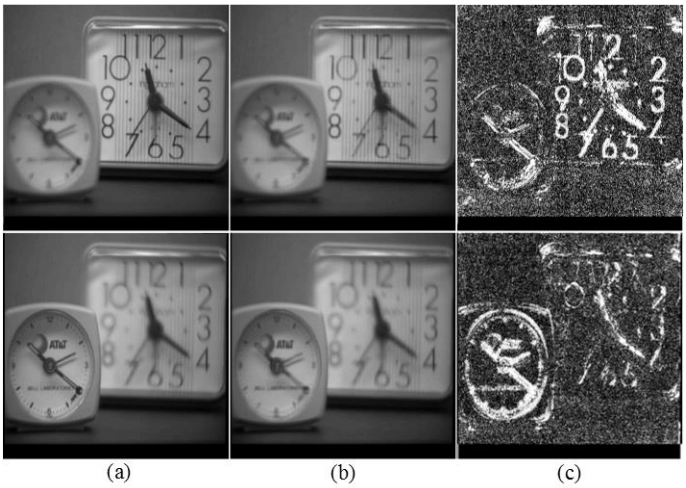

Figure 1. Decomposition of Multi-Focus Images 'Clock' using RPCA. (a) Source Images D, (b) Principal Matrix A, (c) Sparse Matrix E

Candes *et al.*, [11] have extended the RPCA to the background modelling from surveillance video. They correctly identified the moving pedestrians in the foreground by using the sparse component of surveillance video. The sparse matrix E effectively represents the salient feature of the foreground objects. As is known, the focused objects in the foreground are very important for multi-focus image fusion. Motivated by Candes's idea, this paper tries to establish a multi-focus image fusion scheme based on the RPCA decomposition. Figure 1 (a) shows the multi-focus source images 'Clock'. Figures 1 (b) and (c) show the corresponding images of the principal matrix A and sparse matrix E. It is obvious that the salient features of sparse matrix E agree well with the local features of the sharper objects in source images. Thus, the focused regions can be detected by comparing the salient features computed from the sparse matrix E.

2.2. Quad Tree Decomposition

QT is an important data structure where each internal node in the tree has exactly four children and each leaf node in the tree has no children. QT decomposition is an analysis technique which can partition an image into blocks which are more homogeneous than the image itself. In traditional QT decomposition, a square image can be partitioned into four equal sized blocks and then each block is evaluated by some threshold conditions of region homogeneity. The block that meets the threshold conditions will not be subdivided further, while the block that doesn't meet the threshold conditions will be subdivided into four blocks. And then the blocks are evaluated again iteratively until each block meets the threshold conditions [10, 15]. An example of the subdivision of an image in a QT structure is depicted in Figure 2 (a). The whole image is represented by a root node which is split into four blocks when its homogeneity doesn't meet the threshold conditions. The image block whose homogeneity meets the threshold conditions is represented by leaf node. In Figure 2 (a), I_0 is the image at level 0. After initial subdivision, $I_k(k = 1, \cdots, 4)$ are corresponding to regions at level 1. At level 1, the first and the third blocks, namely I_1 and I_3, are subdivided into

smaller blocks I_{1k} and $I_{3k}(k=1,\cdots,4)$ at level 2. According to the rule of QT decomposition, I_{1k} and I_{3k} are further subdivided if they meet the threshold conditions. Other quadrants will be subdivided similarly.

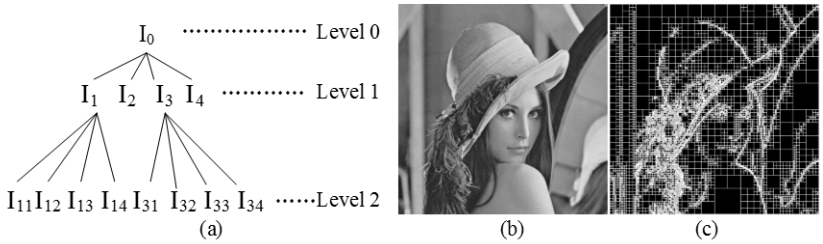

Figure 2. QT Decomposition of an Image. (a) Subdivision of an Image into quad Tree Structure, (b) Source Image 'Lena', (c) QT Decomposition Result of 'Lena'

In the process of QT decomposition, division is first performed on low resolution image. And then the subdivision is performed on high resolution image based on the division of low resolution image. It has been proved that the QT decomposition has the advantages of self-adaptation and high speed [15]. It is obvious that the salient features such as edges and textures of Figure 2 (c) are corresponding to the salient feature of Figure 2 (b). It can be seen that the block size of the subdivision in Figure 2 (c) is changing with the region homogeneity of Figure 2 (b). It is obvious that the QT decomposition can adaptively control the block size of subdivision of image based on the region homogeneity. So QT decomposition can be used to determine the optimal subdivision of blocks. It is useful for eliminating the blocking artifacts of the fused image. In this paper, the maximum differences of the elements are used as the region homogeneity of sparse matrices of the source images.

3. Proposed Method

According to the advantages of RPCA and QT decomposition, we try to establish a scheme to improve the accuracy of the focused regions detection while inhibit the blocking artifacts of the fused image. In this section, a novel adaptive method which does not use fixed block size of division in any portion of the multi-focus images is proposed. The method can obtain an optimal subdivision of the source images by performing QT decomposition on the sparse matrix of the source images. The proposed fusion framework is depicted in Figure 3.

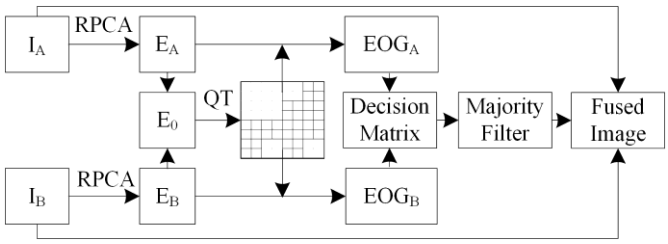

Figure 3. Block Diagram of Proposed Multi-Focus Images Fusion Framework

For simplicity, this paper assumes that there are only two source images, namely I_A and I_B, respectively. The rationale behind the proposed scheme applies to the fusion of more than

two multi-focus images. The source images are assumed to pre-registered and the image registration is not included in the framework. The fusion algorithm consists of the following seven steps:

Step 1: Construct the data matrix. The source images $\{I_A, I_B\}, I_A, I_B \in ¡^{M \times N}$ are converted into column vectors $I_A^c, I_B^c \in ¡^{MN \times 1}$, respectively. Thus, data matrix D is defined as:

$$D = [I_A^c \ I_B^c] \tag{3}$$

Step 2: RPCA decomposition is performed on D to obtain a principal matrix $A \in ¡^{MN \times 2}$ and a sparse matrix $E \in ¡^{MN \times 2}$, respectively. The sparse matrix $E \in ¡^{MN \times 2}$ is computed through inexact augmented Lagrange multipliers algorithm (IALM) of RPCA [11], which is an version implementation for recovering low-rank matrices and has been reported to yield similar results with much less time consumption. The sparse matrix $E \in ¡^{MN \times 2}$ is then converted into matrices $E_A, E_B \in ¡^{M \times N}$ corresponding to the source images I_A and I_B, respectively.

Step 3: Construct the temporary sparse matrix. Since the salient regions of sparse matrices $E_A, E_B \in ¡^{M \times N}$ are corresponding to the focused regions of the source images, respectively. Based on the salient region homogeneity of the sparse matrices, the better partitions of the sparse matrices $E_A, E_B \in ¡^{M \times N}$ are corresponding to the better partition of the source images. In order to better partition the sparse matrices, a temporary fusion sparse matrix E_0 ($E_0 \in ¡^{M \times N}$) is constructed by averaging the two matrices.

Step 4: Partition the temporary sparse matrix E_0 into blocks by using QT decomposition. QT decomposition is first performed on the temporary sparse matrix. The temporary sparse matrix is partitioned into blocks based on the region homogeneity of E_0. And then the two sparse matrices E_A and E_B are split based on the split results of E_0, respectively. To overcome the disadvantages of the small block in traditional block-based image fusion method, the minimum block size is set for terminating the further QT decomposition when the region homogeneity of the block doesn't meet the threshold condition. The region homogeneity is defined as:

$$|\max(E_{(i,j)}^R) - \min(E_{(i,j)}^R)| < T \tag{4}$$

where $E_{(i,j)}^R$ is the value of the element at the position (i, j) in sparse matrix. T is the threshold condition. In this paper, the threshold condition is set as 0.005 and the minimum block size is set as 8×8.

Step 5: Compute the sharpness of the blocks of the sparse matrices E_A and E_B, respectively. Let $E_A^{(k)}$ and $E_B^{(k)}$ denote the k th blocks of sparse matrices E_A and E_B, respectively. Let $EOG_k^{E_A}$ and $EOG_k^{E_B}$ be the EOG of $E_A^{(k)}$ and $E_B^{(k)}$, respectively. The EOG of each sparse matrix block can be defined as:

$$\begin{cases} EOG = \sum_i \sum_j (E_i^2 + E_j^2) \\ E_i = E(i+1, j) - E(i, j) \\ E_j = E(i, j+1) - E(i, j) \end{cases} \tag{5}$$

where $E(i, j)$ indicates the value of the element at the position (i, j) in sparse matrix block.

Step 6: Construct the decision matrix. The EOG of the corresponding blocks of the sparse matrices are compared to determine which pixel of the corresponding block of source images is in focus. A decision matrix $H \in ¡^{M \times N}$ is constructed for recording the comparison results according to the selection rule as follows:

$$H(i, j) = \begin{cases} 1 & EOG_k^{E_A} \geq EOG_k^{E_B} \\ 0 & otherwise \end{cases} \tag{6}$$

where "1" in H indicates the pixel at position (i, j) in the source image I_A is in focus, while "0" in H indicates the pixel at position (i, j) in the source image I_B is in focus.

Step 7: Construct the fused image. A consistency verification process [16] is subsequently applied to refine the decision matrix by using a majority filter, where a pixel is more likely to belong to the label which many of its neighbors also belong. A majority filter [17] is performed on the decision matrix W to obtain a refined decision matrix. Thus, the final fused image F is constructed according to fusion rule as follows:

$$F(i, j) = \begin{cases} I_A(i, j) & H(i, j) = 1 \\ I_B(i, j) & H(i, j) = 0 \end{cases} \tag{7}$$

where $I_A(i, j)$ and $I_B(i, j)$ are the values of the pixels at position (i, j) in the source images I_A and I_B, respectively. In this paper, the majority filter size is 3×3.

4. Experimental Results

In order to evaluate the validity of the proposed method, the experiment is performed on a set of 2 pairs of multi-focus source images [18] differing in content and texture, as is shown in Figure 4. The two pairs are all grayscale images with size of 640×480 pixels. In this paper, all the source images are assumed to have been registered. Experiments are conducted with Matlab in Windows environment on a computer with Intel Xeon X5570 and 48G memory. For comparison, beside the proposed method, some existing multi-focus image fusion methods are also implemented on the same set of source images. These existing methods include discrete wavelet transform (DWT), nonsubsumpled contourlet transform (NSCT), SF, principal component analysis (PCA), RPCA (Wan's method [12]). Due to the lack of original source code, the Eduardo Fernandez Canga's Matlab image fusion toolbox [19] is used as the reference for DWT, SF and PCA. Specifically, the Daubechies wavelet function 'bi97' is used in the DWT. The decomposition level of DWT is 4. The NSCT toolbox [20] is used as the reference for NSCT. The RPCA toolbox [21] is used as the reference for RPCA decomposition. The pyramid filter '9-7' and the orientation filter '7-9' with {4, 4, 3} levels of decomposition are set for the fusion method based on NSCT. The sliding window size in RPCA is 35×35. In order to quantitatively compare the performance

of proposed method and that of the others mentioned above, three metrics are used to evaluate the fusion performance. They are: (i) Structural similarity (SSIM), which reveals the degree of structural similarity between two images in luminance, contrast and structure [22, 23]. (ii) Mutual information (MI) [24], which measures the degree of dependence of the source image and the fused image. (iii) $Q^{AB/F}$ [25], which reflects the amount of edge information transferred from the source images to the fused image. In these methods, a larger value signifies a better fusion result.

Figure 4. Multi-focus Source Images. (a) Near Focused Image 'Disk', (b) far Focused Image 'Disk', (c) Near Focused Image 'Lab', (d) Far Focused Image 'Lab'

4.1. Qualitative Analysis

For qualitative comparison, the fused images 'Disk' and 'Lab' obtained by different methods are shown in Figures 5 (a-f) and Figures 6 (a-f). The difference images between the far focused source image 'Lab' and corresponding fused images obtained by different methods are shown in Figures 7 (a-f).

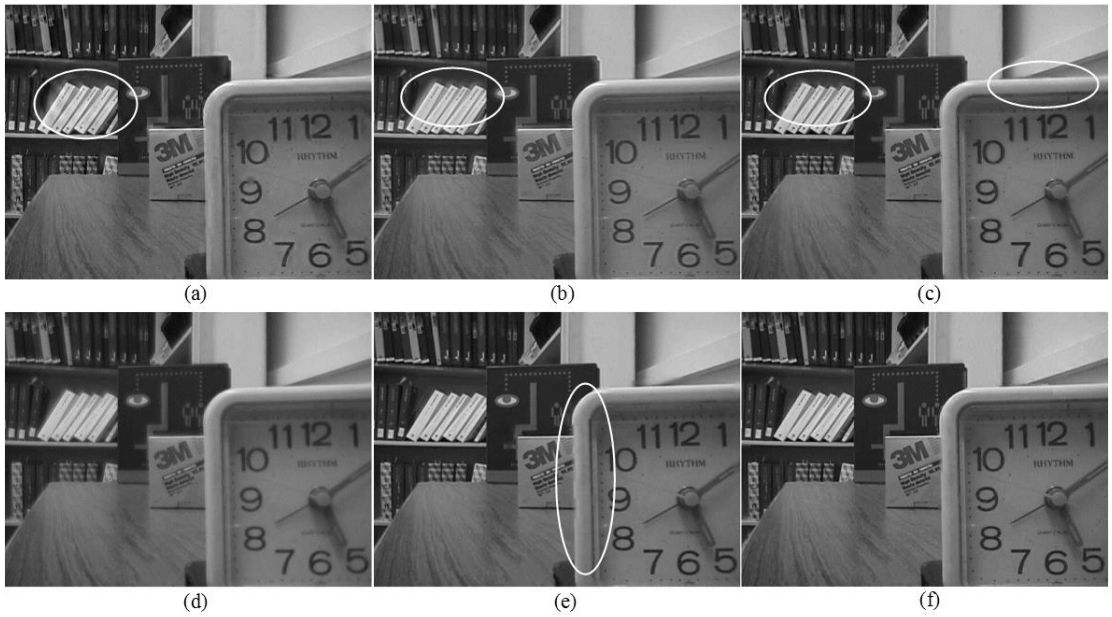

Figure 5. Fused Images 'Disk' Obtained by DWT (a), NSCT (b), SF (c), PCA (d), RPCA (e) and the Proposed Method (f)

The fused images obtained by the other fusion methods demonstrate obvious blur, such as the white books in Figures 5 (a-b) and the upper edge of the student's head in Figures 6 (a-b).

In addition, the obvious blur also appears in Figures 7 (a-b). The obvious blocking artifacts appear in the fused image obtained by SF, such as the upper edge of the clock in Figure 5(c) and the upper edge of the student's head in Figure 6 (c). The blocking artifacts also appear in Figures 7 (c). The low contrast appears in the fused image obtained by PCA, such as Figures 5 (d) and 6 (d). The fused images obtained by RPCA demonstrate incomplete or surplus, such as the incomplete edge of the clock in Figure 5 (e), and the narrow prominent of the student's head in Figure 6 (e). The obvious residual can be seen in Figure 7 (e). It is easy to see that the visual quality of the fused images obtained by the proposed image is better than that of the fused images obtained by the other methods.

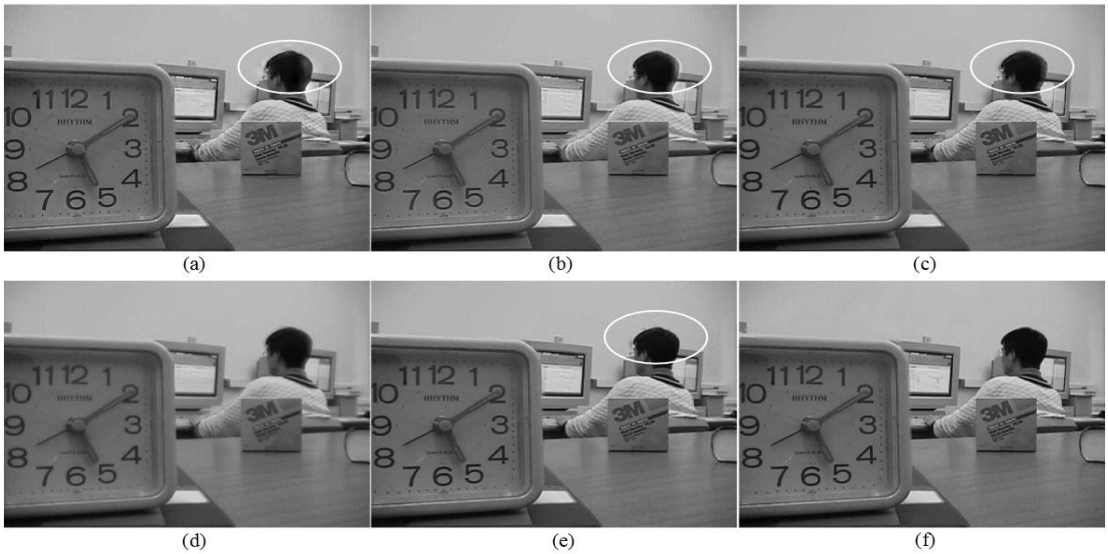

Figure 6. Fused Images 'Lab' Obtained by DWT (a), NSCT (b), SF (c), PCA (d), RPCA (e) and the Proposed Method (f)

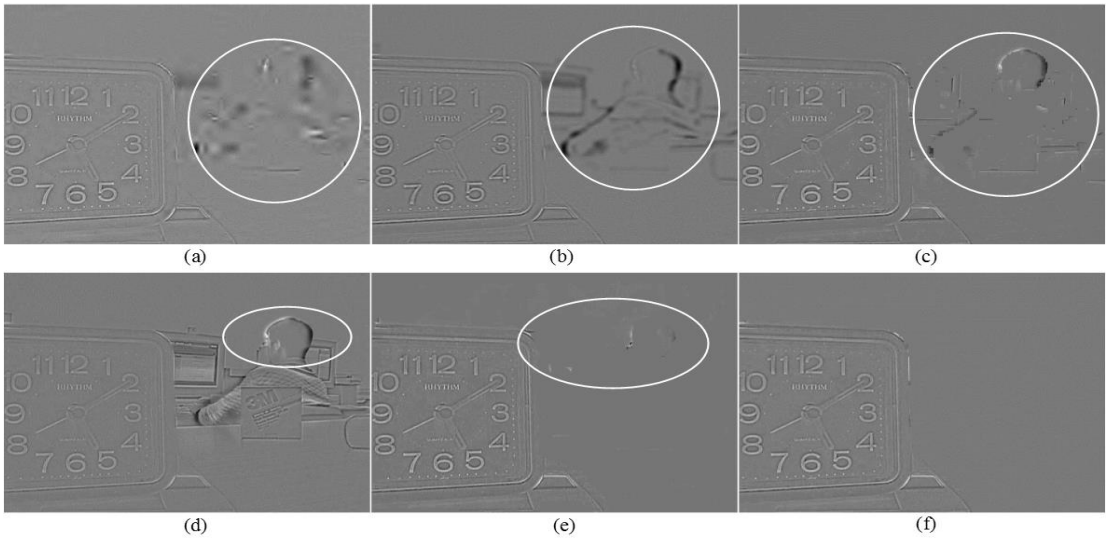

Figure 7. Difference Images between Source Image 'Lab' Far Focused and Corresponding Fused Images obtained by DWT (a), NSCT (b), SF (c), PCA (d), RPCA (e) and the Proposed Method (f)

4.2. Quantitative Analysis

For quantitative comparison, the quantitative results in three quality measures are shown in Table 1. The proposed method gains higher MI and $Q^{AB/F}$ values than the other methods. The running times are also shown in the Table. The fusion method based on PCA is rated highest and the proposed method is rated second when using SSIM measure. One can see that the running time of RPCA-based method is about five times larger than that of the proposed method. Due to the sliding window technique is applied for the detection of focused regions, the computational time of standard deviation of each sliding window in the RPCA-based is longer than that of QT decomposition used in the proposed method. However, the presented method still yields larger computational cost than DWT-based method, SF-based method and PCA-based method, due to the matrix decomposition accounts for the majority of the computational load.

Table 1. Performance of Different Fusion Methods

Method	Disk				Lab			
	SSIM	**MI**	$Q^{AB/F}$	**Run-time(s)**	**SSIM**	**MI**	$Q^{AB/F}$	**Run-time(s)**
DWT	0.84	5.36	0.64	0.64	0.90	6.47	0.69	0.59
NSCT	0.86	5.88	0.67	463.20	0.91	6.95	0.71	468.51
SF	0.87	7.00	0.68	1.01	0.91	7.94	0.72	1.03
PCA	**0.91**	6.02	0.53	0.11	**0.94**	7.12	0.59	0.08
RPCA	0.85	8.12	0.72	60.38	0.91	8.50	0.75	60.80
Proposed	0.90	**8.30**	**0.75**	9.46	0.92	**8.66**	**0.78**	9.52

5. Conclusion and Future work

This paper proposes a novel multi-focus image fusion method using sparse features to improve the fusion quality. The performance evaluations have demonstrated that the proposed method can produce better fused image and significantly inhibit the blocking artifacts. However, the proposed method has high computational cost for the matrix decomposition. In the future, we will consider optimizing the proposed method to reduce the computational cost and extend the developed method to the fusion of medical images.

Acknowledgements

The work is supported by the National Key Technology Science and Technique Support Program (No. 2013BAH49F03), the National Nature Science Foundation of China (No. 61379010), the Key Technologies R&D Program of Henan Province (No. 132102210515), the Natural Science Basic Research Plan in Shaanxi Province of China (No. 2012JQ1012).

References

[1] H. Hariharan, "Extending Depth of Field via Multi-focus Fusion", PhD Thesis, the University of Tennessee, Knoxville, **(2011)**.
[2] S. Li, X. Kang, J. Hu and B. Yang, "Image matting for fusion of multi-focus images in dynamic scenes", Information Fusion, vol. 14, **(2013)**.
[3] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion", Pattern Recognition Letters, vol. 28, no. 9, **(2007)**.

[4] A. M. Eskicioglu, A. M. and P. S. Fisher, "Image quality measures and their performance", IEEE Trans. Communication, vol. 43, no. 12, **(1995)**.
[5] S. Li, J. Kwok and Y. Wang, "Multifocus image fusion using artificial neural networks", Pattern Recognition Letters, vol. 23, no. 8, **(2002)**.
[6] D. Fedorov, B. Sumengen and B. S. Manjunath, "Multi-focus imaging using local focus estimation and mosaicking", Image Processing, 2006 IEEE International Conference on Image Processing, Atlanta, Georgia, **(2006)** October 8-11.
[7] V. Aslantas and R. Kurban, "Fusion of multi-focus images using differential evolution algorithm", Expert System with Application, vol. 37, no. 12, **(2010)**.
[8] W. Wu, X. M. Yang, Y. P. J. Pang and G. Jeon, "A multifocus image fusion method by using hidden Markov model", Opt. Communication, vol. 287, **(2013)**.
[9] A. Goshtasby, "Fusion of multifocus images to maximize image information", Proceedings of SPIE Defense and Security Symposium, Orlando, Florida, **(2006)** April 17-21.
[10] I. De and B. Chanda, "Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure", Information Fusion, vol. 14, no. 2, **(2013)**.
[11] E. Candes, X. Li, Y. Ma and J. Wright, "Robust principal component analysis?", JACM, vol. 58, no. 3, **(2011)**.
[12] T. Wan, C. Zhub and Z. Qin, "Multifocus Image Fusion Based on Robust Principal Component Analysis", Pattern Recognition Letters, vol. 34, no. 9, **(2013)**.
[13] J. Wright, A. Ganesh, S. Rao and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization", Proceedings of Neural Information Processing System, Bangkok, Thailand, **(2009)**, December 1-5.
[14] Z. Lin, M. Chen and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices", UIUC Technical Report UILU-ENG-09-2215, **(2010)**.
[15] P. Jagadeesh, P. Nagabhushan and R. Pradeep Kumar, "A Novel Image Scrambling Technique Based On Information Entropy And Quad Tree Decomposition", International Journal of Computer Science Issues, vol. 10, no. 2-1, **(2013)**.
[16] Y. Zhang and L. Ge, "Efficient fusion scheme for multi-focus images by using blurring measure", Digital Signal Processing, vol. 19, **(2009)**.
[17] H. Li, B. S. Manjunath and S. K. Mitra, "Multisensor image fusion using the wavelet transform", Graphical Models and Image Processing, vol. 57, **(1995)**.
[18] http://www.ece.lehigh.edu/spcrl, online image database, Accessed **(2013)** April 17.
[19] http://www.imagefusion.org/, Image fusion toolbox, Accessed **(2013)** March 20.
[20] http://www.ifp.illinois.edu/minhdo/software/, NSCT toolbox, Accessed **(2013)** January 20.
[21] http://perception.csl.illinois.edu/matrix-rank/sample_code.html, RPCA toolbox, Accessed **(2013)** January 10.
[22] http://foulard.ece.cornell.edu/gaubatz/metrix_mux/metrix_mux_1.1.zip, Assessment Package, Accessed **(2013)** January 10.
[23] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing, vol. 13, no. 4, **(2004)**.
[24] D. J. C. MacKay, "Information theory, inference and learning algorithms", Cambridge university press, **(2003)**.
[25] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure", Electronics Letters, vol. 36, no. 4, **(2000)**.

Authors

Yongxin Zhang is currently pursuing a PhD degree at the School of Information Science and Technology, Northwest University, Xi'an, China. His research interests include image processing and pattern recognition.

Li Chen is a professor at the School of Information Science and Technology, Northwest University, Xi'an, China. Her research interests include intelligent information processing, data mining and pattern recognition.

Zhihua Zhao is currently pursuing a PhD degree at the School of Information Science and Technology, Northwest University, Xi'an, China. His research interests include image processing.

Jian Jia is an associate professor at the Department of Mathematics, Northwest University, Xi'an, China. His research interests include image restoration, multi-resolution representations and intelligent information processing.