Research Article

Airborne Infrared and Visible Image Fusion for Target Perception Based on Target Region Segmentation and Discrete Wavelet Transform

Yifeng Niu,[1,2] Shengtao Xu,[3] Lizhen Wu,[1] and Weidong Hu[2]

[1] College of Mechatronics and Automation, National University of Defense Technology,
   Changsha 410073, China
[2] ATR State Key Laboratory, National University of Defense Technology, Changsha 410073, China
[3] Jiuquan Satellite Launch Centre, Lanzhou 732750, China

Correspondence should be addressed to Yifeng Niu, niuyifeng@nudt.edu.cn

Infrared and visible image fusion is an important precondition of realizing target perception for unmanned aerial vehicles (UAVs), then UAV can perform various given missions. Information of texture and color in visible images are abundant, while target information in infrared images is more outstanding. The conventional fusion methods are mostly based on region segmentation; as a result, the fused image for target recognition could not be actually acquired. In this paper, a novel fusion method of airborne infrared and visible image based on target region segmentation and discrete wavelet transform (DWT) is proposed, which can gain more target information and preserve more background information. The fusion experiments are done on condition that the target is unmoving and observable both in visible and infrared images, targets are moving and observable both in visible and infrared images, and the target is observable only in an infrared image. Experimental results show that the proposed method can generate better fused image for airborne target perception.

1. Introduction

Unmanned aerial vehicles (UAVs) are aircrafts which have the capability of flight without an onboard pilot. UAV can be remotely controlled, semiautonomous, autonomous, or have a combination of these capabilities. UAV can execute given missions in a whole lot of domains [1]. In order to complete these various missions, UAV firstly needs to be equipped with sensor payloads to acquire images of the mission area and realize environment perception. The sensor payloads include infrared sensor, and visible light sensor. Infrared and visible image fusion from the same scene is the basis of target detection and recognition for UAV.

However, the images shot by airborne sensors are dynamic, which increases more difficulties for visible and infrared image fusion. In order to acquire the situation assessment, it is very important to extract the target information. In fact, information of texture and color in visible images are very abundant, while the target information, especially an artificial target, is more outstanding in infrared images. According to this, we can divide the image regions based on target regions, which can utilize the target information more effectively.

Image fusion can be performed at four levels of the information representation, which are signal, pixel, feature, and symbolic levels [2]. An infrared and visible image fusion method based on region segmentation is proposed in [3, 4] and adopted the nonsubsampled contourlet transform (NSCT) to fuse the regions given and optimize the quality of the fused image, while the method did not capture the target information effectively. Yin et al. proposed an infrared and visible image fusion method based on color contrast enhancement which can be propitious to target detection and improve the observer performance [5]. Hong et al. presented a fusion framework for infrared and visible images based on data assimilation and genetic algorithm (GA) [6]. Shao et al. introduced fast discrete curvelet transform (FDCT) and focus measure operators to realize the fusion of infrared and visible images [7]. Other fusion methods for the infrared and color visible images have been compared in [8–10], in which the fusion method based on discrete wavelet transform (DWT) performed well. Though NSCT, FDCT, and other novel transforms are superior to DWT in some performances, the transform should not be the key and core problem for infrared and visible image fusion [11]. In this paper, we only use DWT as a meaning to research the fusion concerning target detection and perception, for DWT has less computational complexity [12]. Yao et al. proposed a new approach to airborne visible and infrared image fusion [13, 14] and researched target fusion detection [15], which used the relevant information in different frames that could meet the demand for low real-time fusion.

Dynamic image fusion has its own characteristics, which require that the fusion method is consentaneous and robust to both time and space [16–18]. In order to utilize the different region features and get more effective target and background information, a method of visible and infrared image fusion in DWT domain based on dynamic target detection and target region segmentation is proposed. First, the image segmentation is done based on the detected candidate target regions, and then the information between the frames can be used to attain the stability and consensus to time. Finally, the different fusion rules are designed according to the characteristics of target regions to complete visible and infrared image fusion.

2. Image Segmentation Based on Target Regions

According to the requirements of airborne image processing for the accuracy and the speed, the moving target in the airborne images can be detected using the frame difference method based on background motion compensation [19]. The algorithm flow of target region detection is shown in Figure 1.

2.1. Motion Information Extraction Based on Frame Difference

On the basis of motion compensation for the image background, target detection can be done by applying the frame difference method to the image sequences. The regions whose pixel values are constant can be regarded as background regions, while the regions whose pixel
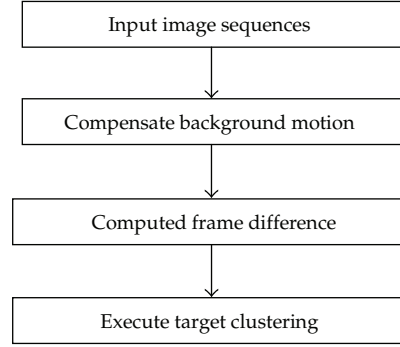
Figure 1: Flow of target region detection.

values change with different points are moving target regions, which include the motion information of targets. Using the converse transform parameters, we can make motion compensation for frame $n + 1$ and compute the difference between the former frame and the current.

$$D_n(x, y) = \left| I_n(x, y) - I_{n+1}(x, y) \right|, \tag{2.1}$$

where $D_n(x, y)$ denotes the frame difference of frame $n+1$ at point (x, y) and $I_n(x, y)$ denotes the pixel value of frame n at point (x, y). The change of pixel values in unmoving regions should be zero; however, because of some random noises, luminance change, and weather change, the differences are fairly small and belong to the salt and pepper noise. In order to extract the moving target regions, we need to select a proper threshold to segment the source images. Thus, we can acquire the moving target regions and relevant motion information. This method is easy to realize and robust to illumination, which could be a proper method for airborne image processing.

2.2. Target Clustering and Image Segmentation

After getting motion information from the neighbor frames, each target region has not been differentiated, and some unpredicted noises may appear in the target regions. In order to distinguish the different moving regions and filter noisy points in the target regions, the target clustering algorithm is proposed, which is shown in Figure 2.

The criterion for judging whether the points belong to a certain cluster is distance. According to the transcendental knowledge, the distances among different clusters can be established. The threshold can be set beyond which a new cluster can be added. When each target is distinguished, the number and the range of points in each cluster need to be computed. The points whose quantity is less than a certain number can be regarded as noisy points or false points and then can be eliminated. For the confirmed target regions, the region size and boundary can be computed, and then each target region can be marked out in the source images.

Suppose the visible image is E_1 and the infrared image is I_1, each image is to be divided into two parts at least, target region of R and background region of B, shown in Figure 3. There exist several target regions in source images. As shown in Figure 3, two source images
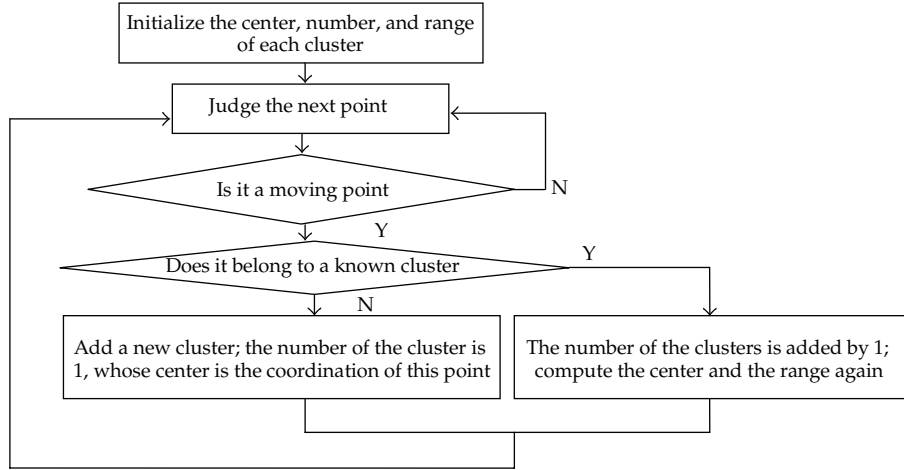
Figure 2: Flow of target clustering.

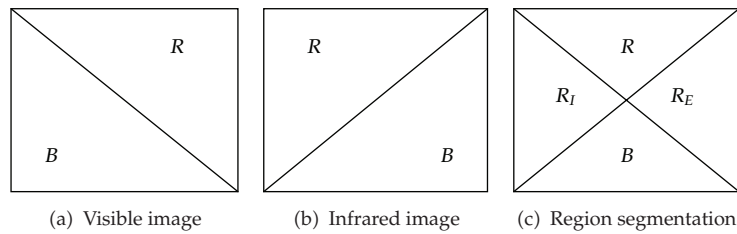(a) Visible image   (b) Infrared image   (c) Region segmentation

Figure 3: Illustration of region segmentation based on target regions.

are superposed, and four regions come into being, including common target region of R, visible target region of R_E, infrared target region of R_I, and common background region of B. For different regions, the different fusion strategies and methods are adapted to gain better fusion effect.

3. Visible and Infrared Image Fusion Based on Target Regions in DWT Domain

First, input the registered source visible image E_1 and infrared image I_1. Find the DWT of each E_1 and I_1 to a specified number of decomposition levels of N. Suppose (x, y) denotes the coordination of any pixel point, $g_i^E(x, y)$ and $f_N^E(x, y)$ denote high-frequency and low-frequency subband coefficients of the visible image, respectively, $g_i^I(x, y)$ and $f_N^I(x, y)$ denote the high-frequency and low-frequency sub band coefficients of the infrared image, respectively, F denotes the fused image, and $g_i^F(x, y)$ and $f_N^F(x, y)$ denote the high-frequency and low-frequency sub band coefficients of the fused image. The fusion rules of different regions are as follow.

Step 1. For the target region that only exists in visible image or infrared image, the high-frequency and low-frequency subband coefficients of F can be computed as follow:

$$g_i^F(x,y) = \begin{cases} g_i^E(x,y) & \text{if } (x,y) \in R_E, \\ g_i^I(x,y) & \text{if } (x,y) \in R_I, \end{cases}$$

$$f_i^F(x,y) = \begin{cases} f_N^E(x,y) & \text{if } (x,y) \in R_E, \\ f_N^I(x,y) & \text{if } (x,y) \in R_I. \end{cases} \tag{3.1}$$

Step 2. For the common target region R, according to similarity measurement of two images, we use the selective or the weighted average rule to fuse the source images. The similarity of the two images (M_{EI}) in region R is defined as

$$M_{EI}(R) = \frac{2 \sum_{(x,y)\in R} I_E(x,y) \cdot I_I(x,y)}{\sum_{(x,y)\in R} [I_E(x,y)]^2 + \sum_{(x,y)\in R} [I_I(x,y)]^2}. \tag{3.2}$$

Then compute the energy of all high-frequency sub band coefficients in region R of the two images, and use it as the fusion measure

$$S^P(R) = \sum_{(x,y)\in R} \sum_i^N g_i^P(x,y)^2, \quad P = E, I. \tag{3.3}$$

If $M_{EI}(R) < \alpha$, where α is a similarity threshold, such as 0.8, then apply the selective fusion rule to the two image

$$g_i^F(x,y) = \begin{cases} g_i^E(x,y), & \text{if } S^E(R) \geq S^I(R), \\ g_i^I(x,y), & \text{if } S^E(R) < S^I(R), \end{cases}$$

$$f_N^F(x,y) = \begin{cases} f_N^E(x,y), & \text{if } S^E(R) \geq S^I(R), \\ f_N^I(x,y), & \text{if } S^E(R) < S^I(R). \end{cases} \tag{3.4}$$

If $M_{EI}(R) \geq \alpha$, then apply the weighted average fusion rule to the two images and define the weighted coefficients

$$\omega_{\min} = \frac{1}{2}\left(1 - \frac{1 - M_{EI}(R)}{1 - \alpha}\right),$$

$$\omega_{\max} = 1 - \omega_{\min}. \tag{3.5}$$

Then fusion process in the common target region R can be described as follows:

$$g_i^F(x,y) = \begin{cases} \omega_{\max}g_i^E(x,y) + \omega_{\min}g_i^I(x,y), & \text{if } S^E(R) \geq S^I(R), \\ \omega_{\max}g_i^I(x,y) + \omega_{\min}g_i^E(x,y), & \text{if } S^E(R) < S^I(R), \end{cases}$$

$$f_N^F(x,y) = \begin{cases} \omega_{\max}f_N^E(x,y) + \omega_{\min}f_N^I(x,y), & \text{if } S^E(R) \geq S^I(R), \\ \omega_{\max}f_N^I(x,y) + \omega_{\min}f_N^E(x,y), & \text{if } S^E(R) < S^I(R). \end{cases} \tag{3.6}$$

Step 3. For the common background region B, take the different fusion strategies for the high-frequency and low-frequency sub band, respectively. For low-frequency sub band, we use the average method for fusion directly

$$f_N^F(x,y) = 0.5f_N^E(x,y) + 0.5f_N^I(x,y). \tag{3.7}$$

For high-frequency sub band, we adopt the fusion rule based on windows proposed by Burt and Kolczynski [20]. First, compute the energy of high-frequency sub band coefficients in local window, and use it as the fusion measure

$$S^P(x,y) = \sum_{(m,n)\in N(x,y)} \omega(m,n) \cdot g_i^P(m,n)^2, \quad P = E, I, \tag{3.8}$$

where $N(x,y)$ is the window whose center is (x,y), $\omega(m,n)$ is a weighted coefficient, and the sum is 1. According to the similarity fusion method in a local window, the similarity is defined as

$$M_i(x,y) = \frac{2 \cdot \sum_{(m,n)\in N(x,y)} \omega(m,n) \cdot |g_i^E(m,n)| \cdot |g_i^I(m,n)|}{S_i^P(x,y) + S_i^I(x,y)}. \tag{3.9}$$

If $M_i(x,y) < \alpha_i$, where α_i is the similarity threshold, then apply the selected fusion rule for the two image

$$g_i^F(x,y) = \begin{cases} g_i^E(x,y), & \text{if } S^E(x,y) \geq S^I(x,y), \\ g_i^I(x,y), & \text{if } S^E(x,y) < S^I(x,y). \end{cases} \tag{3.10}$$

If $M_i(x,y) \geq \alpha_i$, then apply the selected fusion rule to the two image, take the weighted average fusion rule for the two images, and define the weighted coefficients

$$\omega_{\min} = \frac{1}{2}\left(1 - \frac{1 - M_i(x,y)}{1 - \alpha_i}\right),$$

$$\omega_{\max} = 1 - \omega_{\min}. \tag{3.11}$$

Table 1: Evaluation metrics of different image fusion methods.

Fusion method	Entropy	Mutual information	Edge preservation
Weighted average	4.4058	0.9521	0.4152
Principal component analysis	6.1254	1.3214	0.7124
Laplacian pyramid	6.3547	1.4521	0.7457
Gradient pyramid	5.9687	1.3542	0.7012
Contrast pyramid	6.3142	1.4217	0.7315
Discrete wavelet transform	6.5214	1.4754	0.7462
Our fusion method	6.5321	1.5531	0.8213

The fusion process in the common background region B can be described as follows:

$$g_i^F(x,y) = \begin{cases} \omega_{\max}g_i^E(x,y) + \omega_{\min}g_i^I(x,y), & \text{if } S^E(x,y) \geq S^I(x,y), \\ \omega_{\max}g_i^I(x,y) + \omega_{\min}g_i^E(x,y), & \text{if } S^E(x,y) < S^I(x,y). \end{cases} \tag{3.12}$$

4. Results and Analysis of Fusion Experiments

4.1. Target Is Unmoving and Observable Both in Visible and Infrared Images

First, the experiment to fuse the visible and infrared images in which the target is unmoving and observable in both of the two images is conducted. There exists a static tank target in source images. We can get the small rectangle region including the tank-based target region segmentation, as shown in Figures 4(i) and 4(j), other part is the background region, thus, the source image can be divided into different parts, and different fusion rules are adopted to produce the fused image. As the source images are fused using different pyramid methods, the decomposition level is 3, while the fusion method based on discrete wavelet transform adopt the decomposition level of 2.

The evaluation metrics of different fusion method are given in Table 1. The definition of these metrics can be found in [21, 22]. It can be seen that our method not only preserves abundant texture information, but also image edges maintain well and the fusion effect is better than other methods. In fact, result of weighted average method is blurring in background, and target is not clear, which shows that the source information has not been utilized efficiently. The results of pyramid decomposition methods are abundant in background, but the target information has not been preserved. The basic wavelet method preserves both the details and the target information. However, all of above methods are based on one rule, and the whole image use the same fusion rule, which has not used different fusion rules according to different region characteristics and inevitably result in losing information. The result of our method which adopts different fusion rules according to the different regions shows that the background information is more abundant, and the target is more outstanding.

4.2. Target Is Moving and Observable Both in Visible and Infrared Images

Second, the experiment to fuse the visible and infrared images in which the target is moving and observable both in the two images is conducted (note: the source images are from [16]).
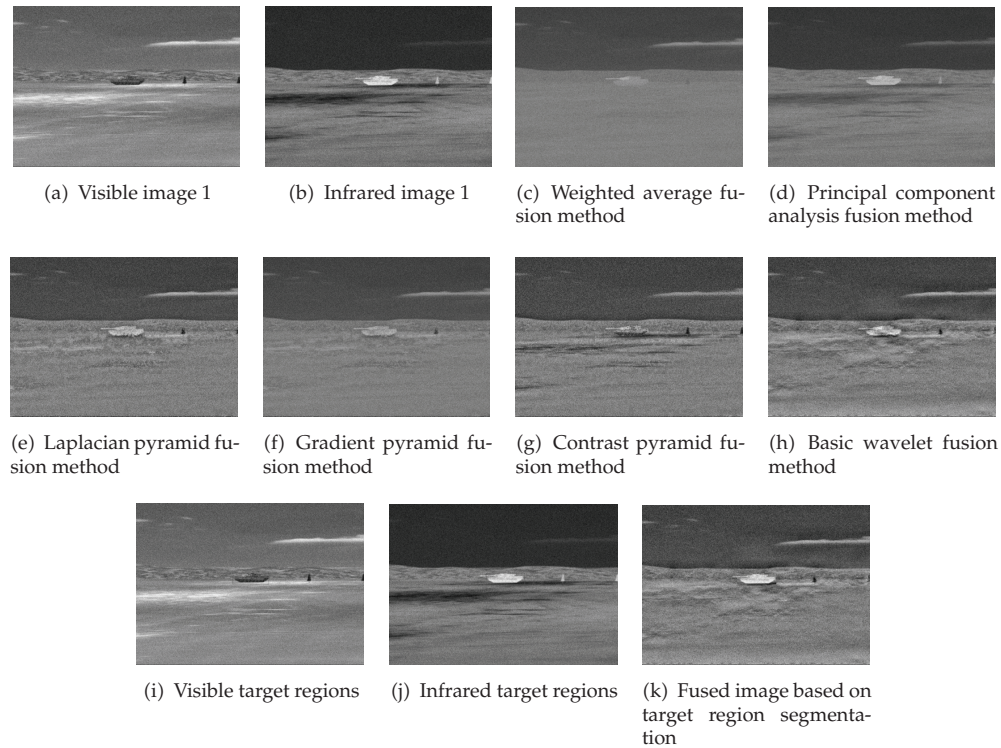
(a) Visible image 1   (b) Infrared image 1   (c) Weighted average fusion method   (d) Principal component analysis fusion method

(e) Laplacian pyramid fusion method   (f) Gradient pyramid fusion method   (g) Contrast pyramid fusion method   (h) Basic wavelet fusion method

(i) Visible target regions   (j) Infrared target regions   (k) Fused image based on target region segmentation

Figure 4: Results on condition that the target is observable both in visible and infrared image.

There exist several moving vehicle targets in source images which increases more difficulties for multiple source image fusion. Through target detection, the source images can be divided into four different regions, that are target R_1, R_2, R_3, and R_4 and the background, which be shown in Figures 5(c) and 5(d). We adopt different fusion rules according to the different regions and make a contrast research to the basic wavelet method. The results in Figures 5(e) and 5(f) show that the background information is more abundant, and the target is more outstanding than those of the basic wavelet method.

4.3. Target Is Observable Only in Infrared Image

In the actual environment, the target can be sheltered from other objects, which can be observed by the infrared sensor, such as Figures 6(a) and 6(b), in which the person target is unobservable in the visible image, while it is observable in the infrared image. From Figures 6(c) and 6(d), it can be seen that the target information changes using the basic wavelet method, which becomes dark inside and blur along the edge, while the target is clear using our method which means that the target information and background information have been mostly preserved and can meet demand of the target extraction.

5. Conclusions

In this paper, we proposed a new approach to infrared and visible image fusion based on detected target regions in DWT domain, which can help UAV to realize environment
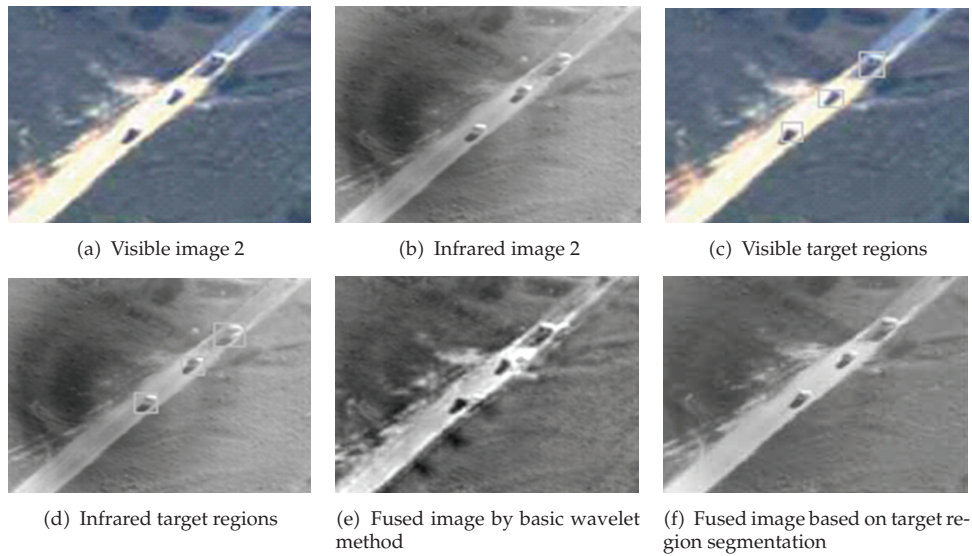
(a) Visible image 2

(b) Infrared image 2

(c) Visible target regions

(d) Infrared target regions

(e) Fused image by basic wavelet method

(f) Fused image based on target region segmentation

Figure 5: Results on condition that targets are moving both in visible and infrared image.

(a) Visible image 3

(b) Infrared image 3

(c) Fused image by basic wavelet method

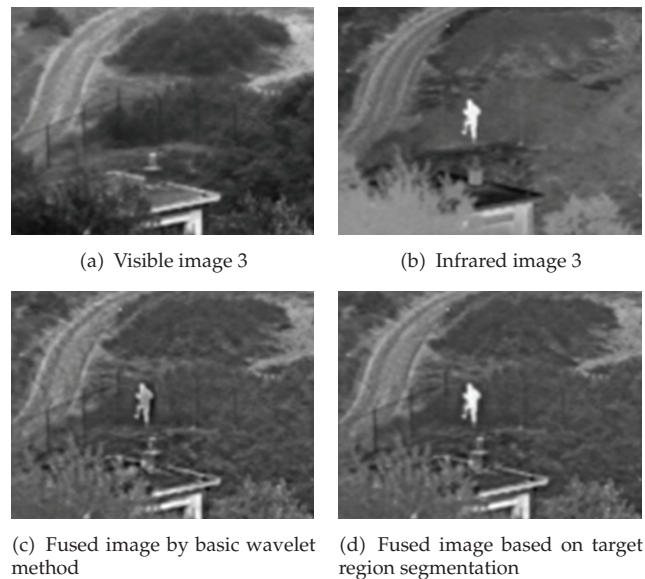(d) Fused image based on target region segmentation

Figure 6: Results on condition that the target is observable only in infrared image.

perception. Other than the conventional fusion methods based on region segmentation, we proposed a frame difference method for target detection, which can be used to segment the source images, and then we design different fusion rules based on target regions to fuse the visible and infrared images, which can gain more target information and preserve more background information in the source images. In the future, the method can be spread to other source image for fusion, and the time performance of our method can be improved based on GPU (Graphics Processing Unit) and other hardware.

Acknowledgment

References

[1] R. R. Pitre, X. R. Li, and R. Delbalzo, "UAV route planning for joint search and track missions-an information-value approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2551–2565, 2012.

[2] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information Fusion*, vol. 4, no. 4, pp. 259–280, 2003.

[3] B. L. Guo, Q. Zhang, and Y. Hou, "Region-based fusion of infrared and visible images using nonsubsampled contourlet transform," *Chinese Optics Letters*, vol. 6, no. 5, pp. 338–341, 2008.

[4] L. Kun, G. Lei, L. Huihui, and C. Jingsong, "Fusion of infrared and visible light images based on region segmentation," *Chinese Journal of Aeronautics*, vol. 22, no. 1, pp. 75–80, 2009.

[5] S. F. Yin, L. C. Cao, Y. S. Ling, and G. Jin, "One color contrast enhanced infrared and visible image fusion method," *Infrared Physics and Technology*, vol. 53, no. 2, pp. 146–150, 2010.

[6] L. Hong, Z. H. He, J. Xiang, and S. Li, "Fusion of infrared and visible image based on genetic algorithm and data assimilation," in *Proceedings of International Workshop on Intelligent Systems and Applications, (ISA '09)*, pp. 1–5, May 2009.

[7] Z. F. Shao, J. Liu, and Q. M. Cheng, "Fusion of infrared and visible images based on focus measure operators in the curvelet domain," *Applied Optics*, vol. 51, no. 12, pp. 1910–1921, 2012.

[8] J. Y. Liu, Q. Wang, and Y. Shen, "Comparisons of several pixel-level image fusion schemes for infrared and visible light images," in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference, (IMTC '05)*, vol. 3, pp. 2024–2027, May 2005.

[9] X. Q. Zhang, Q. L. Chen, and T. Men, "Comparison of fusion methods for the infrared and color visible images," in *Proceedings of 2nd IEEE International Conference on Computer Science and Information Technology, (ICCSIT '09)*, pp. 421–424, August 2009.

[10] J. Saeedi and K. Faez, "Infrared and visible image fusion using fuzzy logic and population-based optimization," *Applied Soft Computing*, vol. 12, no. 3, pp. 1041–1054, 2012.

[11] J. Zhang, S. Y. Chen, S. Liu et al., "Normalized weighted shape context and its application in feature-based matching," *Optical Engineering*, vol. 47, no. 9, Article ID 097201, 2008.

[12] C. Cattani and A. Kudreyko, "On the discrete harmonic wavelet transform," *Mathematical Problems in Engineering*, vol. 2008, Article ID 687318, 2008.

[13] F. H. Yao and A. Sekmen, "Multi-source airborne IR and optical image fusion and its application to target detection," *Lecture Notes in Computer Science*, vol. 5359, no. 2, pp. 651–660, 2008.

[14] F. H. Yao, A. Sekmen, and M. Malkani, "A novel method for real-time multiple moving targets detection from moving IR camera," in *Proceedings of 19th International Conference on Pattern Recognition (ICPR '08)*, pp. 1–4, December 2008.

[15] F. H. Yao, G. F. Shao, A. Sekmen, and M. Malkani, "Real-time multiple moving targets detection from airborne IR imagery by dynamic Gabor filter and dynamic gaussian detector," *Eurasip Journal on Image and Video Processing*, vol. 2010, no. 9, Article ID 124681, pp. 1–22, 2010.

[16] C. Cattani, S. Y. Chen, and G. Aldashev, "Information and modeling in complexity," *Mathematical Problems in Engineering*, vol. 2012, Article ID 868413, 4 pages, 2012.

[17] S. Y. Chen, Y. H. Wang, and C. Cattani, "Key issues in modeling of complex 3D structures from video sequences," *Mathematical Problems in Engineering*, vol. 2012, Article ID 856523, 17 pages, 2012.

[18] S. Y. Chen, J. H. Zhang, Y. F. Li et al., "A hierarchical model incorporating segmented regions and pixel descriptors for video background subtraction," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 118–127, 2012.

[19] S. T. Xu, Y. F. Niu, L. C. Shen et al., "Real-time moving target detection method form unmanned airborne camera image," *Computer Engineering*, vol. 36, pp. 271–273, 2010 (Chinese).

[20] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in *Proceeding of the 4th International Conference on Computer Vision (ICCV '93)*, pp. 173–182, May 1993.

[21] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 493–500, 2007.

[22] Y. H. Yuan, J. L. Zhang, B. K. Chang, and Y. Han, "Objective quality evaluation of visible and infrared color fusion image," *Optical Engineering*, vol. 50, no. 3, Article ID 033202, 2011.

Advances in
Operations Research

Advances in
Decision Sciences

Mathematical Problems
in Engineering

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Submit your manuscripts at
http://www.hindawi.com

Hindawi

Journal of
Complex Analysis

Journal of
Mathematics

International Journal of
Stochastic Analysis

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

Journal of
Applied Mathematics

Journal of
Optimization