Fusion of Multifocus Images Based on the Nonsubsampled Contourlet Transform*

ZHANG Qiang, GUO Bao-long
(*ICIE Institute, School of Electromechanical Engineering, Xidian University, Xi'an* 710071, *China*)

Abstract: Focusing on the deficiencies of existing wavelet based algorithm, a novel algorithm for multifocus images fusion based on the nonsubsampled contourlet transform (NSCT) is proposed. And, the concepts of the local area visibility (LAVI) and the local oriented energy (LOE) are introduced in the contourlet domain. The selection principle of the low frequency subband coefficients based on the LAVI and the selection principle of the bandpass directional subband coefficients based on the LOE are presented respectively. The experimental results demonstrate that the proposed algorithm, compared to the methods based on the discrete wavelet transform and the discrete wavelet frame transform, can effectively reduce the loss of the useful information and the introduction of the artificial information. In addition, the proposed algorithm can extract more useful information from the source images, and make the fused image with higher performance in terms of both visual quality and objective evaluation criteria.

Key words: Image fusion; Nonsubsampled contourlet transform; Local area visibility; Local oriented energy

CLCN: TP391 **Document Code**: A **Article ID**: 1004-4213(2008)04-0838-6

0 Introduction

Optical lenses, particularly those with long focal lengths, suffer from the problem of limited depth of field. Consequently, the image obtained will not be in focus everywhere, i. e., if one object in the scene is in focus, another one will be out of focus. A possible way to alleviate this problem is by image fusion, in which several pictures with different focus points are combined to form a single image. This fused image will then hopefully contain all relevant objects in focus[1].

In recent years, various fusion methods based on multiscale transforms (MST) have been proposed and provide high performance. In previous research, the most commonly used MST for image fusion were the pyramid transform (PT)[2-3] and the discrete wavelet transform (DWT)[4-6]. In general, the DWT-based fusion schemes outperform the PT-based methods[1]. However, most of the DWT-based methods are mainly focusing on the combination of the high frequency subband coefficients, and just employed the simple 'average' scheme for the selection of the low frequency subband coefficients. But the low frequency subband coefficients contain most of the image energy and decide the contour of the image. And then such methods decrease the contrast of the fused image to some extent. In addition, the DWT has limitations such as limited orientations and non-optimal-sparse representation of images. Therefore, some artifacts are easily introduced in the fused images attained by the DWT-based methods, which will reduce the resultant image visual quality consequently [7].

Recently, CUNHA A L, ZHOU Jiang-ping and DO M N presented a new image geometric analysis tool, called the nonsubsampled contourlet transform (NSCT)[8]. The NSCT has such advantages as mulit-scale, good spatial and frequency localization and multi-direction, and can effectively capture the geometric information of images. Therefore, when the NSCT is introduced to image fusion, the characteristics of original images can be taken better and more information for fusion can be obtained. In addition, due to the elimination of the downsampler and upsampler during the decomposition and the reconstruction of the image, the NSCT has the shift-invariance property, which will effectively reduce the effects of the mis-registration on the fused results. Therefore, we apply the NSCT to image fusion and present a novel algorithm for fusion of multifocus images based on the NSCT. And the selection principles of the low frequency subband coefficients and the bandpass directional subband coefficients

* Supported by the National High Technology Research and Development Program of China (2006AA01Z127) and the National Natural Science Foundation of China (60572152)
Tel: 029-81143980 Email: zhangqiang134@163.com
Received date: 2006-09-18

are discussed in detail. The proposed algorithm has been applied to merge several sets of multi-focus images. And the experimental results show that the proposed algorithm is of great validity and of feasibility.

1 Nonsubsampled contourlet transform

The NSCT is a shift-invariant version of the contourlet transform [9]. The contourlet transform employs Laplacian pyramids (LP) for multiscale decomposition, and directional filter banks (DFB) for directional decomposition. To achieve shift-invariance, the NSCT is built upon nonsubsampled pyramids (NSP) and nonsubsampled directional filter banks (NSDFB).

The NSP is a two-channel nonsubsampled filter bank and has no downsampling or upsampling, and hence it is shift-invariant. The multiscale decomposition is achieved by iterating using the nonsubsampled filter banks. Such expansion is conceptually similar to the 1-D nonsubsampled wavelet transform computed with the a trous algorithm[10] and has $J+1$ redundancy, where J denotes the number of decomposition stages. For the next level, all filters are upsampled by 2 in both dimensions. The cascading of the analysis part is shown in Fig. 1. The equivalent

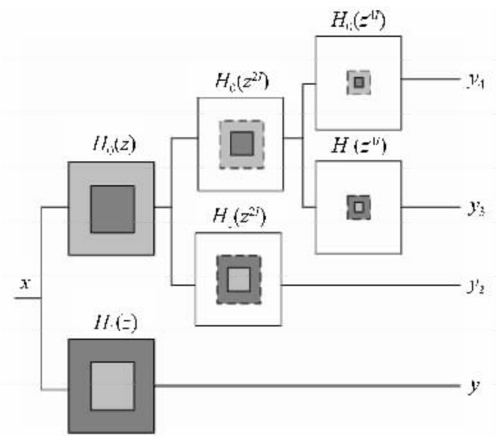

Fig. 1 Iteration of two-channel nonsubsampled filter banks in the analysis parts of the NSP

filters of a k-th level cascading NSP are given by

$$H_n^{\mathrm{eq}}(z)=\begin{cases} H_1(z^{2^{n-1}I})\prod_{j=0}^{n-2}H_0(z^{2^j I}) & 1\leqslant n\leqslant k \\ \prod_{j=0}^{n-2}H_0(z^{2^j I}) & n=k+1 \end{cases} \tag{1}$$

The NSDFB is a shift-invariant version of the critically sampled DFB in the contourlet transform. The building block of a NSDFB is also a two-channel nonsubsampled filter bank. To obtain finer directional decomposition, the NSDFBs are iterated. For the next level, all filters are upsampled by a quincunx matrix given by $D=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$. Fig. 2 illustrates a four - channel

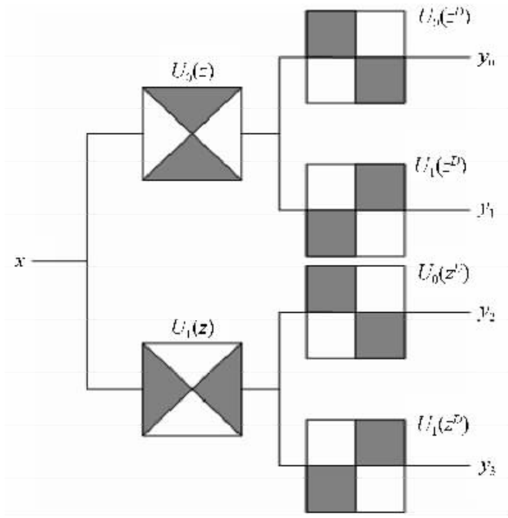

Fig. 2 A four-channel NSDFB constructed with two-channel fan filter banks

decomposition with two-channel fan filter banks. And the equivalent filter in each channel is given by

$$U_k^{\mathrm{eq}}(z)=U_i(z)U_j(z^D) \tag{2}$$

The NSCT is obtained by carefully combining the 2-D NSP and the NSDFB as shown in Fig. 3. If the building block two-channel nonsubsampled filter banks in the NSP and the NSDFB are invertible, then clearly the NSCT is invertible. The NSCT allows any number of 2^l directions in

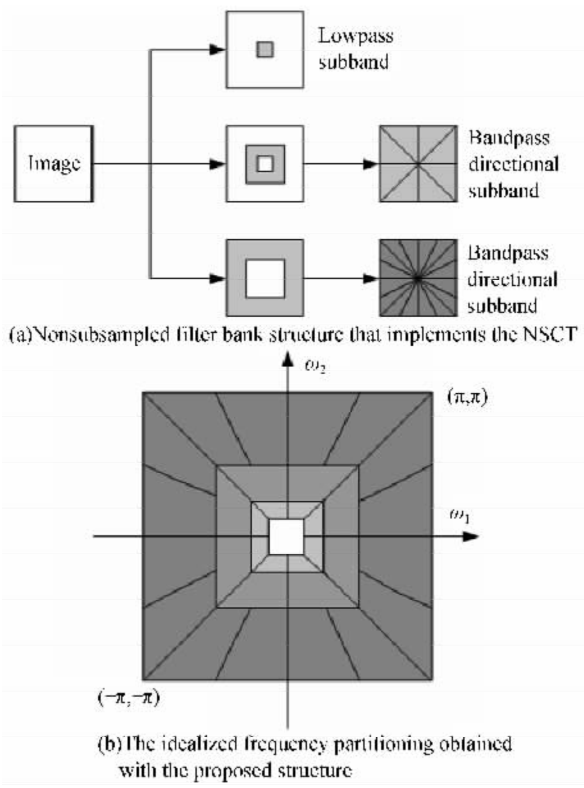

Fig. 3 The NSCT

each scale, and then the NSCT has redundancy given by $1+\sum_{j=1}^{J}2^{l_j}$, where l_j denotes the number of levels in the NSDFB at the j-th scale.

2 Image fusion algorithm

The NSCT has such properties as multi-scale, localization, multi-direction and shift-invariance. Therefore, we apply the NSCT to image fusion and present a novel algorithm for fusion of multi-sensor images based on the NSCT. This section provides an overview of the generalized fusion process. To simplify this discuss, we assume the fusion process is to generate a composite image F from a pair of source images denoted with A and B. Also we assume here that the source images are already perfectly registered. Then the proposed image fusion approach consists of the following steps.

1) Perform a J-level NSCT on source images A and B, and attain the NSCT coefficients $\{C_{j_0}^{A}(k_1,k_2), C_{j,l}^{A}(k_1,k_2)(j\geqslant j_0)\}$ and $\{C_{j_0}^{B}(k_1,k_2), C_{j,l}^{B}(k_1,k_2)(j\geqslant j_0)\}$, where, $C_{j_0}(k_1,k_2)$ is the low frequency subband coefficients and the $C_{j,l}(k_1,k_2)$ is the bandpass directional subband coefficients at the j-th scale and on the l-th direction.

2) Employ some fusion rules to attain the NSCT coefficients of the fused image $F\{C_{j_0}^{F}(k_1,k_2), C_{j,l}^{F}(k_1,k_2)(j\geqslant j_0)\}$.

3) Apply the inverse NSCT on the attained coefficients and thus obtain the fused image.

The low frequency subband coefficients decide the contour of the image, therefore, the selection principle of the low frequency subband coefficients is very important in the fusion algorithm and a proper principle can improve the quality of the fused image greatly. However, in many paper, the commonly used method for the combination of the low frequency subband coefficients is the 'average' scheme. This may reduce the image contrast greatly and then make some useful information of the source images lost.

In this paper, we introduce the concept of the image visibility (VI), which is inspired from the human visual system and is defined as[1,11]

$$\mathrm{VI}=\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{|F(m,n)-\mu|}{\mu^{\alpha+1}} \tag{3}$$

where μ is the mean intensity value of the image, and α is a visual constant ranging from 0.6 to 0.7. The image visibility represents the clarity of an image. In order to represent the local area clarity, we introduce the concept of local area visibility (LAVI) in the contourlet domain. When the local area mean $\bar{C}_{j_0}(k_1,k_2)\neq 0$, the LAVI is defined as

$$\mathrm{LAVI}(k_1,k_2)=\sum_{i=-(M_1-1)/2}^{(M_1-1)/2}\sum_{j=-(N_1-1)/2}^{(N_1-1)/2}\omega(i,j)\cdot\frac{|C_{j_0}(k_1+i,k_2+j)-\bar{C}_{j_0}(k_1,k_2)|}{\bar{C}_{j_0}(k_1,k_2)^{\alpha+1}} \tag{4}$$

where, the local area size $M_1\times N_1$ may be 3×3 or 5×5, $\tilde{\omega}(i,j)$ is the coefficient of the Gaussian template. And when $\bar{C}_{j_0}(k_1,k_2)=0$, the LAVI is also set to 0.

To select the low frequency subband coefficients of the fused image properly and restrain the effects of the noise, we propose a scheme based on the local area visibility combining with the 'average' scheme, and the scheme is defined as Eq. (5), in which the th is an experimental threshold according to the resultant image.

For the bandpass directional subband coefficients, the basic fusion rule is the absolute value maximum selection, i.e., the larger absolute values in the sub-bands are retained for reconstruction. There is the fact that the larger values correspond to the features in the image such as edges, lines, and regions boundaries. However, the NSCT coefficients of the image have great dependency. In order to select the bandpass directional subband coefficients more properly, we present a selection scheme based on the local oriented energy (LOE) in the contourlet domain, or the coefficients corresponding to larger LOE are included for reconstruction. This is based on the fact that the human visual system detects features at the points where LOE reaches maximum value[12]. The proposed scheme is given by Eq. 6, in which the LOE is defined by Eq. 7 and l_j denotes the number of direction decomposition levels in the NSCT at the j-th scale.

$$C_{j_0}^{F}(k_1,k_2)=\begin{cases} C_{j_0}^{A}(k_1,k_2) & \mathrm{LAVI}^{A}(k_1,k_2)-\mathrm{LAVI}^{B}(k_1,k_2)>\mathrm{th}\\ C_{j_0}^{A}(k_1,k_2)\times 0.5+C_{j_0}^{B}(k_1,k_2)\times 0.5 & |\mathrm{LAVI}^{A}(k_1,k_2)-\mathrm{LAVI}^{B}(k_1,k_2)|\leqslant\mathrm{th}\\ C_{j_0}^{B}(k_1,k_2) & \mathrm{LAVI}^{B}(k_1,k_2)-\mathrm{LAVI}^{A}(k_1,k_2)>\mathrm{th}\end{cases} \tag{5}$$

$C_{j,l}^{F}(k_1,k_2)=$

$$\begin{cases} C_{j,l}^{A}(k_1,k_2) & \mathrm{LOE}_j^A(k_1,k_2) \geqslant \mathrm{LOE}_j^B(k_1,k_2) \\ C_{j,l}^{B}(k_1,k_2) & \mathrm{LOE}_j^A(k_1,k_2) < \mathrm{LOE}_j^B(k_1,k_2) \end{cases} \quad (6)$$

$$\mathrm{LOE}_j(k_1,k_2) = \sum_{l=1}^{2^{l_j}} |C_{j,l}(k_1,k_2)|^2 \quad (7)$$

After determining all the NSCT coefficients of the fused image F, we can attain the fused image F by taking the inverse NSCT.

3 Experiments and analysis

A set of multifocus images have been used for evaluation of the image fusion algorithm presented in this paper. For comparison with other fusion methods, the fusion scheme is also performed based on the discrete wavelet transform (DWT-based method), the discrete wavelet frame transform (DWFT-based method) and the NSCT (NSCT-simple-based method), in which, the low frequency subband coefficients of the fused image are all simply attained by the 'averaging' scheme and the bandpass subband coefficients are all attained by the absolute-value-maximum-selection approach. In addition, the mean square error (Emse), difference coefficient (dDC)[13], the correlation coefficient (C) and the similarity based quality metric[14] are use to evaluate the performance of the fusion results quantitatively.

Smaller Emse, dDC values and a higher C value between the fused image and the reference image indicate that more information of the original images has been preserved during the fusion process and the fused image is closer to the reference image.

Re. [14] presents an improved structural similarity based quality metric for image fusion. The global similarity metric between the fused image F and the reference image R, is defined as

$$\mathrm{MSSIM}(R,F) = \frac{1}{|W|}\sum_{w\in W}\mathrm{SSIM}(R,F|w) \quad (8)$$

where, $\mathrm{SSIM}(R, F \mid w)$ denotes the structural similarity metric for the corresponding regions in the fused image and the reference image, or $\mathrm{SSIM}(R, F \mid w) = \frac{(2\overline{w}_R\overline{w}_F + C_1)}{(\overline{w}_R^2+\overline{w}_F^2+C_1)}\frac{(2\sigma_{w_R}\sigma_{w_F}+C_2)}{(\sigma_{w_R}^2+\sigma_{w_F}^2+C_2)} \cdot \frac{(\sigma_{w_Rw_F}+C_3)}{(\sigma_{w_R}\sigma_{w_F}+C_3)}$ and C_1, C_2, C_3 are small constants, with $C_3 = C_2/2$, w denotes the sliding window or the window under consideration, $\overline{w}$, σ_w^2 and $\sigma_{w_Rw_F}$ denote the mean, the variance and the covariance respectively, and W is the family of all sliding windows. In addition, the similarity metric between the fused image F and the input image A are defined as

$$\mathrm{MMS}(F,A) = \mathrm{MI}(F,A) \times \mathrm{MSSIM}(F,A) \quad (9)$$

where $\mathrm{MI}(F,A)$ denotes the mutual information between the image F and the image A. The higher the $\mathrm{MSSIM}(R,F)$ value, the higher quality of the fused image is. And higher $\mathrm{MMS}(F,A)$ value demonstrates that more information has been extracted from the source image and injected into the fused image.

Fig. 4(a) and Fig. 4(b) illustrate a pair of test images with different focus and Fig. 4(c) shows a reference image in which the camera was focused on each object. Fig. 5 illustrates the fusion results by the four fusion methods above. Compared with the source images, all of the four fusion results are satisfied. All the fusion methods can eliminate the effects resulting from the different focus of the camera, and can make all the objects in the fused images clear. However, by a more careful comparison, we can find that the fused image attained by the proposed method in this paper performs the best.

(a) Focus on the left part　(b) Focus on the right part　(c) The referenced image

Fig. 4　Source images

To explain this clearer, we extract parts of the fused results from Fig. 5(a)～(d) and put them into Fig. 6. In Fig. 6(a), we can find that many 'artifacts' are introduced. On the contrary, the artifacts are greatly reduced in Fig. 6(b) and Fig. 6(c), especially in Fig. 6 (c), which owes to the perfect properties of the NSCT. In Fig. 6 (d), almost all the useful information of the source images are preserved and few artifacts are introduced, and then the fused image has the best visual quality, which owes to the NSCT as well as the proper coefficient selection principles.

Table. 1 gives the quantitative results. According to the data in Table I, we can see that the methods based on the NSCT, especially the proposed method in this paper, significantly outperform the DWT-based method and the DWFT-based method. Our proposed method has the largest MSSIM, MMS and C values and the smallest Emse and dDC values, which means that our proposed fusion method can extract more information from the source images and preserve

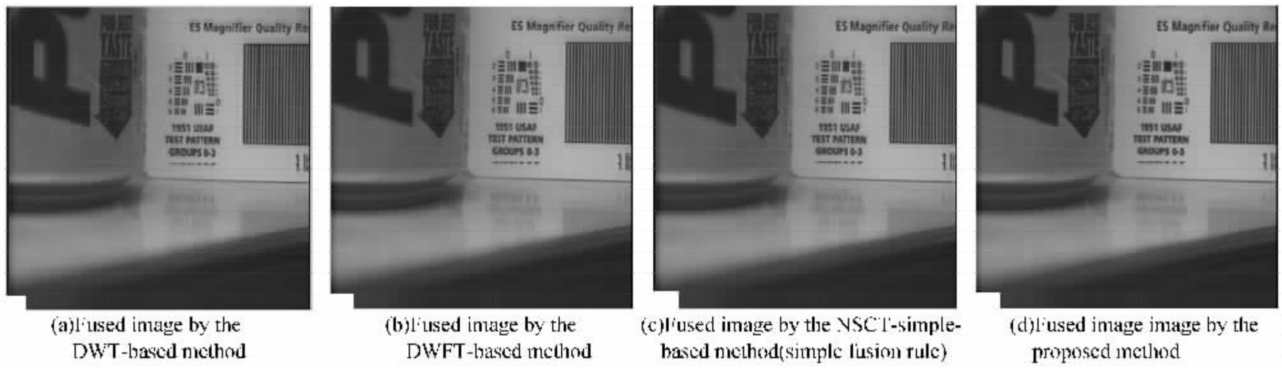

Fig. 5 Fused results

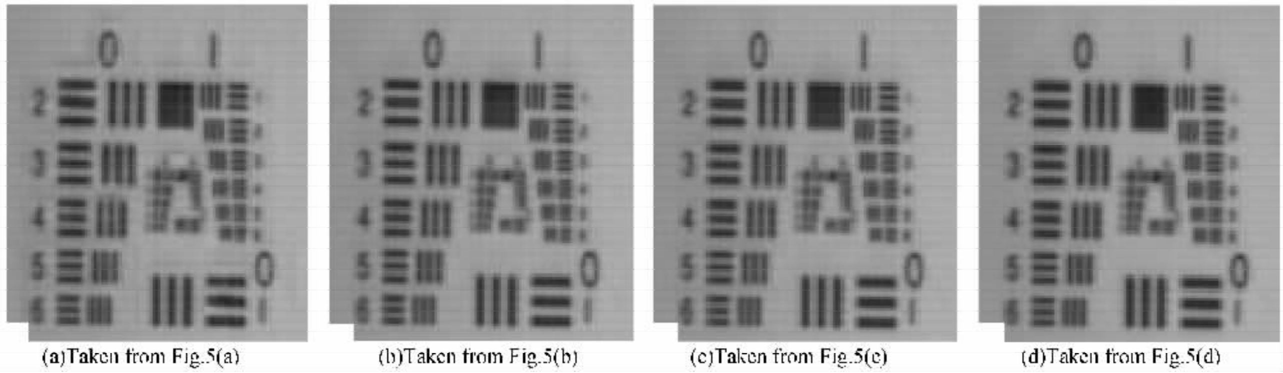

Fig. 6 Locally zoomed parts of the fused images

Table. 1 Performance of different fusion methods

Method	Emse	dDC	C	MSSIM	MMS(F,I_L)	MMS(F,I_R)	$\overline{\text{MMS}}$
DWT-based Method	6.1449	0.0349	0.9964	0.9316	3.3578	5.0746	4.2162
DWFT-based Method	5.4542	0.0331	0.9969	0.9368	3.4704	5.4122	4.4413
NSCT-simple-based method	5.0688	0.0307	0.9971	0.9471	3.7604	5.7941	4.7772
Proposed method	3.2854	0.0272	0.9980	0.9543	4.1270	6.0371	5.0820

them into the fused image. In addition, the fused image by the proposed method is closer to the reference image.

4 Conclusion

The multiscale decomposition and reconstruction tool of the image and the fusion rules are the two most important factors of the fusion algorithms based on the multiscale decomposition. The NSCT has many good properties such as multi-scale, perfect spatial and frequency localization, multi-direction and shift-invariance. Therefore the NSCT can present images better and is a more suitable multiscale analysis tool of the image in many cases such as image fusion. This paper proposes a novel image fusion scheme based on the NSCT and discusses the fusion rules thoroughly. And the proposed method has been used to merge several sets of multifocus images. The experimental results demonstrate that the proposed algorithm is of great validity and of feasibility.

References

[1] LI S T, KWOK J K, WANG Y N. Multifocus image fusion using artificial neural networks[J]. *Pattern Recognition Letters*, 2002, **23**: 985-997.

[2] BURT P J, ADELSON E. The laplacian pyramid as a compact image code[J]. *IEEE Trans Commum*, 1983, **31**(4): 532-540.

[3] TOET A. Image fusion by a ratio of low-pass pyramid[J]. *Pattern Recognition Letters*, 1989, **9**: 255-261.

[4] PAJARES G, DE LA CRUZ J M. A wavelet-based image fusion tutorial[J]. *Pattern Recognition*, 2004, **37**: 1855-1872.

[5] YAN Wen-hao, MA Cai-wen, ZHANG Ming, *et al*. A new way for image fusion based on wavelet transform[J]. *Acta Photonica Sinica*, 2006, **35**(4): 638-640.

[6] LI Hui-hui, GUO Lei, LIU Hang. Evaluation of multisensor image fusion using different wavelet transform[J]. *Acta Photonica Sinica*, 2006, **35**(8):1263-1267.

[7] ZHANG Qiang, GUO Bao-long. Fusion of multisensor images based on the curvelet transform[J]. *Journal of*

Optoelectronics · Laser, 2006, **17**(9): 1123-1127.

[8] CUNHA A L, ZHOU Jiang-ping, DO M N. The nonsubsampled contourlet transform: theory, design, and applications[J]. *IEEE Trans on Image Processing*, 2006, **15**(10): 3089-3101.

[9] DO M N, VETTERLI M. The contourlet transform: An efficient directional multiresolution image representation[J]. *IEEE Trans on Image Processing*, 2005, **14**(12): 2091-2106.

[10] SHENSA M J. The discrete wavelet transform: Wedding the a trous and Mallat algorithms[J]. *IEEE Trans on Signal Processing*, 1992, **40**(10): 2464-2482.

[11] LI Zhen-hua, JING Zhong-liang, SUN Shao-yuan. Pixel-clarity-based multifocus image fusion[J]. *Chinese Optics Letters*, 2004, **2**(2): 82-85.

[12] LIU Z, TSUKADA K, HANASAKI K, *et al*. Image fusion by using steerable pyramid[J]. *Pattern Recognition Letters*, 2001, **22**: 29-939.

[13] LI Hui-hui, GUO Lei, LIU Hang. Research on image fusion based on the second generation curvelet transform[J]. *Acta Optica Sinica*, 2006, **26**(5): 657-662.

[14] DI Hong-wei, LIU Xian-feng. Image fusion quality assessment based on structural similarity[J]. *Acta Photonica Sinica*, 2006, **35**(5): 766-771.

一种基于非下采样 Contourlet 变换多聚焦图像融合算法

张强,郭宝龙

(西安电子科技大学 机电工程学院 智能控制与图像工程研究所,西安 710071)

收稿日期:2006-09-18

摘　要:针对现有小波类图像融合算法的不足,提出了一种基于非下采样 Contourlet 变换多聚焦图像融合算法,并在 Contourlet 域中引入了局部区域可见度以及局部方向能量的概念.针对低频子带系数和各带通方向子带系数分别提出了基于局部区域可见度以及基于局部方向能量的系数选择方案.通过对多聚焦图像融合的仿真实验,表明该算法相对于传统的基于离散小波变换和离散小波框架变换融合算法能够有效减少有用信息的丢失以及虚假信息的引入,同时能够从源图像中提取更多的有用信息并注入到融合图像中,得到更好视觉效果和更优量化指标的融合图像.

关键词:图像融合;非下采样 Contourlet 变换;局部区域可见度;局部方向能量

ZHANG Qiang　was born in 1979. Now he is a lecturer at School of Electromechanical Engineering, Xidian University and is pursuing the Ph. D. degree in circuits and systems at Xidian University. His research interests focus on image multiscale geometric analysis and image fusion.