Research on Target Type Recognition Algorithm of Aerial Infrared Image

Bin Liu and Yangyu Fan
Institute of Electronic and Information / Northwestern Polytechnical University, Xi'an, China
Email: liubin2908@126.com, fan_yangyu@sina.com

Jian Guo
Luoyang Optoelectro Technology Development Center, Luoyang, China
Email: guojian014@163.com

***Abstract*—In order to improve the target type recognition rate of aerial infrared image under the new requirements of omnidirectional crossing and multitarget types, a recognition algorithm which has four steps is researched in this paper. Firstly the maximum between-cluster variance (Ostu) algorithm is applied to segment target from the infrared image. Secondly a new edge detection algorithm is proposed to get the target edge in the segmentation image. Thirdly the edge points are fitted to be a polygon and its Fourier descriptor is extracted to obtain the target feature. Finally the target type is identified by a classification recognition algorithm based on the BP neural network and the Fourier descriptor of target. The simulation results show that the recognition rate of the target type recognition algorithm is more than 80% for the typical aerial target on every complex condition, and the processing time is just 0.1s. So the researched algorithm can meet the requirements of real-time performance and high recognition rate.**

***Index Terms*—Target Recognition; Image Segmentation; Edge Detection; Fourier Descriptors; Neural Network;**

I. INTRODUCTION

At the end of the trajectory, for the high speed moving target, the traditional lateral detection optical fuze which is a part of infrared imaging guidance air-to-air or surface-to-air missile can't timely calculate detonation delay time and control detonation. A possible solution is that the missile uses its hardware resources of imaging guidance detection system to realize the forward looking detection and recognition, and it is also called as guidance integrated fuze (GIF) technology [1, 2]. The aim of imaging guidance detection system is only to track target, so it just needs obtain the target center coordinate in the infrared image and doesn't need recognize the target type [3]. So the image processing algorithms which used in imaging guidance detection system are not suitable for application of forward looking detection and recognition [4]. In this paper, a target type recognition algorithm is proposed to meet its new requirements of real-time performance and high recognition rate.

The proposed algorithm has four steps. Firstly the Ostu algorithm is applied to segment target from the infrared image. The infrared image reflects the temperature distribution of target and its background. In general, the infrared radiation intensity of target is more than the infrared radiation intensity of background, especially high speed moving target. Besides the temperature of target is higher than the temperature of background because the friction is occurred for high moving target. So the gray value of target is more than the gray value of background in infrared image and the image segmentation algorithm is feasible to separate target from background. By comparison and analysis many image segmentation algorithms, such as iterative segmentation algorithm, maximum entropy algorithm and simple statistics algorithm, the Ostu algorithm is selected because its process speed is very fast. Secondly a new edge detection algorithm is proposed to get the target edge in the segmentation image. The proposed edge detection algorithm has high detection accuracy, and its processing speed is faster than other edge detection algorithms. Thirdly the edge points are fitted to be a polygon and its Fourier descriptor is extracted to obtain the target feature. The edge points contain the shape information of target, and the information could be used to recognize the target type. The Fourier descriptor extraction algorithm can get the target type information because it has translation invariance, rotation invariance and scale invariance. Finally the target type is identified by a classification recognition algorithm based on the BP neural network and the Fourier descriptor of target. The BP neural network has a lot of advantages, such as distributed parallel fast processing, nonlinear mapping and robustness. The simulation indicates the researched algorithm can meet the requirements of real-time performance and high recognition rate.

II. IMAGE SEGMENTATION AND EDGE DETECTION ALGORITHM

A. Image Threshold Segmentation Algorithm

Image threshold segmentation algorithm is a very important method of image segmentation. The key of threshold segmentation is to find optimal segmentation threshold. The image can be divided into target and background by optimal segmentation threshold.

doi:10.4304/jmm.8.5.468-474

According to the characteristics of grayscale image, the Otsu algorithm [5, 6] is adopted to segment image after the simulation and comparison for the various algorithms. The algorithm is shown as follows.

Assuming the grayscale value of an image is 0 to $I-1$, the probability of the pixel grayscale value i is p_i, gray value is divided into two groups $C_0=\{0,1,\cdots,t\}$, $C_1=\{t+1,t+2,\cdots,I\text{-}1\}$ by integer t, then the probability of C_0 is

$$\omega_0=\sum_{i=0}^{t}p_i=\omega(t) \quad (1)$$

the average value of C_0 is

$$\mu_0=\frac{\sum_{i=0}^{I-1}ip_i}{\omega_0}=\frac{\mu(t)}{\omega(t)} \quad (2)$$

and the probability of C_1 is

$$\omega_1=\sum_{i=t+1}^{I-1}p_i=1-\omega(t) \quad (3)$$

the average value of C_1 is

$$\mu_1=\frac{\sum_{i=t+1}^{I-1}ip_i}{\omega_1}=\frac{\mu-\mu(t)}{1-\omega(t)} \quad (4)$$

where μ is the statistical mean of the overall image gray, it can be written as

$$\mu=\sum_{i=0}^{I-1}ip_i=\omega_0\mu_0+\omega_1\mu_1 \quad (5)$$

Then the variance between the two groups is

$$\begin{aligned}\sigma^2(t)&=\omega_0(\mu_0-\mu)^2+\omega_1(\mu_1-\mu)^2\\&=\omega_0\omega_1(\mu_1-\mu_0)^2\end{aligned} \quad (6)$$

Changing t from 0 to l-1, t is optimal threshold value when the variance is the maximum.

$$t=Arg\max_{0\le t\le l-1}\sigma^2(t) \quad (7)$$

B. Edge Detection Algorithm

The edge detection algorithm records the coordinate sequence of the binary image by tracking each pixel of the boundary line which constitutes the target closed region [7, 8, 9, 10]. The boundary line tracking which uses the clockwise search method, treats background as a plane, and target as a building, then makes the right-hand contact with the wall of the building and moves forward, finally the right-hand returns to the starting point. There are eight adjacent points of each pixel (x, y) in the digital image, so the point can only extend along the following eight directions: right (0), upper right (1), upper (2), upper left (3), left (4), lower left (5), lower (6), lower right (7), as shown in Figure 1. Assuming in the binary image, the target area is represented by 1, the background area is represented by 0. The array $\mathbf{X}$ is used to store the extracted edge points, assuming X_s is the edge starting point which has been searched out. The algorithm has two main steps, namely the search of the starting point and the search of the target boundary points.

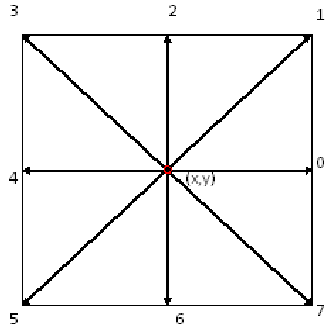
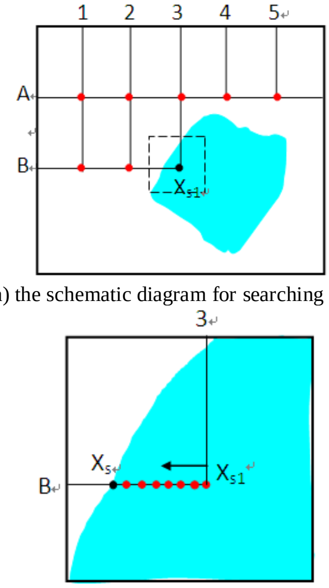


Figure 1. Adjacent points distribution of the pixel

Assuming the shadow part in Figure 2(a) is the target region, the search of the starting point also has two steps. Firstly search one pixel of the target region with long step length from top to bottom and from left to right. As shown in Figure 2(a), we can get (B, 3), the reference location of the starting point X_{s1}, after searching the pixels (A, 1), (A,2), … , (B, 2). Then search the starting point with short step length from right to left. As shown in Figure 2(b) which has been zoomed by the region separated by the closed dash line in Figure 2(a), the starting point X_s can be found.



(a) the schematic diagram for searching Xs1

(b) the schematic diagram for searching Xs

Figure 2. The schematic diagram for the starting point

Before descript the search of the target boundary points, some rules is defined as below: 2 represents a processing background point, and its value before processing is 0; 3 represents a processing target region point, and its value before processing is 1; 4 represents a processed target region point, namely the searched boundary pixel, and its original value is 1. The changes of value 0→2, 2→0, 1→3, 3→4 are legitimate during the process of search, the rest of changes aren't legitimate.

As shown in figure 3, the search of the boundary points is shown as follows: (1) Set the starting point A with value 1 as the current point and search a background point B in the eight adjacent point of A. Then change the value of A to 1 and the value of B to 2. (2) Search a point C with value 1 by clockwise reading the eight adjacent point values of A from B. Then change value of A to 4 and put its coordinates into X, and change value of B to 0 and set C as the current point A. (3) return to (1) until the search point is the starting point again. Finally X will record all boundary point coordinates of target.

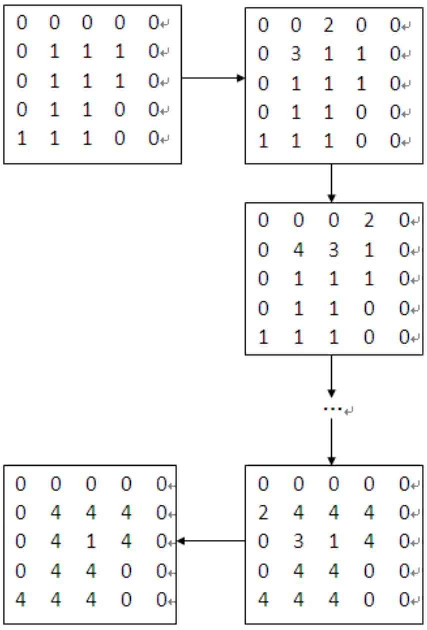


Figure 3. The schematic diagram for searching edge

III. Feature Extraction and Classification Recognition Algorithm

There are many image recognition methods based on shape characteristic, such as basing on boundary characteristic points, invariant moment, Fourier descriptor, autogressive model and so on. The analysis results indicate that the Fourier descriptor based on the coordinate sequence information of the object contour has the optimal function of shape recognition. Target infrared grayscale image obtained at the end of ballistic has larger translation, rotation and scaling during the changes of imaging distance, orientation, location and other factors. Therefore, features extraction should have invariability for translation and rotation, and the Fourier descriptor [11, 12] just has above advantages.

The time for the signal processing is very short because the imaging GIF works at the end of ballistic. The target feature extraction and recognition algorithm must be finished within a very short frame interval. If taking the target contour points as the shape characteristics of the target recognition and extracting Fourier descriptor directly, the total calculation of the feature extraction will be very large and the process time will be very long. So polygon is used to contour fitting, it takes polygon vertices as an approximation of the original contour data to describe the shape of the object. The block diagram of target feature extraction and type recognition algorithm is shown in Figure 4.

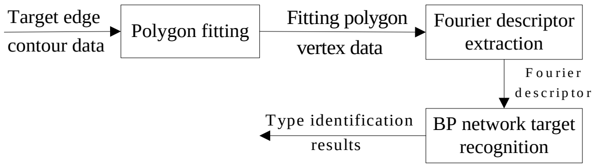


Figure 4. Target type recognition block diagram

A. Target Feature Extraction Algorithm

1) Polygonal Representation of the Contour Shape

After completing the image segmentation and edge detection, an ordered set of boundary point representing complete contour of the target is obtained. But the amount of this point set is too large, and the quantity of the contour data is usually not the same for different targets and it may be not a power of 2. So the fast Fourier transform algorithm can not be used directly to get the Fourier descriptor. In order to improve the calculation speed of the Fourier descriptor, the quantity of contour data can be reduced by the polygon approximation method. Then the vertex data of approximation polygon are taken Fourier transform instead of original contour data. By this way, the storage of the shape boundary information and the computation of Fourier transform can be greatly reduced because the vertices of polygonal approximation are less than the points of the original digital curve. For the target contour data, in the given approximation accuracy (such as 2 to 3 pixels), the extracted polygon can approximately instead the original contour curve after extracting the contour data. This algorithm not only greatly reduces the computation of the Fourier transform, but also retains the shape feature. Besides, it can eliminate the impact of the contour curves noise to some extent.

2) The Fourier Descriptor Feature Extraction

The Fourier descriptor is the Fourier transform coefficient of the shape edge curve, and it is the result of frequency domain analysis to the edge curve. Besides the Fourier descriptor is based on the global information of the object contour data and associated with the scale of the shape, direction and the starting point of contour position. The Fourier descriptors can be not sensitive to the starting point of the contour data by normalizing them, and the ability of the target recognition will be greatly improved with them.

The normalized Fourier descriptor characteristic of target image can be obtained by following method.

1) $(x[k], y[k])$ is a group of ordered set of points which can clearly describe the target edge curve, one-dimensional sequence is obtained with describing the $(x[k], y[k])$ to complex plane, it can be written as

$$f(k) = x(k) + jy(k) \tag{8}$$

where $k = 0, 1, \cdots, N-1$, N is the number of the closed boundary points.

2) Takes Fourier transform of $f(k)$,

$$F(\omega) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\exp[-j2\pi k\omega / N] \tag{9}$$

where $\omega = 0, 1, \cdots, N-1$, $F(w)$ is the Fourier description of the boundary.

3) $F(0)$ is the DC component, it only represents the center position of the image. $F(0)$ should be removed in order to obtain the invariance of the image.

4) Make modulo operator of $F(\omega)$ and retain previous 12 transformed coefficients, then they are divided by $F(1)$ to normalize them. Thus, the feature vectors required by the target type recognition algorithm of neural network are constituted, namely the 12-dimensional Fourier descriptors vectors $s(k)$, $k = 1, 2, \cdots, 12$.

The low-frequency components of the Fourier coefficient are mainly determined by the overall contour of the image. The high frequency components which are not useful enough to recognize the target just portray the detail of the shape and easily degenerate into noise. Therefore, the boundary shape can be described approximately by low-frequency components of the Fourier coefficient which achieve the significant purpose of data compression.

B.Target Classification Recognition Algorithm

1) The Choice of the Neural Network Classifier

The back propagation (BP) neural network and the radial basis function (RBF) neural network are widely used in the mode recognition. The simulation comparison indicates the convergence speed of the RBF network is quicker, the required network training time is shorter, and the classification capability of the network is better. But it is a large-scale network and it needs so many radial basis neurons when the input space is large. Because all the parameters of the BP neural network need to be adjusted in the training process, the convergence speed is slow. But its network size is small and its recognition processing time is very short. Because the requirement of the speed of the recognition algorithm is very fast, the BP neural network is more suitable than the RBF neural network [13, 14, 15, 16].

2) Target Type Classification Recognition Algorithm

The three layers feed forward network with a single hidden layer is used in building BP neural network. The dimension of Fourier feature vector is 12, so the number of network input nodes is designed to be 12. Three types of the targets need to be identified, the number of the output nodes is designed to be 3. The number of hidden layer nerve cell is set to be 12, and the logsig function is adopted for the hidden layer nerve cell. The output layer uses linear sigmoid function, and its output is the linear weighted sum of the hidden layers. The logsig function can be written as

$$f(x) = 1/(1+e^{-x}) \tag{10}$$

IV. SIMULATION

A. Simulation and Analysis

The computer is configured to P4 3.0G CPU and 1.5G bit memory for simulation, and programming tool is Matlab.

The original infrared image of the real aircraft whose size is 128×128 is shown in Figure 5, while the histogram of original image, the image segmentation result and the edge detection result are shown in Figure 6～Figure 8 respectively.

Figure 5. Original image

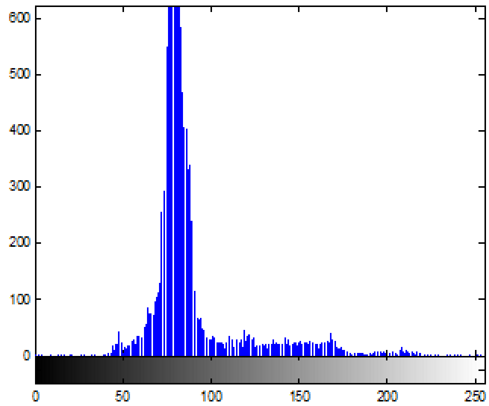


Figure 6. Histogram of original image

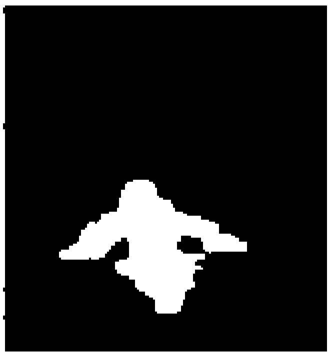


Figure 7. Threshold segmentation result

The target and background are at a fuzzy area in Figure 5 and its optimal segmentation threshold infrared image is nearby 100 in Figure 6. The segmentation threshold t of this image calculated by the segmentation algorithm is

112, so it's consistent with the observations of the histogram. After the segmentation, the target and the background have been completely separated, and the target can be clearly described and no noises in the Figure 7. The edge detection result of the segmentation is very clearly and the aircraft shape can be accurately depicted in Figure 8. The processing time of the proposed edge detection algorithm is 0.031s, and the processing time of the Sobel is 0.211s. So the proposed edge detection algorithm is much faster than the Sobel.

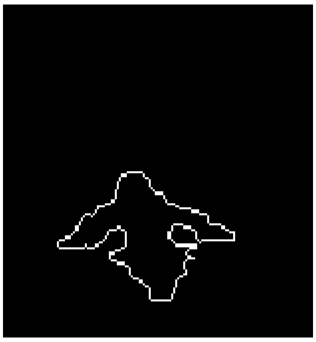


Figure 8. Edge detection result

B. Simulation of Target Feature Extraction Algorithm

Figure 9 is the image of original edge contour, and Figure 10 is the reconstructed image by taking the previous 12 Fourier transformed coefficients. It can be seen that the aircraft curve in the reconstructed image has a little deformation and it can replace the aircraft curve in the original edge curve to descript the aircraft information.

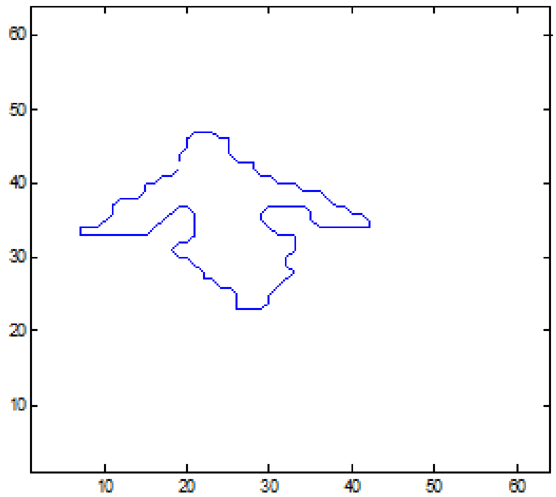


Figure 9. Image of original edge contour

C. Training and Testing of Classification Recognition Algorithm

In order to build a BP neural network, the Fourier descriptors of typical target are taken as input vector to training network. Collect 80 infrared gray images of three types of targets (cruise missile, helicopter, fighter plane) with different attitudes, different locations, different proportions, where 50 infrared images are used for the neural network training and the rest for neural network testing.

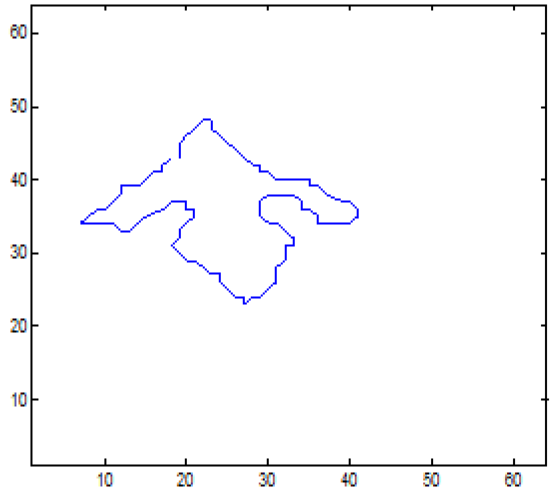


Figure 10. Reconstructed image by taking the first 12 items

The simulation results of training skeleton map of typical target are shown in Figure 11, Figure 13 and Figure 15 respectively. The simulation results of test skeleton map of typical target are shown in Figure 12, Figure 14 and Figure 16 respectively.

The scaled conjugate gradient reverse propagation algorithm is used in BP neural network training. By adjusting the weight vector, the width and the connection weight coefficient between the hidden layer and the output layer continuously, it enables the sum squared error of network falls below the target error. Once the output sum squared error of network is less than 0.001, stop training.

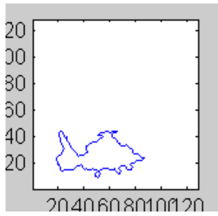


Figure 11. Training skeleton map of helicopter

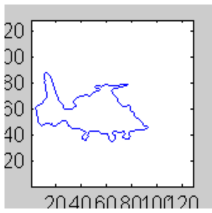


Figure 12. Test skeleton map of helicopter

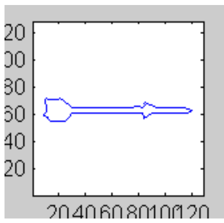


Figure 13. Training skeleton map of cruise missile

It can be seen from Figure 17, the performance of network satisfies requirement when the training achieves the 77th step. After training, the Fourier descriptors of a test set of typical target are taken as neural network input to test its recognition rate. The simulation results show that the recognition rate is more than 80% by the BP neural network. Besides, the RBF neural network is also built and its recognition rate is 70%. So the recognition rate of the BP neural network is higher than the RBF neural network.

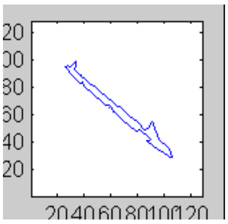


Figure 14. Test skeleton map of cruise missile

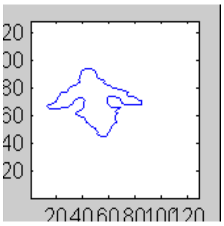


Figure 15. Training skeleton map of flight plane

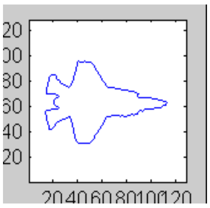


Figure 16. Test skeleton map of flight plane

Finally, the processing time of the target type recognition algorithm is monitored and its value is 0.1s. The time will reduce one order of magnitude if the proposed algorithm is realized by FPGA.

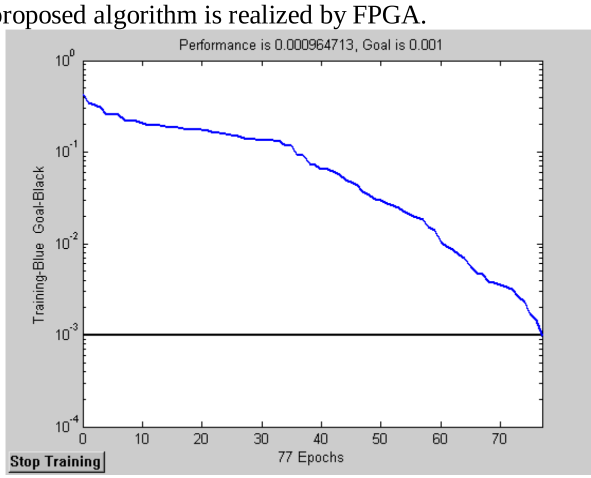


Figure 17. Training error curve

V. CONCLUSIONS

According to the characteristics of the aerial infrared target image, a group of image segmentation, edge detection, feature extraction, type recognition algorithms are put forward in this paper after analysis and comparison of many algorithms. The simulation results indicate that the set of algorithms for typical infrared targets have higher ability of target type classification identification, and its real-time performance meets the requirement of GIF Fuze.

REFERENCES

[1] Z. Yuan, Z. J. Sun. The design of air to air missile fuze warhead system. *Beijing: National Defense Industry Press*, 2007.

[2] B. Liu, Y. Y. Fan. “GIF technology and its typical application,” *Aero Weaponry*, vol. 3, pp. 33–36, 2006.

[3] Z. J. Liu, Q. Li, Z. W. Xia, et al. “Target recognition for small samples of ladar range image using classifier ensembles,” *Journal of Optical Engineering*, vol. 51, pp. 087201-1–087201-12, 2012.

[4] B. Liu, Y. Y. Fan, J.W Zhou. “A burst control algorithm based on imaging guidance and laser ranging,” *Journal of Detection & Control*, vol. 34, pp. 57–60, 2012.

[5] Z. S. Dong, F.C Zhang. “Furnace flame image segmentation modeling based on OSTU,” *Journal of Microcomputer Information*, vol. 26, pp. 201–203, 2010.

[6] L. Q. Zhu, W. Z. Chen, R. Xing, et al. “A robust method to extract features from palm image,” *Journal of Zhejiang University (Engineering Science)*, vol. 43, pp. 1412–1418, 2009.

[7] X. Qiu, J. Huang, F. Yang, et al. “Image enhancement based media-adventitia border detection in intravascular ultrasound images,” *Journal of Image and Graphics*, vol. 17, pp. 537–545, 2012.

[8] J. Wu, G. M. Zhang, J. Xia, et al. “Gray cerebrovasular image skeleton extraction algorithm using level set model,” *Journal of Multimedia*, vol. 5, pp. 208–215, 2010.

[9] C.W. Lu, Z. B. Hu. “A fuzzy search algorithm for structured P2P network based on multi-dimensional semantic matrix,” *Journal of Networks*, vol. 7, pp. 377–384, 2012.

[10] S. C. Zhang, T. Y. Liu. “Research on IR targets extraction and tracking under low altitude background,” *Journal of Laser & Infrared*, vol. 40, pp. 546–548, 2010.

[11] L. L. Zhang, S. X. Liu, J. J. Wu, et al. “Wavelet extraction of GPR signals based on fractional fourier transform,” *Journal of Jilin University(Earth Science Edition)*, vol. 42, pp. 569–574, 2012.

[12] C. Y. Li, Q. X. Cao. “Extraction method of shape feature for vegetables based on depth image,” *Transactions of the Chinese Society for Agricultural Machinery*, vol. 43, pp. 242–245, 2012.

[13] L. R. Li, C. L. Shen, C. Q. Wang. “Recognition method of aerial targets based on BP neural network,” *Journal of Control & Command Control*, vol. 37, pp. 132–139, 2012.

[14] J. H. Qu, S. B. Qu, Z. J Wang. “Feature based fuzzy neural network approach for target classification and recognition in remote sensing images,” *Journal of Remote Sensing*, vol. 13, pp. 67–74, 2009.

[15] H. H. Sun, L. Yao, Q. H. Zhang, H. L. Zhang. “Research on images identification method based on neural network,” *Optical Technique*, vol. 34, pp. 173–174, 2008.

[16] G. Xie, T. R. Cao, C. D. Yan. "Texture features extraction of chest HRCT images based on granular computing," *Journal of Multimedia*, vol. 5, pp. 639–674, 2010.

Bin Liu is a doctoral student in the Institute of Electronic and Information of Northwestern Polytechnical University. He was born in 1963. His research interests include Signal Processing, imaging GIF fuze.

Yangyu Fan received his Ph.D. in the School of Marine Engineering from Northwestern Polytechnical University in December 1999.He was born in 1960. He is currently a Professor at the School of Electronics and Information of the University and the head of Multimedia & Virtual Reality Laboratory, the Institute of Signal Processing & Wireless optical Communication. His research interests include Image Processing, Signal Processing, Virtual Reality, Pattern Recognition, Wireless Optical Communications. He is a member of the Information Processing Branch Association and the DSP Branch Association of China Electronics Society, one of the principals of the Shaanxi Signal Processing Society.

Jian Guo was born in 1963. He is currently a senior engineer in the Luoyang Optoelectro Technology Development Center. His research interests include Imagel Processing, imaging GIF fuze.