Image Fusion Algorithm Based on Contourlet Transform and PCNN for Detecting Obstacles in Forests

Zheng Yu, Lei Yan, Ning Han, Jinhao Liu

School of Technology, Beijing Forestry University, Beijing 100083, China
Email: mark_yanlei@bjfu.edu.cn   hn217@bjfu.edu.cn

Abstract: *In this paper the image fusion algorithm based on Contourlet transform and Pulse Coupled Neural Network (PCNN) was proposed to improve the performance of the image fusion in the detection accuracy of obstacles in forests. At the same time, the wavelet transform and the Principal Component Analysis (PCA) were simulated for comparison with the proposed algorithm. Then visible and infrared thermal images were collected in a forest. The experimental results have shown that the fused images using the method proposed provided a better understanding of the reality, enhanced images' clarity and eliminated factors which provided shelters for targets.*

Keywords: *Contourlet transform, image usion, obstacles in forests, pulse coupled neural network.*

1. Introduction

With the technological advances of forestry, the harvester has gradually been applied in the developed countries to replace manual work. The production efficiency, therefore, can be promoted by more than 80 times. However, obstacles in forests, such as stones, clustered trees, animals and humans, will affect the automatic operations of a harvester, cause a dangerous accident and decrease the efficiency. Thus, detecting obstacles to improve the efficiency and safety of the harvester, is of great value.

Image fusion is a branch of data fusion which refers to the acquisition, processing and synergistic combination of the information gathered by various knowledge sources and sensors to provide better understanding of a phenomenon

[1]. The fusion of visible and infrared thermal images is often used to improve situational awareness. When it is dark or misty in the forest, it becomes difficult to distinguish the objects in the visible images, whereas the infrared thermal images that provide information about temperature, are complementary to the visible images. Furthermore, after a period of extensive cooling (e.g., after a long period of rain or early in the morning), the visible images may represent the background in great details (vegetation or soil areas, texture), whereas the infrared images are less detailed due to the low thermal contrast in the scene. In these situations, 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 imagery on a single display may then allow both the detection and the unambiguous localization of the target with respect to the context [2].

The fusion rules have been well improved during the past 10 years. Therefore, image fusion has made inroads into many fields and deeply affected the lives of people. G y a o u r o v a et al. [3] use image fusion to recognize faces and obtain a robust performance. A fusion system to enhance the contrast of medical images has been developed by H e [4]. T s a g a r i s, A n a s t a s s o p o u l o s [5] has discriminated objects in the dark by fusing night-time infrared images with visible images. W a n g et al. [6] has used fused images to locate license plates. Z h a o et al. [7] has provided clear information of the local battlefield for the chain of command using fused images. This study provided an approach for detecting obstacles in forests using image fusion based on Contourlet transform and Pulse Coupled Neural Network (PCNN).

2. Methods of image fusion

2.1. Image capture

Visible and infrared thermal images were captured in a forest with stones, humans, and trees. The images were taken before sunrise, at midday and after sunset, during which different conditions of the forest were presented. Before sunrise, most of the objects in the scene have similar temperature after having lost their excess heat by the radiation during the night. At midday, most factors which affect the fusion can be found at the same time. After sunset, the light is not enough to support the recognition of the objects.

2.2. Image fusion

To further improve the capabilities of discrimination, the process of fusion was implemented in three steps. First, the source images were decomposed by the Contourlet transform. At this step, the images were decomposed into low-pass sub-band and high-pass sub-bands. Then, the coefficients of decomposition were fused by the rule of PCNN, during which the decomposed coefficient was determined by PCNN rule. Finally, the fused images were reconstructed by Contourlet inverse transform. At this step the determined coefficient was transformed into a pixel value of an image. Hence, we arrange the procedure of image fusion as shown in Fig. 1.
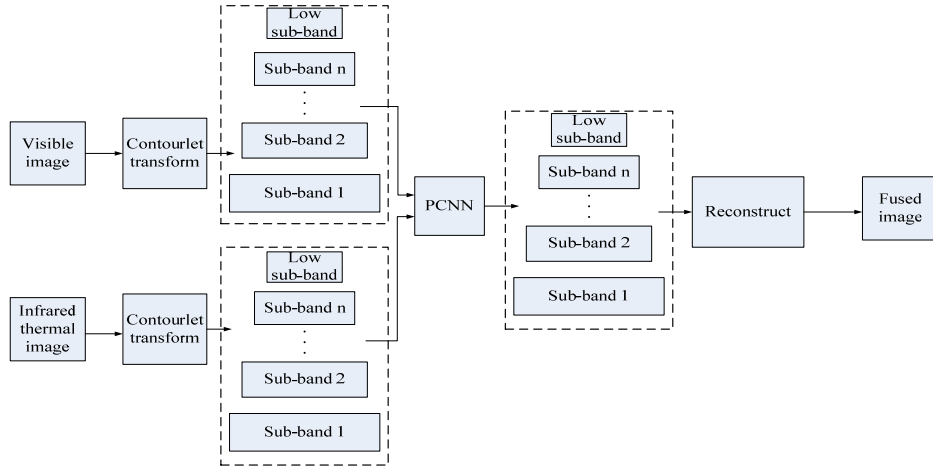
Fig. 1. Procedure for the fusion of visible and infrared thermal images

2.2.1. Contourlet transform

Contourlet transform is first proposed by D o and V e t t e r l i in 2002 [8]. Compared with wavelet transform, it deals efficiently with piecewise smooth images having smooth contours. The Contourlet transform decomposes images into low-pass sub-bands and high-pass sub-bands. Low-pass sub-band is the so-called approximate image. High-pass sub-band reflects contours of targets in detail.

A double filter bank structure is constructed, in which at first the Laplacian Pyramid (LP) is used to capture the point discontinuities, followed by a Directional Filter Bank (DFB) to link point discontinuities into linear structures [9]. Therefore, the image can be decomposed into sub-band images in multi-scale and multi-direction, which provide the detailed information of original images, especially information about the contours of the targets.

2.2.2. Fusion rules

To preserve the information as far as possible, algorithms, such as wavelet transform, PCA and PCNN have been widely used in the image fusion. L u et al. [10] used PCA and wavelet transform to fuse images and obtained a clearer image. F a n et al. [11] proposed an improved algorithm based on wavelet decomposition which enhanced the fusion efficiency and improved fusion accuracy. Z h a o and Q u [12] fused a visible image and an infrared image using PCNN and obtained an image of higher contrast ratio and clarity. This study fused images with these mentioned rules and obtained the best result using PCNN. The theory of PCNN is to decide which pixel at the position of (*i, j*) (a visible image or an 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.

Based on the research of the visual cortex neuron of mammals, Eckhorn firstly proposed a pulse coupled neural network (see [13]). Fig. 2 shows the neuron model of PCNN.
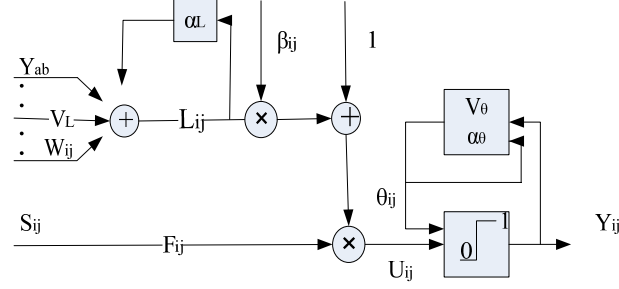
118

Fig. 2. Neuron model of PCNN

In Fig. 2: 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; F_{ij} is the feeding input; β is the linking strength; 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 the time constant; Y_{ij} is the output of the neuron; n is the number of cycles. The rule can be conducted as follows [14]:

(1)
$$F_{i,j}^k(n) = S_{i,j}^k(n),$$

(2)
$$L_{i,j}^k(n) = \exp(-\alpha_L)L_{i,j}^k(n-1) + V_L\sum_{a,b}\omega_{ij,ab}Y_{ab}(n-1),$$

(3)
$$U_{i,j}^k = F_{i,j}^k(n)\times(1+\beta L_{i,j}^k(n)),$$

(4)
$$\theta_{i,j}^k = \exp(-\alpha_\theta)\theta_{i,j}^k(n-1) + V_\theta \times Y_{i,j}^k(n),$$

(5)
$$Y_{i,j}^k = \text{step}(U_{i,j}^k(n) - \theta_{i,j}^k(n)).$$

Each decomposition coefficient can be regarded as a neuron and the neuron is connected with 8 neurons nearby as Fig. 3 shows.
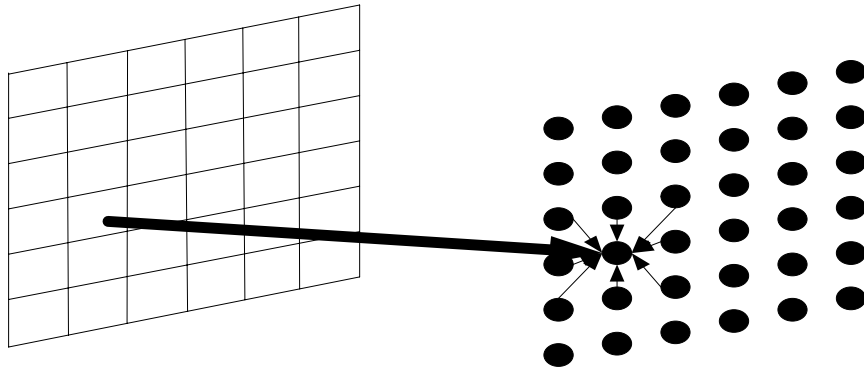

Fig. 3. Diagram of image processing using PCNN ($\square$ denotes a pixel, $\bullet$ denotes a neuron)

The k-th decomposed sub-band in the Contourlet coefficients is denoted by k, and (i, j) indicates the pixels' position in the sub-image. It is conducive to capture the information produced by the neurons nearby. In this model v_θ is a constant

which is much larger than the other parameters. When neuron A outputs a pulsing signal, the value of Y_A turns to 1, then θ_A will increase rapidly, based on formulas (4) and (5). Thus, the neuron will not produce a pulsing signal in the next loop. At that moment, as formulas (2) and (3) show, the neurons B next to neuron A will capture the signal, and U_B increase so that output their signal (Y_B turns to 1) ahead of time. In this way the pulse transmits between neurons and the pulses a neuron outputs are seriously related to the neuron's capacity to describe the feature that is associated to the input parameter. Therefore, after N loops, the more pulsing signals a neuron produces, the better feature (compared with what inputs are described) the neuron describes. By analyzing this, we could obtain the feature in an image.

Based on the experiments, we set α_L =0.2, α_θ =0.2, V_L =1.5, V_θ =25, β =0.2, N=2500, $\omega_{ij} = \begin{vmatrix} 0.707 & 1 & 0.707 \\ 1 & 0 & 1 \\ 0.707 & 1 & 0.707 \end{vmatrix}$. Sum-Modified-Laplacian (SML) that measures the clarity of images can be defined as follows [15]:

(6)
$$ML(i,j) = |2I(i,j) - I(i\pm1,j)| + |2I(i,j) - I(i,j\pm1)| + \\ + |2I(i,j) - I(i\pm1,j\pm1)| + |2I(i,j) - I(i\pm1,j\mp1)|,$$

(7)
$$SML(i,j) = \sum_m \sum_n \omega(m,n)[ML(i+m, j+n)]^2,$$

where $I(i, j)$ is the decomposition coefficient at the position of (i, j). To describe the areas which did not change acutely, we set $\omega(m,n) = \dfrac{1}{15}\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix}$.

In low-pass regions, we select SML as an input of PCNN, which makes the network's output related to clarity. The more pulsing signals a neuron produces, the clearer the neuron presents. Therefore, by comparing the times of pulsing signals the corresponding neurons created, we can identify which pixel contains the most abundant information and which one is the clearest one. And in high-pass regions, the coefficients produced by Contourlet transform are adopted as an input to the network. The more pulsing signals a neuron produces, the stronger contrast ratio the neuron presents [16].

After 300 loops, we record the number of pulsing signals a neuron produces as $T^k(i, j)$, where k denotes a visible image or an infrared thermal image, and (i, j) indicate the position of (i, j). Thus, the fusion rule can be written as:

(8)
$$\begin{cases} I(i,j) = I^A(i,j) & \text{if} \quad T^A(i,j) > T^B(i,j), \\ I(i,j) = I^B(i,j) & \text{if} \quad T^B(i,j) \geq T^A(i,j), \end{cases}$$

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

3. Experimental results and discussion

The images used in this paper are collected in Beijing Olympic Forest Park by a Fluke Ti 55 infrared thermal camera.
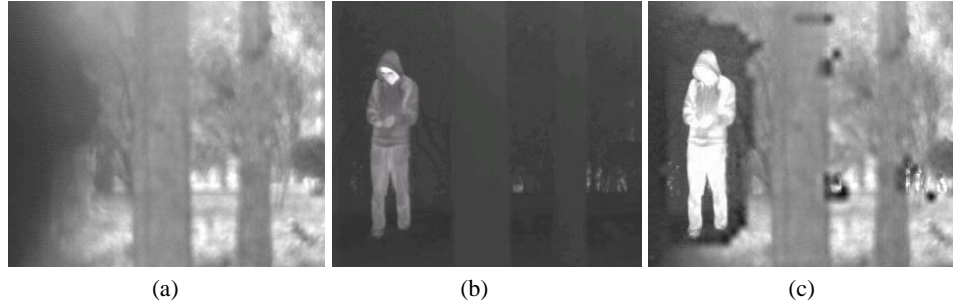

(a)   (b)   (c)

Fig. 4. Example of detecting hidden obstacles: A visible image (a); an infrared thermal image (b); a fused image (c)
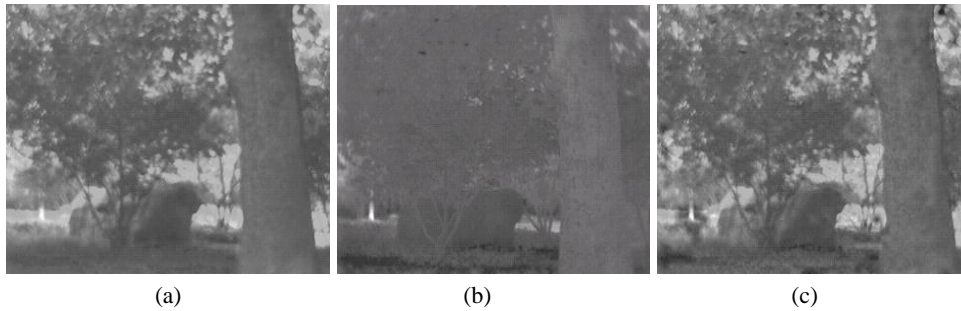

(a)   (b)   (c)

Fig. 5. Example of enhancing clarity: A visible image (a); an infrared thermal image (b); a fused image (c)

Figs 4 and 5 are results of image fusion. In Fig. 4a the person could not be seen clearly in the visible image, whereas the infrared thermal image having temperature information was more effective to distinguishing the person as Fig. 4b shows. In addition, the contrast ratio of Fig. 4b was obviously weaker than that of Fig. 4a, it was difficult to tell the trees from the background in Fig. 4b. Thus, the fused image as shown in Fig. 4c captured the advantages of both visible image and infrared thermal image. In Fig. 5b, because of the low thermal contrast, it was hard to distinguish the stones and the clustered trees. Nevertheless, Fig. 5a provided a complement to recognize the targets. In addition, compared with Fig. 5a, the fused image as shown in Fig. 5c performed better in the contrast ratio and clarity.

The results suggest 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.

Three criteria were applied for the quantitative assessment of the fusion performance.

By showing the average information included in the image and reflecting the detailed information of the fused image, entropy is a measure that directly concludes the performance of image fusion. Thus, the greater the entropy of the

121

fused image is, the more abundant information the fused image provides, and the greater the quality of the fusion is [17, 18]. The entropy of a fused image was calculated according to

$$(9) \qquad E = -\sum_{i=0}^{255} P_i \log_2 P_i,$$

where E is the entropy of an image, and P_i is the probability of i in the image, i means value (which is a number between 0 up to 255).

Then we used the average gradient to measure the performance of the fused image. It is defined as follows [10]:

$$(10) \quad \overline{\text{grad}} = \frac{1}{(l-1)(c-1)} \sum_{m-1} \sum_{n-1} \sqrt{\frac{\left(X(i,j) - X(i+1,j)\right)^2 + \left(X(i,j) - X(i,j+1)\right)^2}{2}},$$

where $\overline{\text{grad}}$ is the average gradient of an image, and $X(i, j)$ is the pixel value of the image at position (i, j); l, c decide the area (the biggest time is 3×3). The average gradient measures the clarity of the image. A larger average gradient means a higher spatial resolution.

Standard deviation measures the degree that the data points deviate from the mean. The **st**andard **d**eviation (std) was calculated according to

$$(11) \qquad \text{std} = \sqrt{\frac{\sum (X(i,j) - \overline{X})^2}{(t-1)}},$$

where std is the standard deviation of an image, and $X(i, j)$ is the pixel value of the image at the position of (i, j); $\overline{X}$ is the average pixel value; t is the number of the pixel of an image. A larger standard deviation means a better visible effect.

Table 1. Quality assessment of the image fusion

Figure	Entropy			Average gradient			Standard deviation		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Fig. 4	7.35	5.02	7.45	8.29	1.57	9.56	43.81	18.31	49.90
Fig. 5	6.45	5.53	6.52	5.34	3.00	6.11	27.57	12.77	28.35

Table 1 suggests that fused images captured more abundant information, had a higher spatial resolution and showed a better visible effect, compared with single original images.

Fig. 6 shows results of different fusion rules, including a wavelet, PCA and PCNN. Wavelet performed worst in capturing curve edges. In addition, compared with PCA, PCNN performed better in describing the detail information. For example, PCNN better represented the people standing on the right of the image. Table 2 shows the quality assessment for different fusion rules.
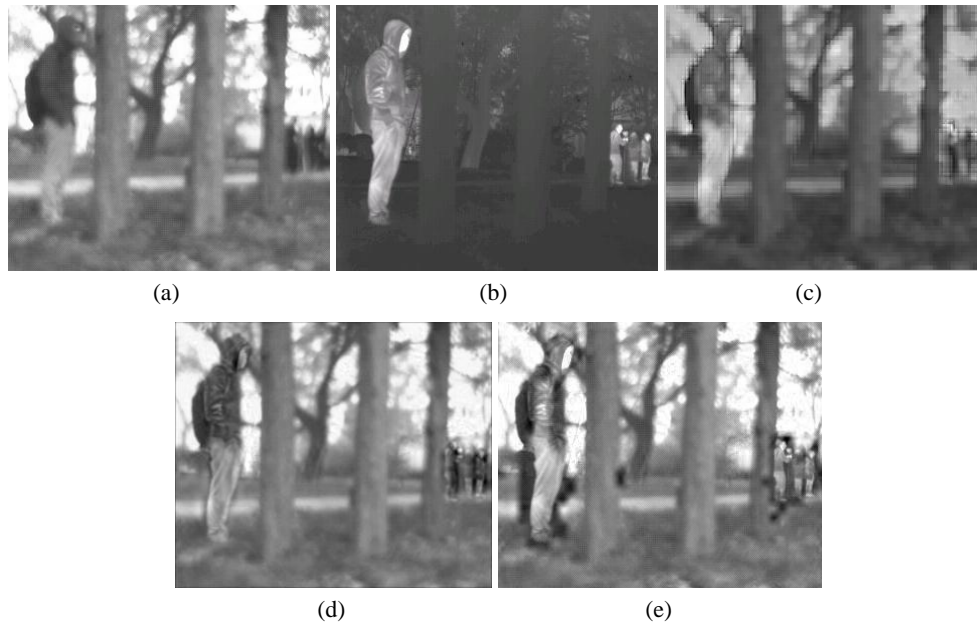
(a) (b) (c)

(d) (e)

Fig. 6. Performance of image fusion using different fusion rules: A visible image (a); an infrared thermal image (b); a fused image using wavelet (c); a fused image using PCA (d); a fused image using PCNN (e)

Fig. 6 shows results of different fusion rules, including a wavelet, PCA and PCNN. Wavelet performed worst in capturing curve edges. In addition, compared with PCA, PCNN performed better in describing the detail information. For example, PCNN better represented the people standing on the right of the image. Table 2 shows the quality assessment for different fusion rules.

Table 2. Quality assessment of different fusion rules

Fusion rule	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

Table 2 suggests that compared with the wavelet and PCA, PCNN yielded more abundant information and higher clarity, and therefore performed better in depicting a complicated situation. Image fusion based on PCNN provided better understanding of the reality in forests under any circumstances.

4. Conclusion

In this paper image fusion based on Contourlet and PCNN was proposed to detect obstacles in forests. The fusion is divided into three steps. At the first step, the original images are decomposed into low-pass sub-band and high-pass sub-bands

by a Contourlet transform. Then coefficients of decomposition are fused by the rule of PCNN. Finally, the fused image is reconstructed by Contourlet inverse transform. Compared with the rule of wavelet transform and PCA, the proposed algorithm performs better in capturing the targets in images. 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. In addition, the further fusion of fused images and other valid data, such as distance data and temperature data, might provide the reality precisely to guarantee the safety and efficiency, and facilitate the production activities of a harvester in forests.

References

1. V a r s h n e y, P. K. Multisensor Data Fusion. – Electronics & Communication Engineering Journal, Vol. **9**, 1997, No 6, pp. 245-253.
2. T o e t, A., J. K. I j s p e e r t, A. M. W a x m a n, M. A g u i l a r. Fusion of Visible and Thermal Imagery Improves Situational Awareness. – Displays, Vol. **18**, 1997, No 2, pp. 85-95.
3. G y a o u r o v a, A., G. B e b i s, I. P a v l i d i s. Fusion of Infrared and Visible Images for Face Recognition. – In: Computer Vision-ECCV2004, Vol. **3024**, 2004, pp. 456-468.
4. H e, K a i. Study on Pixel-Level Medical Image Fusion. Northwestern Polytechnical University, 2006.
5. T s a g a r i s, V., V. A n a s t a s s o p o u l o s. Fusion of Visible and Infrared Imagery for Night Color Vision. – Displays, Vol. **26**, 2005, No 4-5, pp. 191-196.
6. W a n g, R a n, G. W a n g, J. L i u, Z. C h e n, Y u S h i. A Novel and Fast Method for Precise License Plate Location from Digital Images Based on Fusion of Salient Features. – International Journal of Digital Content Technology and its Applications, Vol. **6**, 2012, No 19, pp. 511-519.
7. Z h a o, F u l i, W. D i n g, G. W a n g, H. L i. Experimental Tests of Image Fusion Method for Unmanned Aerial Vehicle. – International Journal of Advancements in Computing Technology, Vol. **4**, 2012, No 16, pp. 420-427.
8. D o, M. N., M. V e t t e r l i. Contourlets: A Directional Multiresolution Image Representation. – In: Proc. of 2002 International Conference on Image Processing, Vol. **1**, 2002, pp. 357-360.
9. L i n, L i y u. Contourlet Transform. – In: Image Processing. Science Press, China, 2008.
10. L u, H u a n, Q. W u, C. J i a n g. Color Image Fusion Based on PCA and Wavelet Frame Transform. – Computer Stimulation, Vol. **24**, 2007, No 9, pp. 202-205, p. 296.
11. F a n, L i, Y. Z h a n g, Z. Z h o u, D. P. S e m a n e k, S. W a n g, L. W u. An Improved Image Fusion Algorithm Based on Wavelet Decomposition. – Journal of Convergence Information Technology, Vol. **5**, 2010, No 10, pp. 15-21.
12. Z h a o, J., S. Q u. 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, Vol. **29**, 2011, No 6, pp. 849-853.
13. L i n d b l a d, T., J. M K i n s e r. Image Processing Using Pulse-Coupled Neural Networks. Higher Education Press, China, 2008.
14. M a, Y., Y. Z h a i, P. G e n g, P. Y a n. A Novel Algorithm of Image Fusion Based on PCNN and Shearlet. – International Journal of Digital Content Technology and Its Applications, Vol. **5**, 2011, No 12, pp. 347-354.

15. Q u, X., J. Y a n, G. Y a n g. Multifocus Image Fusion Method of Sharp Frequency Localized Contourlet Transform Domain Based on Sum-Modified-Laplacian. – Optics and Precision Engineering, Vol. **17**, 2009, No 5, pp. 1203-1212.
16. L i, H. Study on the Method of Image Fusion Based on Nonsubsampled Contourlet Transform and PCNN. Chongqing University, 2009.
17. Y a k h d a n i, M. F., A. A z i z i. Quality Assessment of Image Fusion Techniques for Multisensor High Resolution Satellite Images (Case Study: IRS-P5 and IRS-P6 Satellite Image). – In: ISPRS TC VII Symposium – 100 Years ISPRS, Vol. **7**, 2010, pp. 204-209.
18. L i u, Z., E. B l a s c h, Z. X u e, J. Z h a o, R. L a g a n i e r e, W. W u. Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study. – IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. **34**, 2012, No 1, pp. 94-109.