Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Multi-focus Image Fusion Based on Muti-scheme

Li-xiong LIU
School of Computer Science and Technology
Beijing Institute of Technology
Beijing, China
lxliu@bit.edu.cn

Bin Liao
School of Electric and Electronic Engineering
North China Electric Power University
Beijing, China
nathan@ncepu.edu.cn

***Abstract*—Multi-focus fusion is an important technique to integrate focal information from a set of input registered images. In this effort, a novel image fusion algorithm based on multi-scheme is presented. Given an over-complete Gabor dictionary, through matching pursuit signal decomposition algorithm, each source image can be described by sparse combinations of these atoms. Then the coefficients of the fused images are constructed according to different fuse rules. The last experiment results show that our algorithm can achieve better fusion effect than traditional wavelet image fusion method or spatial frequency method, whatever from subjective visual effect or objective metric.**

***Keywords*—matching pursuit; image fusion; spatial frequency; gabor dictionary**

I. INTRODUCTION

Modern optic imaging systems carry several fundamental limitations, and one of them is the low depth of the field. Generally, certain objects at particular distances are focused while other objects are blurred unavoidably depending on their distances from the camera. In other words, it's impossible for us to gain a clear image of all objects in the same scene. Therefore, a need has arise for statistical and mathematical methods that are capable of capturing complementary information and merging it in an efficient way. Such kind of technique is named as multi-focus image fusion.

In recent years, many research achievements are made in the image fusion field. Simple techniques perform fusion directly pixel by pixel on source images (e.g., weighted average method), which often have side effects such as reduced contrast of the fused image. Other methods include pyramid decomposition method [1, 2] and wavelet image fusion method. Wavelet image fusion and its modified versions [3, 4, 5] are emerging in recent years, which have been proved to have better performance over pixel based fusion. But natural images are very complicated, which may contain any kinds of spatial frequencies features. Traditional methods based on orthogonal linear transforms are not suitable for the multiple components present in the natural image. For example, it is hopeless to analyze a mixture of curves and textures with wavelet decomposition, because each phenomenon needs its own appropriate basis. This deficit may affect the quality of the wavelet fused image in certain degree.

Therefore, how to extract the particular features more efficiently from source images and fuse them together becomes the key of image fusion. Over-complete representations with redundant dictionary came forth in 1990s. In 1993, Mallat and Zhang[6] originally proposed the idea of signal sparse decomposition with redundant dictionaries and first introduced the Matching Pursuit algorithm(MP) in the signal processing community. Since then, matching pursuit algorithm has been applied widely in the low bitrate video coding [7, 8], voice recognition and denoising etc. In this paper, we present a novel image fusion method based on multi-scheme. Through matching pursuit algorithm, the strong high frequency features from different source images are extracted and fused together. Because the MP algorithm is redundant and overlapped, blocking or rings artifacts can be efficiently avoided. The last experiments show that the new fusion method is effective and the fused image has better visual effect than that of wavelet fusion method.

II. MATCHING PURSUIT THEORY

Matching pursuit algorithm decomposes any signal f into a linear expansion of waveform selected from a redundant dictionary of functions called hereafter atom. Let $D = \{g_\gamma\}_{\gamma\in\Gamma}$ be the over-complete dictionary containing some predefined normalized functions, where $\Gamma = \{1, \cdots, N\}$ is the set of all indices, and N is the size of the dictionary. A given arbitrary signal f can be decomposed into

$$f = < f, g_{\gamma_0} > g_{\gamma_0} + Rf \tag{1}$$

where $g_{\gamma_0} \in D$ is chosen in such a way that the absolute inner product $\left|< f, g_{\gamma_0} >\right|$ is maximum. Therefore the norm of residual signal $\|Rf\|$ is minimized. The greedy procedure continues iteratively for the residual function Rf . The residual at iteration n+1 is computed using the following equation:

$$R^{n+1} f = R^n f - \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n} \tag{2}$$

where $R^n f$ is the residual at iteration n, and $g_{\gamma_0} \in D$ the function whose inner product with the residual $R^n f$ is at a maximum.

After M iterations, the decomposed signal can be represented in terms of the successively matched atoms as:

$$f = \sum_{n=0}^{m-1} < R^n f, g_{\gamma_n} > g_{\gamma_n} + R^m f \qquad (3)$$

The inner product between the residual at iteration m+1 and the atom at the iteration m is given by

$$\left\langle R^{m+1} f, g_{\gamma_m} \right\rangle = \left\langle R^m f - \left\langle R^m f, g_{\gamma_m} \right\rangle, g_{\gamma_m} \right\rangle = 0 \qquad (4)$$

which means the vectors are orthogonal to each other. The energy of the signal can be written as the sum of the different contributions:

$$\|f\|^2 = \sum_{n=0}^{m-1} \left| < R^n f, g_{\gamma_n} > \right|^2 + \left\| R^m f \right\|^2 \qquad (5)$$

III. IMAGE FUSION BASED ON MULTI-SCHEME

For each registered image from same scene, the contour and background information are always similar, which belong to low frequency components, except that the edges and texture information of specified objects. Therefore, we can use the definition method [9,10] to process the low frequency fusion directly, and the detail information will be captured and fused by MP. Considering the computation complexity of MP depends strongly on the size of the dictionary, the Gabor dictionary with limited number of functions after pruning is used for capturing the high frequency features.

A. *Fusion of low frequency domain*

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

Firstly, 5x5 Gaussian low-pass filtering is performed on the source images respectively and the results are denoted as $f_A^0(x,y)$ and $f_B^0(x,y)$. The fusion rule is combination of average and definition scheme [10]. The evaluating function of definition is the sum of Laplacian operator in the eight directions. Computing the pixel definition in low frequency domain, the neighborhood area around a pixel is generally taken into account, in order to avoid the block effect or the gray value varing maybe appearing in the border between clear and fuzzy areas. The clarity degree of the pixel (x,y) is computed as

$$S(x,y) = \sum_{i=-p}^{p} \sum_{j=-p}^{p} C(x+i, y+j) \qquad (6)$$

where the neighborhood area is represented as [-p, p] and $C(x+i, y+j)$ is computed according to the following equation:

$$C(i,j) = \frac{\partial^2 f(i,j)}{\partial x^2} + \frac{\partial^2 f(i,j)}{\partial y^2} \qquad (7)$$

The local definition within the neighborhood Ω of the center pixel (x_0, y_0) is defined as:

$$S_A(x_0, y_0) = \sum_{i=-p}^{p} \sum_{j=-p}^{p} C(x_0+i, y_0+j) \qquad (8)$$

Similarly, the local definition of the image B can be defined. The match is defined as the local normalized correlation within the neighborhood Ω:

$$M_{AB}(x_0, y_0) = \frac{2 \sum_{(i,j)\in\Omega} p(i,j) \cdot f_A^0(x,y) \cdot f_B^0(x,y)}{S_A(x_0,y_0) + S_B(x_0,y_0)} \qquad (9)$$

where the $p(i,j)$ is the value of weight which is inversely proportional to the distance from the center. The value of the matrix M_{AB} reflects the degree of the correlation between the low frequency of the image A and B. And the fusion function is defined as:

$$f_C^0(x,y) = w_A f_A^0(x,y) + w_B f_B^0(x,y) \qquad (10)$$

The weights of w_A and w_B are decided by the following equations:

$$\begin{cases} w_{\min} = 0, \;\; w_{\max} = 1 & M_{AB}(x,y) < T \\ w_{\min} = \frac{1}{2} - \frac{1}{2}\left(\frac{1 - M_{AB}(x,y)}{1-T}\right), \;\; w_{\max} = 1 & M_{AB}(x,y) \ge T \end{cases} \qquad (11)$$

where the threshold $T \in [0.5, 1]$ is usually set as 0.8. Then the large weight is assigned to the source image with good definition: If the $S_A > S_B$ then $w_A = w_{\max}$ and $w_B = w_{\min}$ else $w_A = w_{\min}$ and $w_B = w_{\max}$.

In this way, the background information of the source images can be efficiently fused selectively depending on the degree of their difference.

B. *Fusion of high frequency domain*

After the fusion of the low frequency component, the residual images between the source images and the fused image $f_C^0(x,y)$ are calculated:

$$R_A^0(x,y) = f_A(x,y) - f_C^0(x,y) \qquad (12)$$

$$R_B^0(x,y) = f_B(x,y) - f_C^0(x,y) \qquad (13)$$

The residual images contain plenty of edges and context information. Based on the well developed theory of the MP algorithm, a novel iterative fusion of the residual images is proposed. The residual images will be divided into 16x16 blocks. At each iteration, the maximal block energy metric [11] is used to decide the initial search position of MP algorithm which maybe in the image A or image B, and the best atom is to be selected in the neighborhood of the search position. Then the residual images are updated. The above procedure will be repeated until the energy of the residual image are low than certain threshold, which means no more important details left. The output of the MP algorithm is a list of 2D atoms. The detected set of atoms is used to reconstruct the fused residual image. The main steps of image fusion on MP are described as following:

1) Initialize the Gabor dictionary D, and set iteration index k=0. Let R_A^k and R_B^k represents the residual images at iteration k.

2) Divide the residual images into 16x16 blocks and calculate the block energy matrixes:

$$E_A(i,j)=\sum\left(R_A^k(x,y)\right)^2,\ E_B(i,j)=\sum\left(R_B^k(x,y)\right)^2 \tag{14}$$

where (i, j) denoted as block index.

3) Find out the block with max energy in each residual image and compare them. Set the search position to be the center of block with max energy.

4) Around the neighborhood Ω of position, search the best atom $g_{(\gamma_k,x,y)}$ with maximal inner product:

$$\left|<R^k, g_{(\gamma_k,x,y)}>\right| = \sup_{\gamma\in\Gamma,(m,n)\in\Omega}\left|<R^k, g_{\gamma,m,n}\right|> \tag{15}$$

5) Add the atom into the atom list Λ and update the residual images:

$$R_A^{k+1} = R_A^k - <R^k, g_{(\gamma_k,x,y)}> g_{(\gamma_k,x,y)} \tag{16}$$

$$R_B^{k+1} = R_B^k - <R^k, g_{(\gamma_k,x,y)}> g_{(\gamma_k,x,y)} \tag{17}$$

6) If $\min(\|R_A^k\|,\|R_B^k\|) > \varepsilon$ then return to step 2.

where ε denoted as a threshold.

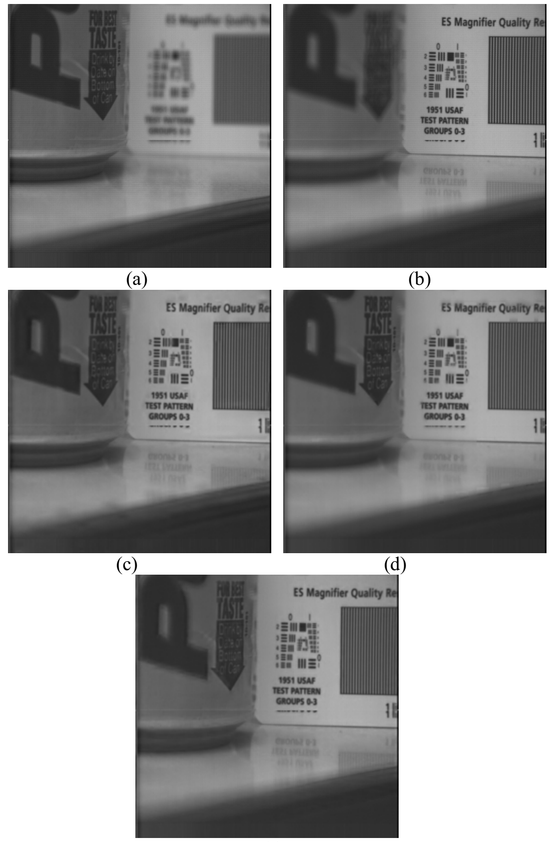
7) Reconstruct the residual image $R_C^0(x,y)$ using the atoms in the set Λ.

8) Merge the fused residual image with the low frequency part to get the fused image.

$$f_C(x,y) = f_C^0(x,y) + R_C^0(x,y) \tag{18}$$

IV. Experiment results

In order to verify the validity of proposed method, we have compared our results with traditional wavelet fusion and spatial frequency fusion method[12] on two groups of multi-focus images. The experimental results are showed in Figs.1-2. In each figure, the original multifocus images are given first, followed by the fused images obtained by wavelet fusion method (Bior. 9/7, 4 scales) and our proposed method. Further, a clear comparison of local details is also given.

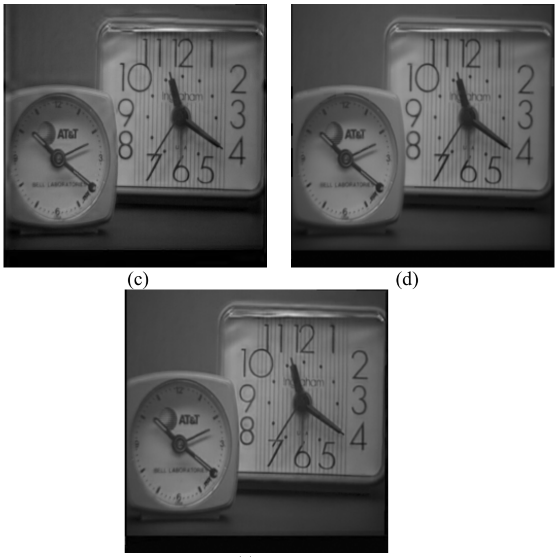


(a) (b) (c) (d) (e)

Figure.1 The 'Mulfocus' source images (512 × 512) and fusion results: (a) focus on the left. (b) focus on the right. (c) wavelet fusion result. (d) Spatial frequency fusion result. (e) mult-scheme fusion result.

Through careful inspection of Fig.1 and Fig.2, the results of proposed method are obviously better than that of wavelet fusion. However this is a subjective measure of quality and may not be universally acceptable. Considering the gradient or derivative operators are useful tools to detect the variation of intensity and the objects in focus always have high gradient values, a quantitative measure of gradient similarity measure is adopted to appreciate the quality of the fused image with results shown in Tab.I.

(a) (b)



(c) (d)

(e)

Figure.2 The 'Clock' source images (512 × 512) and fusion results: (a)focus on the left. (b) focus on the right. (c) wavelet fusion result. (d) Spatial frequency fusion result. (e) mult-scheme fusion result.

TABLE I. THE GRADIENT SIMILARITY MEASURE

	Clock	***Mulfocus***	***Lab***	***Disk***
Wavelet fusion	0.8294	0.8579	0.8816	0.8751
Spatial frequency fusion	0.8412	0.8699	0.8724	0.8484
Mult-scheme fusion	0.8490	0.8903	0.8935	0.8831

CONCLUSIONS

In this paper we have proposed a new multi-focus image fusion method based on multi-scheme. In this method, the source images are first decomposed into low frequency domain and high frequency domain. The definition scheme is used to fuse the low frequency information, and matching pursuit algorithm is exploited to fuse the high frequency information. Finally, the fused low frequency and high frequency information are merged to obtain the resultant fused image. Experimental results show that this method performs better in preserving the edge and texture information than that of wavelet fusion. At the same time the algorithm can be straightforwardly extended to handle more than two source images and has wide applications in medical diagnosis and remote sensing.

ACKNOWLEDGMENT

This paper was partially supported by the National Natural Science Foundation of China under grants No.60602050 and No.60805004.

REFERENCES

[1] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code". IEEE Transaction on Communication, 1983, 31(4):532-540.

[2] A. Toet, "Hierarchical image fusion". Machine Vision and Application, 1990, 3(2): 1-11.

[3] H. Li, B. Manjunath, S. Mitra, "Multisensor image fusion using the wavelet transform". Graph. Models Image Process. 1995, 57(3):235-245.

[4] J. Nunez, X. Otazu, O. Fors, et al. "Multiresolution-based Image Fusion with Additive-wavelet Decomposition". IEEE Trans actions on Geoscience and Remote Sensing, 1999,37(3):1204-1211.

[5] F.Guijiang, L. Yanjun, Z. Fang and C. Rui, "Image fusion algorithm based on wavelet transformation". Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation, 2006:2064-2068.

[6] S. Mallat and Z. Zhang, "Matching Pursuits with Time-Frequency Dictionaries". IEEE Transactions On Signal Processing, 1993,41(12):3397-3415

[7] L. Bin, X. Gang, W. Yuguo, "Adaptive image coding based on matching pursuit". Journal of Computer-aided Design & Computer Graphics, 2003, 15(9):1084-1090.

[8] L. Bin, X. Gang, W. Yuguo, "Multi-layered image representation and coding based on mixed transforms". Journal of China Institute of Communications, 2004, 25(6):120-125.

[9] A. M. Eskicioglu, P. S. Fisher, "Image Quantity Measures and Their Performance". IEEE Trans. on Conmmun, 1995,43(12):2959 -2965.

[10] H. A. Eltoukhy, S. Kavusi, "A Computationally Efficient Algorithm for Multi-Focus Image Reconstruction". SPIE, 2003, 5017:332-341.

[11] D. Ishita, C. Bhabatosh, "A simple and efficient algorithm for multifocus image fusion using morphological wavelets". Signal Processing, 2006:924-936.

[12] The Image Fusion Toolkit for Matlab developed by Eduardo Canga.http://www.imagefusion.org.