Improved Image Fusion Algorithm for Detecting Obstacles in Forests

Lei Yan, Zheng Yu, Ning Han, Jinhao Liu
School of Technology
Beijing Forestry University
Beijing, China,
liujinhao@vip.163.com

ABSTRACT: *The harvester has been widely used in forestry operations. However the existence of obstacles in forests, such as stones, animals, human and clustered trees, may leads to safety accident and decreases the operating efficiency of the harvester. Therefore it becomes vital to show the obstacles accurately to provide guidance for the harvester. In this paper, an improved image fusion algorithm based on Contourlet transform and pulse coupled neural network (PCNN) is proposed to enhance the fused images' clarity and capture more abundant information of the reality. As for comparison, algorithms, such as wavelet transform, Principal Component Analysis (PCA) and PCNN are simulated to evaluate the proposed algorithm. Visible and infrared thermal image are captured in forest in this experiment. The experimental results demonstrate that the proposed algorithm outperform the mentioned methods by the objective criteria of Entropy, Average gradient and Standard deviation.*

Categories and Subject Descriptors:
I.2.10 [Vision and Scene Understanding]: Video Analysis;
I.4.10 [Image Representation]

General Terms:
Video Frame Processing, Content Processing

Keywords: Obstacles in Forests, Image Fusion, Improved Algorithm, Contourlet Transform, Pulse Coupled Neural Network

1. Introduction

With the development of forestry technology, manual work has gradually been replaced by harvester, which leads to a promotion of forestry production efficiency by more than 80 times. However, the existence of obstacles, such as stones, animals, human and clustered trees, will seriously affect the automatic operations of harvester, even cause safety accident. So it is crucial to distinguish the obstacles in forests to improve the efficiency and safety of harvester.

Image fusion is a branch of data fusion which refers to the acquisition, processing and synergistic combination of information gathered by various knowledge sources and sensors to provide a better understanding of a phenomenon [1]. Visible images reflect the visual reality and infrared thermal images demonstrate the truth about temperature. Because of the complicated surroundings in forest, it is difficult to detect targets guided by single display (visible images or infrared thermal images). However, a target that is camouflaged for visible or infrared detection cannot be detected, but may be clearly represented in the other images. The fusion of visible and thermal image on a single display may allow both the detection and the unambiguous localization of the target with respect to the context [2].

With the development of fusion algorithm, it has emerged in almost every corner of people's life. A fusion system to enhance the contrast of medical images has been developed by He [3]. Then Li fuses multi-focus images

using an improved algorithm, and obtains a fused image with integrate information and strong correlations [4]. Wen has used fused images in forensic science [5]. Ma has distinguished inner information of the tissue by fused images [6]. Kavitha has provided clear information in medical images by fusing CT images and MRI images [7]. In this paper, an improved image fusion algorithm based on PCNN and Contourlet transform is proposed to detect obstacles in forests.

2. Fusion Methods

In order to further improve the capabilities of discrimination, the process of fusion was implemented by three steps.

(1) Source images were decomposed by the Contourlet transform. In this step, images were decomposed into low-pass sub-band and high-pass sub-bands.

(2) The coefficients of decomposition were fused by rule of PCNN, during which the decomposed coefficient was determined by the rule of PCNN.

(3) The fused images were reconstructed by Contourlet inverse transform. In this step, the determined coefficient was transformed into pixel value of an image.

The producer was arranged as shown in figure 1:

A double filter bank structure is constructed in which at first the *Laplacian pyramid* (LP) is used to capture the point discontinuities, and followed by a directional filter bank (DFB) to link point discontinuities into linear structures [9]. After this process, the source images can be decomposed into low-pass sub-band and high-pass sub-bands. The low-pass sub-band image can be called approximate image, which shows the most information of the source image. High-pass sub-bands images capture the detail contours of the targets in multi-direction in original image.

2.3. Fusion rules

In order obtain a fused image with better clarity and more information, fusion rules, such as wavelet transform, PCA and PCNN, have been made inroads into image fusion. Liu used wavelet transform to fuse images and got a good fusion result [10]. Gong used PCA and wavelet transform to fuse images and obtained an image with both high resolution and rich spectrums [11]. Singh proposed an improved algorithm based on wavelet decomposition which enhanced the fusion efficiency and improved fusion accuracy [12]. Duan proposed a license plate recognition system with the help of wavelet transform [13]. Zhao fused visible image and infrared image using PCNN and obtained an image of higher contrast ratio and clarity [14].

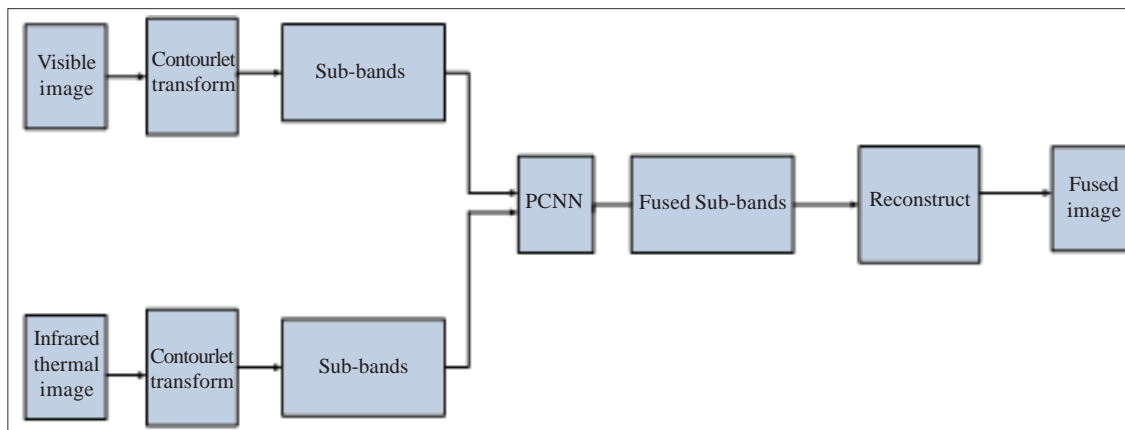In this study, the mentioned fusion rules were simulated

Figure.1 Procedure for image fusion

2.1 Image capture
Both visible and infrared thermal images were collected by Fluke Ti55 infrared thermal camera in the National Olympic Forest Park, where trees, animals, stones, and human can be captured. And the images were taken before sunset, at midday and after sunset, when visible and temperature information are mutual complement.

2.2 Contourlet transform
Contourlet transform is first proposed by M . N. Do and Martin Vetterli in 2002 [8]. It is improved from the research of wavelet transform with an advantage of dealing with piecewise smooth images having smooth contours. The structure of filter banks of Contourlet transform is shown in figure 2.
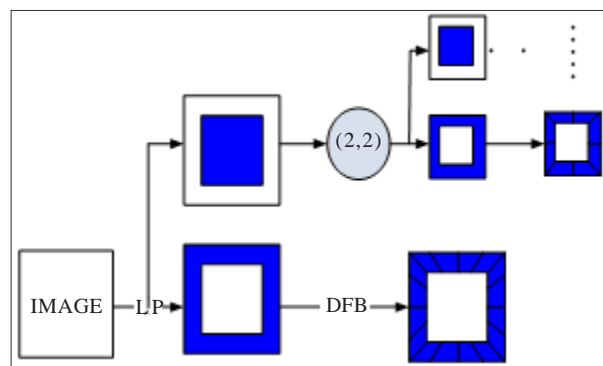

Figure.2 Structure of filter banks of Contourlet transform

to estimate the proposed method. The result show that the fused image using proposed method is of a higher clarity and more information compared with the result of other methods.

2.3.1 Theory of PCNN
Based on the research of visual cortex neuron of mammals, Eckhorn firstly proposed pulse coupled neural network [15]. The neuron model of PCNN is shown in Figure 3.
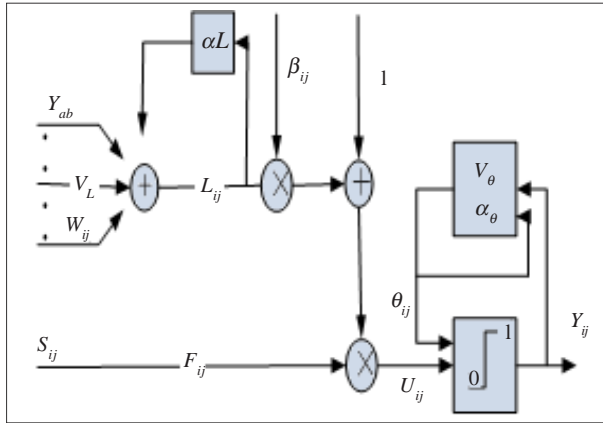

Figure 3. Neuron model of PCNN

Every PCNN neuron is consist of three parts: the receptive field, modulation field and pulse generator.

In the first part, Y_{ab} is the output of the neuron nearby, ω_{ij} is the synaptic gain strength between neurons, V_L is a constant, S_{ij} is the input of a model, α_L is the attenuation time constant, L_{ij} is the linking input.

In the second part, F_{ij} is the feeding input, β is the linking strength.

And in the last part, U_{ij} is the result of modulation, θ_{ij} is the dynamic threshold which changes with the variation of the neuron's output pulse, V_θ is the amplitude gain, α_θ is time constant, Y_{ij} is the output of the neuron. The rule can be conducted as follows [16]:

$$
\begin{cases}
F_{i,j}^k(n) = S_{i,j}^k(n) \\
L_{i,j}^k(n) = \exp(-\alpha_L) L_{i,j}^k(n-1) + V_L \sum_{a,b} \omega_{ij,ab} Y_{ab}(n-1) \\
U_{i,j}^k = F_{i,j}^k(n) \times (1 + \beta L_{i,j}^k(n)) \\
\theta_{i,j}^k = \exp(-\alpha_\theta) \theta_{i,j}^k(n-1) + V_\theta \times Y_{i,j}^k(n) \\
Y_{i,j}^k = step(U_{i,j}^k(n)) - \theta_{i,j}^k(n))
\end{cases}
$$

When applied to image processing, PCNN is a partial linked network (one layer and two dimensions). Each decomposition coefficient can be regarded as the input of a neuron, and the neuron is connected with 8 neurons nearby, as figure 4 shows. The output of a neural is 1 or 0 (fire or not fire). During the iteration, when a neural

produces pulse, the neural nearby will be affected by the signal and output its signal ahead of time. Therefore the information is transmitted and coupled automatically. That is the foundation of PCNN's application in image fusion.
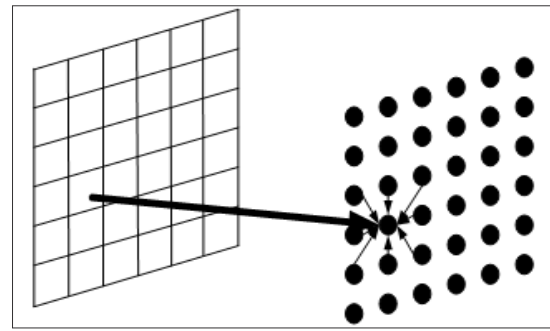

Figure 4. Diagram of image processing using PCNN (□denotes pixel, ●denotes neuron)

As shown in formula (1), it becomes easy for a neuron to capture the pulse generated by a neuron nearby. So it will produce its pulse ahead of time, then the dynamic threshold will increase rapidly which cease the neuron from producing pulse. In the next loop, the neurons nearby will capture its signal and output its signal ahead of time. Therefore, after N loops, the more pulsing signals a neuron produces, the better feature (compared with what inputs describe) the neuron describes.

Mostly, the larger the pixel's value is, the more possible the pixel generates pulse, and the more signals the pixel may produce. In [14, 16, 17], the theory of PCNN is to decide which pixel at the position of (i, j) (visible image or infrared thermal image) performs best in capturing targets. Then the value of the best pixel is determined as the pixel value of the fused image at the corresponding position.

2.3.2 Improved rule of PCNN
The above method deals effectively with the fusion of multi-focus images, however its weakness emerges when dealing with the fusion of multi-sensor images, especially when the grey-level of the source images are quite different. In that situation, the obvious difference between parts from different source images decreases the visual effect of the fused image. Thus, an improved fusion algorithm is proposed.

Based on experiments, we set $\alpha_L = 0.2$, $\alpha_\theta = 0.2$, $V_L = 1.5$,

$$
\beta = 0.2, N = 3000, \omega = \begin{vmatrix} 0.707 & 1 & 0.707 \\ 1 & 0 & 1 \\ 0.707 & 1 & 0.707 \end{vmatrix}
$$

Sum-modified-Laplacian (SML) that measures the clarity of images can be defined as follows [18]:

$$
\begin{aligned}
ML(i, j) = &|2I(i,j) - I(i \pm 1, j)| + |2I(i, j) - I(i, j \pm 1)| \\
&+ |2I(i, j) - I(i \pm 1, j \pm 1)| + |2I(i, j) - I(i \pm 1, j \mp 1)|
\end{aligned} \quad (2)
$$

$$
SML(i, j) = \sum_m \sum_n \omega(m, n) [ML(i+m, j+n)]^2 \quad (3)
$$

Where $I(i, j)$ denotes the decomposition coefficient at position of (i, j). To describe the areas which did not change acutely, we set $\omega(m, n) = \frac{1}{15} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix}$. This method overcomes the drawback, including gray level distortion and blocky effect, which are resulted from the discontinuous boundary between clear district and vague district, and better describes the clarity of the image.

In low-pass sub-band images, SML which represents clarity of an image is chosen as input of PCNN. In this situation, the more signals a neuron produces, the more energy the pixel contains, and the better the corresponding pixel dose in capturing information. Then we record the number of pulsing signal a neuron produce as $T^k(i, j)$, where k denotes visible image or infrared thermal image, and (i, j) indicates the position of (i, j).

In order to eliminate the weakness of the rule of PCNNsuch blocky effect resulted from the difference between the gray-level of source images, we divide this sub-band image into three parts by comparing the times of pulsing signals: $P_A \in \{P \mid T^A(i, j) - T^B(i, j) > S\}$, $P_B \in \{P \mid T^B(i, j) - T^A(i, j) > S\}$, and $P \in \{P \mid T^A(i, j) - T^B(i, j) \mid \leq S\}$. Where S is a threshold which is set based on experiments as 5.

And different part has its own fusion rule. The rules of each part are described as formula (4) shows:

$$\begin{cases} I(i, j) = I_A(i, j), & T^A(i, j) - T^B(i, j) > S \\ I(i, j) = \rho_A I_A(i, j) + \rho_A I_A(i, j), \mid T^A(i, j) - T^B(i, j) \mid \leq S \quad (4) \\ I(i, j) = I_B(i, j), & T^B(i, j) - T^A(i, j) > S \end{cases}$$

Where, $\rho_A = \dfrac{T^A(i, j)}{T^A(i, j) + T^B(i, j)}$, and

$$\rho_B = \dfrac{T^B(i, j)}{T^A(i, j) + T^B(i, j)}.$$

In high-pass sub-band images, the decomposed coefficients are adopted as input of the network.The more pulsing signals a neuron produces, the stronger contrast ratio the neuron presents [19]. The rule of this sub-band can be written as:

$$\begin{cases} I(i, j) = I_A(i, j), & if \ T^A(i, j) > T^B(i, j) \\ I(i, j) = I_B(i, j), & if \ T^B(i, j) > T^A(i, j) \end{cases} \quad (5)$$

Where A and B denotes the visible image and the infrared thermal image, $I(i, j)$ is the decomposition coefficient at position of (i, j).

3. Experimental Results and Discussion

We have selected three kinds of quantitative assessment criteria to estimate the fusion performance, including Entropy, Average gradient and Standard deviation. Entropy is a measure that directly concludes the performance of

(a) Visible image   (b) Infrared thermal image   (c) Fused image with wavelet

(d) Fused image with PCA   (e) Fused image with PCNN   (f) Fused image with proposed method
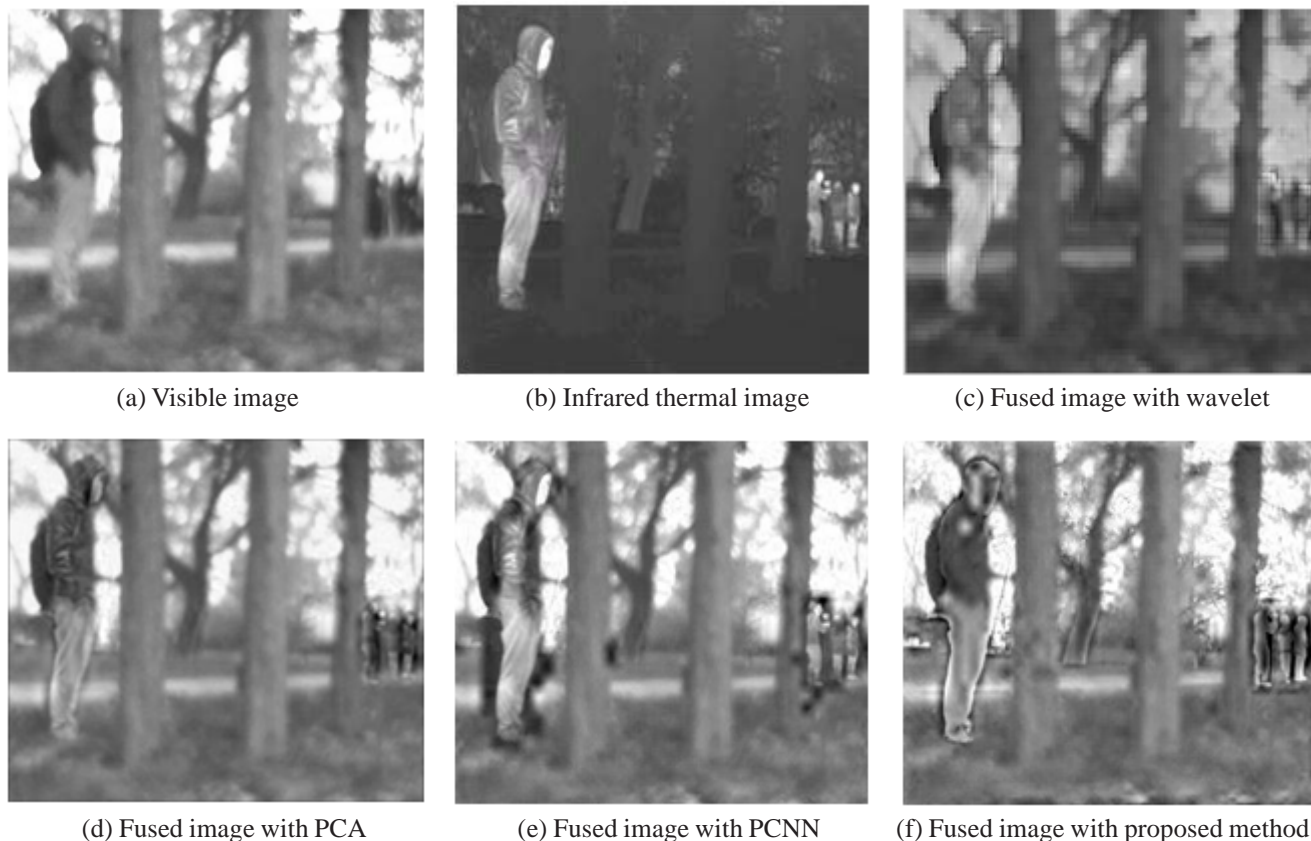
Figure 5. Performance of image fusion using different fusion rules

image fusion. The greater the Entropy is, the more abundant information the image maintains, which represents a higher quality of fusion [20, 21].

Average gradient is a measure which represents clarity of an image. A larger average gradient indicates a higher spatial resolution [22]. In addition, standard deviation measures the degree that the data points deviate from the mean. A lager standard deviation means a better visual effect. Then the criteria above were calculated as follows:

$$E = - \sum_{i=0}^{225} P_i \log_2 P_i \qquad (6)$$

Where E represents Entropy of an image, P_i is the probability of pixel i in an image.

$$\overline{grad} = \frac{1}{(m-1)(n-1)} \sum_{m-1} \sum_{n-1} \sqrt{\frac{\Delta F}{2}} \qquad (7)$$

In this formula, $\Delta F = (F(i,j) - F(i+1,j))^2 + (F(i,j) - F(i,j+1))^2$.

$\overline{grad}$ indicates average gradient, and $F(i,j)$ is the gray-value at the position of (i,j), m and n represent the size of the image.

$$std = \sqrt{\frac{\sum (F(i,j) - \bar{F})^2}{(n-1)}} \qquad (8)$$

Where std denotes the standard deviation of an image, and $F(i,j)$ is the pixel value of the image at position of (i,j), $\bar{F}$ is the average pixel value, n is the number of pixel of an image.

It is suggested in Table 1 that compared with the mentioned methods, the fused image using proposed method captures more abundant information from the source images. What is more, the clarity, spatial resolution and visual effect of the fused image using proposed method are obviously enhanced compared with the results of

wavelet transform, PCA and PCNN. Compared with the result of PCA, the human in fused images using proposed method are obviously clearer and more evident.

Even, from figure 5(e) and figure 5(f) we can conclude that the weakness of PCNN especially the blocky effect is eliminated without decreasing the fusion effect. The edge of trees is much sharper in figure 5(f), and the blocky effect around the people in the left of figure 5(e) is avoided in figure 5(f).

Figure 6 is examples of fusion using proposed method, (1) of each line are visible images taken in National Olympic Forest Park, and (2) of each line are infrared thermal images taken at the same time, and the same view. And they reflect information about temperature. (3) of each line are the fused images using proposed method. Because of the similarity between the temperature of trees' contour and the surroundings, the edge of trees can't be detected accurately in infrared thermal images, as is shown in (2) of each line, especially in a(2) and c(2), some trees are obviously far more thinner than the reality as the visible images show,which do promote the possibility of accidents in forests.However, visible images provide a complement to recognize the targets accurately. Then the mentioned drawback is well avoided in the fused images. In addition, because of the somber light, the contrast ratio of the Figure 6 (a1) was obviously weaker than usual, and it was not enough to support the recognition. Thus, fused image as shown in the Figure 6 (a3) captured the advantages of both visible image and infrared thermal image.The targets in fused images become much easier to be recognized compared with that in eitherthe visible images orthe infrared thermal images.

In table 2, the assessment criteria of fused images arebetter than that of visible images and infrared thermal images, which suggest that the fused images capture more abundant information, have a higher spatial resolution and show a better visible effect, compared with single original

	Entropy	Average gradient	Standard deviation
Wavelet	7.10	6.38	107.98
PCA	7.27	5.21	49.66
PCNN	7.33	8.92	53.06
Proposed Method	7.33	10.82	52.24

Table 1. Quality assessment of different fusion rules

Entropy			Average gradient			Standard deviation		
(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
6.09	3.20	6.42	5.59	3.18	7.86	22.13	24.36	36.19
6.89	3.62	7.14	10.32	2.85	11.78	43.74	16.53	43.71
7.31	2.55	7.48	15.35	2.82	16.85	52.31	29.42	52.89

Table 2. Quality assessment of image fusion

(a1) Visible image   (a2) Infrared thermal image   (a3) Fused image

(b1) Visible image   (b2) Infrared thermal image   (b3) Fused image

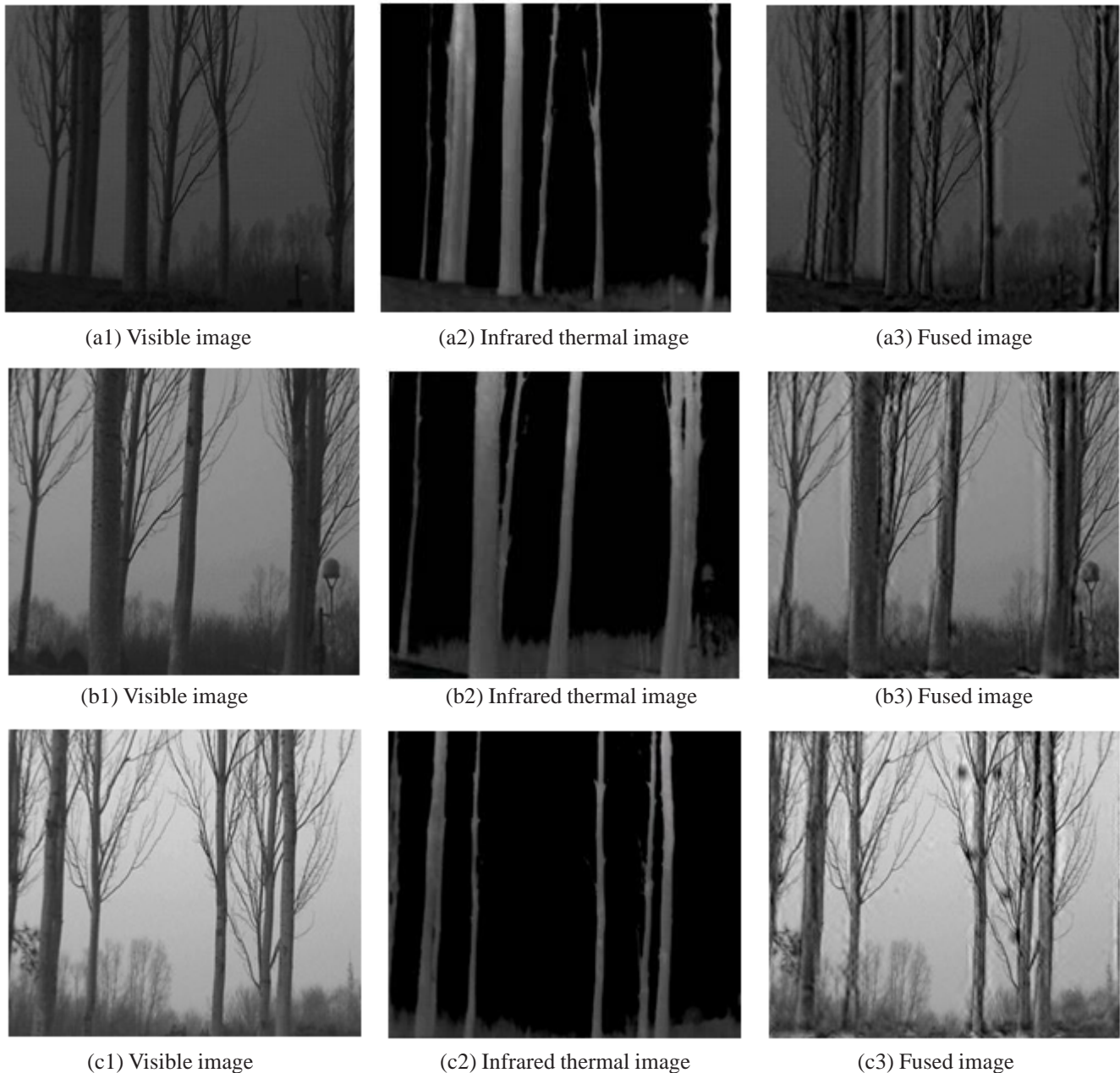(c1) Visible image   (c2) Infrared thermal image   (c3) Fused image

Figure 6. Examples of image fusion with proposed method

images. The results indicate that targets in both visible and infrared thermal images could be more clearly presented in one fused image, and the fused image could better characterize the reality in forests. So the proposed method is available in detecting obstacles in forests.

4. Conclusion

In this paper an improved image fusion algorithm based on Contourlet and PCNN was proposed to detect obstacles in forests. Compared with other former fusion rules, such as wavelet transform, PCA and PCNN, the proposed method performs better in capturing obstacles in images by obtaining a fused image with higher clarity and better spatial resolution. And it is proved by three quantitative assessment criteria. At the same time, the limitation of PCNN like blocky effect is efficiently avoided.

The fusion of visible and infrared thermal images could counteract the effects of light, shelter, fog and other things in forests. Therefore, fused images could better depict true phenomena and might provide clearer visual information of complicated circumstances. So when set on the harvester, it will works as a guide which guarantee the safety and efficiency of people and the harvester.Even further fusion of fused images and other valid data, such as distance data and temperature data, might provide the reality more precisely.With the help of this technology, the efficiency and safety of harvester may be well improved, and the forestry production efficiency might be highly promoted.

5. Acknowledgment

References

[1] Varshney, P. K. (1997). Multisensor data fusion, *Electronics & Communication Engineering Journal*, 9 (6) 245-253.

[2] Toet, A., Ijspeert, J. K., Waxman, A. M., Aguilar, M. (1997). Fusion of visible and thermal imagery improves situational awareness, *Displays*, 18 (2) 85-95.

[3] He, K. (2006). Study on Pixel-level Medical Image Fusion, *Northwestern Polytechnical University*.

[4] Li, Sh. T., Kang, X. D., Hu, J. W., Yang, B. (2013). Image matting for fusion of multi-focus images in dynamic scenes. *Information Fusion*, 4, p. 147-162.

[5] Wen, C. Y., Chen, J. K. (2004). Multi-resolution image fusion technique and its application to forensic science, *Forensic Science International*, 140, p. 217-232.

[6] Ma, L.Y., Feng, N. Z. (2012). Nonsubsampledcontourlet transform based image fusion for ultrasound tomography, *Journal of Nanoelectronics and Optoelectronics*, 7 (2) 216-219.

[7] Kavitha, C. T., Chellamuthu, C., Rajesh, R. (2012). Medical image fusion using combined discrete wavelet and ripplettransfomrs, *Procedia Engineering*, 38, p. 813-820.

[8] Do, M. N., Vetterli, M. (2002). Contourlets: A Directional Multi-resolution Image Representation, *In*: Proceedings 2002 International Conference on Image processing, 1, p. 357-360.

[9] Lin, L. Y. (2008). Contourlet Transform —image processing, *Science Press*, China.

[10] Liu, Q., Du, H., Xie, Q. Z. (2011). Research on an image fusion algorithm based on the wavelet transform for infrared and visible light images, *Vehicle and Technology*, (3) 48-51.

[11] Gong, Y. X., Yang, W. K., Fan, W. D. (2012). Image fusion based on PCA and Symmetric fractional B-spline wavelet, *Computer Engineering and Applications*, 48 (4).

[12] Singh, R., Khare, A. Fusion of multimodal medical images using Daubechies complex wavelet transform—A multi-resolution approach, *unpublished*.

[13] Duan, P., Xie, K. G., Song, N., Duan, Q. Ch. (2010). A method of vehicle license plate De-noising and location in low light level, *Journal of Networks*, 5 (12) 1393-1400.

[14] Zhao, J. C., Qu, S. R. (2011). A Better Algorithm for Fusion of Infrared and Visible Image Based on Curvelet Transform and Adaptive Pulse Coupled Neural Networks (PCNN), *Journal of Northwestern Polytechnical University*, 29 (6) 849-853.

[15] Lindblad, T., Kinser, J.M. (2008). Image Processing Using Pulse-Coupled Neural Networks, *Higher Education Press*, China.

[16] Huang, W., Jing, Zh. L. (2007). Multi-focus image fusion using pulse coupled neural network, *Pattern Recognition Letters*, 28, p. 1123-1132.

[17] Shi, C., Miao, Q. G., Xu, P. F. A novel algorithm of remote sensing image fusion based on Shearlets and PCNN, *unpublished*.

[18] Qu, X. B., Yan, J. W., Yang, G. D. (2009). Multi-focus image fusion method of sharp frequency localized Contourlet transform domain based on sum-modified-Laplacian, *Optics and Precision Engineering*, 17 (5) 1203-1212.

[19] Li, H. F. (2009). Study on the Method of Image Fusion Based on NonsubsampledContourlet Transform and PCNN, *Chongqing University*.

[20] Yakhdani, M. F., Azizi, A. (2010). Quality Assessment of Image Fusion Techniques for Multi-sensor High Resolution Satellite Images, *In*: ISPRS TC VII Symposium – 100 Years ISPRS, 7, p. 204-209.

[21] Liu, Z., Xue, Z. Y., Zhao, J. Y., Wu, W. (2012). Objective Assessment of Multi-resolution Image fusion Algorithms for Context Enhancement in Night Vision: A Comparative study, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 34 (1) 94-109.

[22] Lu, H., Wu, Q. X., Jiang, C. S. (2007). Color Image Fusion Based on PCA and Wavelet Frame Transform, Computer Stimulation, 24 (9) 202-205, 296.