A Region-Based Image Fusion Algorithm for Detecting Trees in Forests

Zheng Yu, Lei Yan[*] and Ning Han

School of Technology, Beijing Forestry University, Beijing, China

Abstract: Multi-sensor image fusion has been widely used in many areas. In this paper, a region-based method using NSCT and fuzzy logic is proposed to enhance trees' contour and enrich the information in images. Firstly, the required source images are segmented, which are then decomposed using the NSCT transform. In low frequency domain, the fuzzy fusion method is used. And in high frequency domain, a region-based rule is proposed to highlight the objects' detail within images. Finally, the fused images are obtained by an inverse NSCT transform. Nine pieces of objective criteria and computation time are adopted to evaluate the fusion result. Moreover 30 groups of collected images are utilized to prove the robustness of the proposed method. The results show that the algorithm could provide fused images which have a higher visual effect, combine more information and consume less time. Furthermore, this process could better capture the contour and texture of the trees.

Keywords: Fuzzy logic, live wire segmentation, NSCT, region based image fusion, tree detection.

1. INTRODUCTION

In recent years, forestry management has become an important research topic. As forests are located within complicated surroundings, it is difficult to recognize targets with a single sensor. Therefore, image fusion algorithm has been applied in this area. Image fusion is a branch of multi-sensor fusion, the goal of which is to combine composite information into a signal image. Infrared images always contain information about objects' contour which cannot be seen in visible images. However, the details are well exposed in visible images. Therefore, the main part is to extract and combine useful information.

Image fusion algorithm can be classified into three levels: pixel-level, feature-level and decision-level. Pixel-level algorithms have been widely used in military detection, medical research and agricultural engineering. They not only extend the information an image covers, but also prompt its visual effect. In [1], the fused image was used to supervise fruit. And in [2], with the help of fused images, Yao collected and analyzed the wetland ecosystem information. Furthermore, Xu fused medical images and consequently obtained accurate information [3]. In recent years, many novel algorithms have been proposed. For example, Jiang proposed a new algorithm based on a modified self-generating neural network and fuzzy logic, which obtained a better visual performance [4]. Moreover, with the help of fused images, Yan promoted the accuracy of recognition [5].

Although a pixel level algorithm is effective in capturing the targets in an image, the amount of time this method takes seriously affected its application. In this paper, a feature-level image fusion algorithm, which not only improves the images' quality but also consumes less time, is proposed to detect trees in forests.

*Address correspondence to this author at the School of Technology, Beijing Forestry University, Beijing, China; Tel: 86-010-62338095/62338097; E-mail: mark_yanlei@bjfu.edu.cn

2. IMAGE FUSION METHODS

2.1. Image Segmentation Method

The live wire segmentation method is a widely used tool to extract targets from images. It was firstly proposed by Barrett and Mortensen in 1992 [6], and can be divided into two parts: defining the cost function and searching for the optimal path. The cost function which measures the boudaries of objects, is defined based on the 8 neighboring pixels as shown in formula 1[7]:

$$f(\text{p},\text{q}) = \omega_{Log} f_{Log}(\text{q}) + \omega_{canny} f_{canny}(\text{q}) \\ + \omega_G f_G(\text{q}) + \omega_D f_D(\text{p},\text{q}) \tag{1}$$

Where ω represents the weight of each feature, $f_{Log}(q)$ is the result of Log edge detection in pixel q, $f_{canny}(q)$ represents the canny edge detection, $f_G(q)$ refers to the gradient magnitude, and $f_D(\text{p},q)$ refers to the gradient direction from pixel p to q. A low cost function means there is a strong edge. Thus, $f_{Log}(q)$ and $f_{canny}(q)$ are set as the adverse aspects of the edge detection result. As for the gradient magnitude and gradient direction, the following measures are taken [8]:

$$f_G(q) = 1 - \frac{G}{\max(G)} \tag{2}$$

Where $G = \sqrt{I_x^2 + I_y^2}$, and I_x, I_y are the gradient components in each direction.

$$f_D(p,q) = \frac{1}{\pi} \{\cos[d_p(p,q)]^{-1} + \cos[d_q(p,q)]^{-1}\} \tag{3}$$

Where $d_p(p,q) = D'(p) \times L(p,q)$, $d_q(p,q) = L(p,q) \times D'(q)$ and $D'(p) = (I_y(p), -I_x(p))$, in which $L(p,q)$ is the unit vector between pixel p and q.

According to the cost function, the boundary is acquired from the start node to the goal node by searching for the minimal cumulative cost value.

Fig. (1). Results of the image segmentation. (From left to right: infrared image, visible image and segmentation result).

It is well known that the infrared image mainly demonstrates the contour in an image. Therefore it is adopted to separate image into different areas. Fig. (**1**) shows the result of image segmentation.

2.2. Image Fusion

In feature level image fusion, it is very crucial to segment images into different areas, and which is regarded as the most difference between pixel and feature lever algorithm. The next step is to transform the original images into multi-scale and multi-directional domains. The high pass sub-band is related to the detail contour and the low pass sub-band is called an approximate image, which shows the most information from the source image. The procedure of this method is shown in Fig. (**2**).
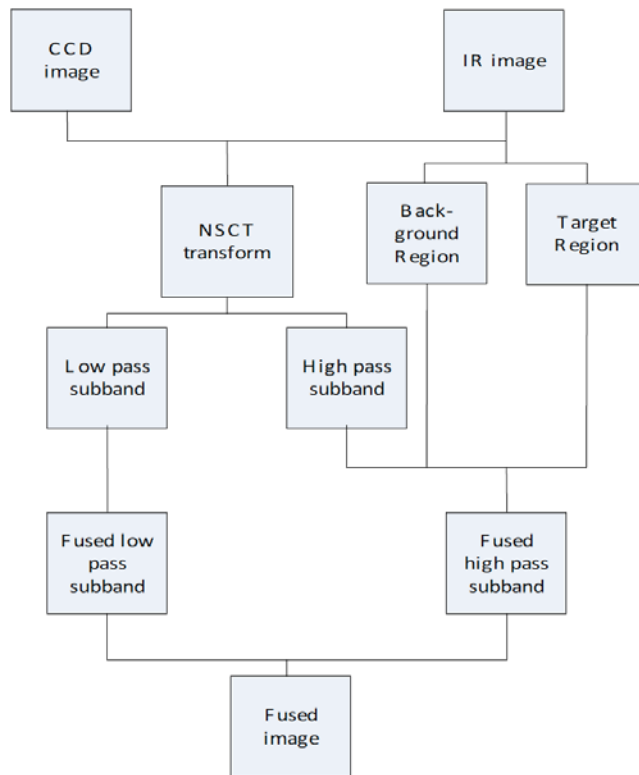

Fig. (2). Procedure for the proposed method.

2.2.1. The Non-subsampled Contoured Transform

NSCT is improved from contourlet transform which is proposed by M.N. Do and Martin Vetterli in 2002 [9]. Con-

tourlet transform is constructed firstly by the Laplacian pyramid (LP) filter, which is used to capture the point discontinuities, and then followed by a directional filter bank (DFB), which links point discontinuities into linear structures. It is excellent at dealing with images which have smooth contours. However, the Gibbs phenomenon caused by being lack of a shift-invariance, seriously affect the its result. Thus, *NSCT* abandons the down-sample and up-sample procedures. With the help of *NSCT*, the images are decomposed into a low pass sub-band and high pass sub-bands.

2.2.2. The Fusion Rules

A large number of algorithms have been applied in this area in order to improve the quality of the fused image. Lewis proposed a novel region-based method, which produced a fused image with a high contrast ratio and much more useful information [10]. Furthermore, Jiang fused images with modified Self-Generating Neural Network and fuzzy logic [4]. Liu improved the PCNN algorithm to fuse images and achieved a better result [11]. Finally, Jin obtained high quality fused image using a novel fusion algorithm called teaching learning based optimization method [12]. In this paper, a novel region based algorithm is proposed to detect trees in forests.

(1) The low pass domain

More useful information is maintained in low pass sub-band, so the fusion in this domain becomes crucial for targets enhancement and information expansion. The main goal is to preserve texture and background captured in visible images. In the past research, the most used method can be classified into two categories: coefficient selection and weighted average. Because of the high contrast ratio between our collected images, the former method becomes not feasible. In contrast, the latter approach is excellent at decreasing the influence of noise and false contours, which seriously affect the fusion quality. So the weighted average method is adopted. In previous research, it is difficult to determine whether a pixel or region belongs to the target or background. So the fuzzy fusion rule is utilized to work out the degree of membership, which digitally measures a pixel or region's belonging degree.

The right image in Fig. (**3**) is the degree of membership in which the pixel value is ranged from 0 to 1. Here a larger value means a higher probability the pixel belonging to target area. Therefore, an uncertain problem is digitally solved.

As for determining the membership function, the infrared image provides a possible approach. It is well known that the

Fig. (3). Degree of membership. (From left to right: infrared image, visible image and degree of membership).

infrared image is related to thermal information, and as shown in Fig. (**3**), the trees are obviously warmer than their surroundings. That is to say the pixel value in the infrared images, which is related to temperature, is likely to offer a feature to distinguish trees from the other surroundings. Then, the Gaussian membership function was used to describe the degree of membership:

$$\lambda_1(i,j) = \exp[-\frac{(f_1(i,j) - \mu)^2}{2(k\sigma)^2}] \tag{4}$$

Where σ is the standard deviation, f is the low pass decomposed coefficient of the infrared image, μ is the mean value of this sub-band, and based on an experiment, the constant k is set as 1.5 [14].

Therefore, the fusion rule in the low pass sub-band is characterized as shown in formula 5:

$$F(i,j) = \lambda_1(i,j)f_1(i,j) + (1 - \lambda_1(i,j))f_2(i,j) \tag{5}$$

Here, f_1 refers to the infrared low pass sub-band, and f_2 denotes the visible low pass sub-band.

(2) High pass domains

The high pass sub-bands mainly describe the detail contour of the target in multi-direction and multi-scale. Generally, a larger high frequency coefficient means a stronger mutation between the neighboring pixels in the original image. Therefore, the goal of the fusion in this domain is to intelligently acquire more abundant detial information among source images. The most widely used method is simply selecting the maximal coefficient [13, 14]. However, it is easy for the aforementioned method to neglect details, especially texture on trees. Consequently, the aforementioned simple fusion rule may be unsuitable.

Then, a region-based rule is proposed in this paper. As mention in section 2.1, the targets are separated from their background. In each region, the local standard deviation is used to measure the activity of each coefficient as shown in formula 6. Furthermore, due to the influence caused by its neighboring pixels, the Gaussian filter is chosen to improve the accuracy of the activity and decrease influence of noise.

$$F^l(i,j \mid d) = \sqrt{\frac{1}{W} \sum_{-m}^{m} \sum_{-n}^{n} (\left| y^l(i+m, j+n \mid d) \right| - \left| \overline{y}(i,j \mid d) \right|)^2} \tag{6}$$

$$NF^l(i,j \mid d) = \sum_{-m}^{m} \sum_{-n}^{n} G(m,n) \times F^l(i+m, j+n \mid d) \tag{7}$$

Where, $G(m,n)$ refers to a window weight obtained by the Gaussian filter, and l, d refer to the decomposed level and direction.

Subsequently, on the basis of *NF*, a parameter which measures the activity of each pixel is obtained with the below formulas:

$$P^l(i,j \mid d) = \begin{cases} 1 & if & NF_A^l(i,j \mid d) > NF_B^l(i,j \mid d) \\ 0 & if & NF_A^l(i,j \mid d) < NF_B^l(i,j \mid d) \end{cases} \tag{8}$$

Where, *A* and *B* denote the infrared image and visible image.

Then, region-based activity is used to choose the most important high frequency coefficient among the source images. In order to avoid the disadvantage that the parameter is sensitive to high coefficient, a new method to generate this parameter was proposed in [15]. Based on *P* and the segmentation result obtained above, the region based activity is generated using formula 9.

$$R^l(R \mid d) = \begin{cases} 1 & if & (\sum_R P^l(n \mid d)) > \frac{R}{3} \\ 0 & if & (\sum_R P^l(n \mid d)) < \frac{R}{3} \end{cases} \tag{9}$$

In each region, *R* is defined by analyzing *P* and the size of this region, if the number of one in *P* beyond one third of the total pixels' number in this region, the *R* of this region is set as 1. On the contrary, it is set as 0.

Once *P* and *R* are obtained, the fusion rule is chosen among the formula below according to the diversity (*DIS*) between the source sub-band images. Moreover, the diversity is defined as shown in equations (11) and (12).

$$\begin{cases} y_1^l(i,j \mid d) = P^l(i,j \mid d) \times y_A^l(i,j \mid d) + (1 - P^l(i,j \mid d)) \times y_B^l(i,j \mid d) \\ y_2^l(i,j \mid d) = R^l(i,j \mid d) \times y_A^l(i,j \mid d) + (1 - R^l(i,j \mid d)) \times y_B^l(i,j \mid d) \\ y_3^l(i,j \mid d) = \frac{NF_A^l(i,j \mid d) \times y_A^l(i,j \mid d) + NF_B^l(i,j \mid d) \times y_B^l(i,j \mid d)}{NF_A^l(i,j \mid d) + NF_B^l(i,j \mid d)} \end{cases} \tag{10}$$

$$DNF^l(i,j \mid d) = \left| NF_A^l(i,j \mid d) - NF_B^l(i,j \mid d) \right| \tag{11}$$

$$DIS^l(i,j \mid d) = \sqrt{\sin(\frac{\pi \times DNF^l(i,j \mid d)}{2 \times T})}, \\ T = \max_{i,j} DNF^l(i,j \mid d) \tag{12}$$

Here, l and d refer to the decomposed level and direction, A and B denote the infrared image and visible image, and y is the decomposed coefficient.

If the diversity at a given position is high ($DIS \in [\frac{2}{3}, 1]$),

then select the first fusion rule in equation (10);

If the diversity at a given position is medium ($DIS \in (\frac{1}{3}, \frac{2}{3})$), then select the second fusion rule in equation (10);

If the diversity at a given position is low ($DIS \in [0, \frac{1}{3}]$),

then select the third fusion rule in equation (10).

The high pass sub-bands of Fig. **4**(**a**) and (**b**) is fused using both the maximal selection method and the proposed method. As a widely used image quality assessment criterion, entropy is adopted to evaluate the result. Larger entropy means more information an image maintains. Subsequently, the criterion of the former result was 2.35 and that in the latter one increase by 45.1%. Therefore, it is obvious that the latter method captures more details especially in selected target regions.
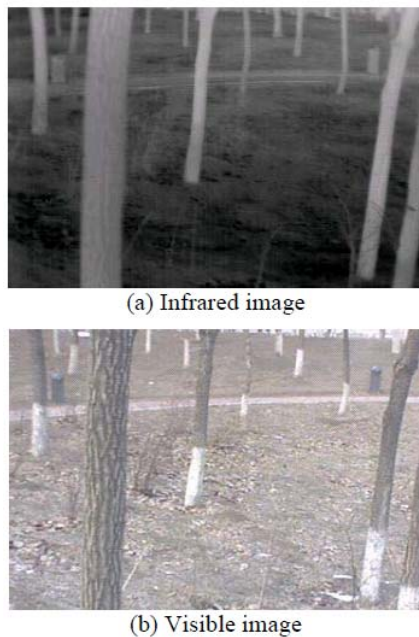

(a) Infrared image

(b) Visible image

Fig. (4). The fusion result of the high pass domain.

3. THE EXPERIMENTAL RESULT AND DISCUSSION

In order to objectively assess the quality of fused images, many quality assessment criteria have been proposed in recent research. Consequently, nine types of them are selected to thoroughly assess the fusion effect, including the entropy [16], average gradient, cross entropy, standard deviation, RMSE, mutual information (MI) [17], SNR, PSNR and spatial frequency.

Fig. (**5**) and Table **1** display the fused result of Fig. (**2**). Furthermore, the data in Table **1** is mentioned in Ref. [14, 15], the first five rows are cited from [15], the sixth row is captured from [14], and the seventh is calculated according

to the method in [13]. Here the entropy which directly concludes the performance of the image fusion is described. It digitally shows the amount of information the image maintains, and evaluates the fusion quality. The computation time is only calculated during the processing of the low frequency domain in a familiar hardware and software situation. From Table **1**, it can be concluded that although the entropy contained in Saeedi's work beyond that in the proposed approach, it takes more than twice time. Furthermore, compared with other methods, the proposed method presents more information without wasting too much time. Therefore, the proposed method is significant and has a certain practical value.

Fig. (5). Fused images. (**a-b**: method in paper [14], proposed method).

Table 1.　Quality assessment.

Method	Entropy	Computation time
LT	6.71	0.771
WA	6.72	0.821
WBV	6.76	0.984
Lewi	6.66	1.703
Saeedi	7.41	2.253
Yan	7.16	1.570
Yin	6.97	0.973
Proposed method	7.00	1.056

In addition, some original images have been collected in the National Olympic Forest Park using Fluke Ti55 infrared thermal camera, in winter and spring when visible and thermal information are mutual complement.

Fig. (**6**) shows two groups of images, from left to right, infrared image and visible image. In the infrared images, the contour of the trees is obviously stronger than that in the visible images. However, some detailed information such as the texture on trees and the surroundings is portrayed better in the visible images. This can be regarded as complementary information. Fig. (**7**) is the fusion result of Fig. (**6**) using different fusion rules.

Obviously the contour and surroundings in the fused images are enhanced compared with every single source image. However, in contrast with the fused images resulted from other fusion methods, Fig. **7**(**d**) does well in capturing the detailed texture of the trees and the background, presenting a strong edge at the same time.

Fig. (6). Collected images.

$(a - d)$

Fig. (7). Fused images. ((**a**) – (**d**): method in [1], improved method of [1], method in [13], proposed method).

Table 2.   Quality assessment.

Method	Entropy	Average gradient	Standard deviation	Cross entropy	SNR	RMSE	Spatial frequency	PSNR	MI	Run time
(a)	6.64	11.72	35.75	**0.79**	13.60	50.87	23.49	29.89	4.60	7.21
(b)	6.85	12.86	36.14	1.34	13.55	51.09	24.50	31.86	4.58	92.19
(c)	7.01	10.63	46.52	1.25	13.41	53.16	21.15	32.10	4.62	24.85
Yan	**7.31**	**15.09**	**54.01**	2.51	10.42	60.00	**28.96**	29.42	4.63	157.4
proposed	7.07	13.19	42.97	1.79	**16.31**	**45.48**	25.73	**35.07**	**5.15**	85.78

In addition, we have deal with 30 groups of images to prove the feasibility of proposed method. Subsequently, an average quality assessment is summarized as shown in Table **2**.

The fourth method in Table **2** is mentioned in [14]. The bold font denotes the best assessment of each item. From this table it can be seen that the proposed algorithm gets four best assessments out of the nine items, while four items rank the second. This result indicates that the proposed algorithm could further improve the quality of the fused image in both the preservation of information and the promotion of visual effect. Compared with the other methods, the proposed algo-rithm could better characterize the reality in forests. Moreo-ver, the contour and texture in the original images is more clearly presented. In addition, although the run time of pro-posed method is much beyond or equal to that of methods (a), (b) and (c), its quality is much better. Furthermore, in contrast with Yan's method, in spite of having a familiar quality, the proposed method saves 50% of the time. There-fore, the proposed algorithm could better present both fea-tures of the source images whilst taking up an appropriate amount of time.

CONCLUSION

In this paper, a region based multi-sensor image fusion algorithm using NSCT and fuzzy logic is proposed. Firstly, source images are segmented into targets and their background using the live wire method. Then the source images are decomposed into multi-scale and multi-direction. In high frequency domain, a region-based method is adopted to enhance targets, for the purpose of presenting more information on trees and make their recognition easier. And in low frequency domain, the fuzzy fusion rule is utilized to avoid the influence of noise and false contours. Finally, the fused image is obtained by reverse NSCT transform.

Compared with some former image fusion algorithms, the proposed method could obtain fused images with higher clarity and better spatial resolution, which makes it performs better in capturing trees in images. Additionally, the quality of the fused image has been proved by nine types of quantitative assessment criteria. Moreover, a shorter running time of the proposed method makes it usable in real-time application. Finally, the robustness of the proposed approach is validated by investigating 30 groups of images collected in forests. The fused image efficiently avoids the influence caused by light, fog and shelter in forests, and provides more information about the actual situation.

The results show that the proposed algorithm could better describe the reality of a complicated situation without spending too much time on computation. Therefore, when used in forestry production and management, it is significant for the promotion of work efficiency.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   D.M. Bulanona, T.F. Burksa, and V. Alchanatisb, "Image fusion of visible and thermal images for fruit detection," *Biosystems Engineering*, vol. 103, no. 1, pp. 12-22, 2009.

[2]   Y. Wei, and H. Min, "Estimation and analysis of wetland ecosystem state parameters via remote sensing technique," *Journal of Dalian University of Technology*, vol. 51, no. 3, pp. 422-427, 2011.

[3]   Z. Xu, "medical image fusion using multi-level local extrema," *Information Fusion*, vol. 19, pp. 38-48, 2014.

[4]   H. Jiang, and Y. Tian, "Fuzzy image fusion based on modified self-generating neural network," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8515-8523, 2011.

[5]   L. Yan, X. Ding, J. Kong, Z. Yu, and J. Liu, "Parameters optimization algorithms for improving the performance of obstacles identification in forest area," *INMATEH-Agricultural Engineering* , vol. 40, no. 2, pp. 43-52, 2013.

[6]   E. N. Mortensen, B. S. Morse, W. A. Barrett, and J. K. Udupa, "Adaptive boundary detection using'live wire' two-dimensional dynamic programming," *IEEE proceedings of Computers in Cardiology,* USA, Durham, pp. 635-638, 1992.

[7]   M. Farber, E. Jan, and H. Heinz, "Live-wire-based segmentation using similarities between corresponding image structures," *Computerized Medical Imaging and Graphics*, vol. 31, no. 7, pp. 549-560, 2007.

[8]   L. Guo, L. Gao, J. Wu, K. Lu, and K. Wang. "New algorithm of live wire based on PCNN simulation research," *Journal of System Simulation*, vol. 20, no. 14, pp. 3707-3711, 2008.

[9]   M. N. Do, and M. Vetterli, "Contourlets: a directional multi-resolution image representation," *Proceedings of International Conference on Image Processing*, vol. 1, pp. 357-360, 2002.

[10]   J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D.R. Bull, and N. Canagarajah, "Pixel and region-based image fusion with complex wavelets," *Information Fusion*, vol. 8, no. 2, pp. 119-130, 2007.

[11]   L. Fu, L. Jin, and C. Huang, "Image fusion algorithm based on simplified PCNN in nonsubsampled contourlet transform domain," *Procedia Engineering*, vol. 29, pp. 1434-1438, 2012.

[12]   H. Jin, and Y. Wang, "A fusion method for visible and infrared images based on contrast pyramid with teaching learning based optimization," *Infrared Physics and Technology*, vol. 64, pp. 134-142, 2014.

[13]   S. Yin, L. Cao, Q. Tan, and J. Guofan, "Infrared and visible image fusion based on nsct and fuzzy logic," *Proceedings of the IEEE International Conference on Mechatronics and Automation*, pp. 671-675, 2010.

[14]   L. Yan, Z. Yu, N. Han, and J. Liu, "Improved image fusion algorithm for detecting obstacles in forests," *Journal of Digital Information Management*, vol. 11, no. 5, pp. 378-384, 2013.

[15]   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.

[16]   M. F. Yakhdani, and A. Azizi, "Quality assessment of image fusion techniques for multi-sensor high resolution satellite images," *In IS-PRS TC VII Symposium – 100 Years ISPRS,* Vienna, Austria, IAPRS, vol. XXXVIII, Part 7B, pp. 204-209, 2010.

[17]   M. Deshmukh, and U. Bhosale. "Image fusion and image quality assessment of fused images," *International Journal of Image Processing*, vol.4, pp. 484-508, 2010.