An Efficient of Coal and Gangue Recognition Algorithm

Kuidong Gao[1], Changlong Du[2], Haoxiang Wang[3] and Shibo Zhang[4]

[1, 2, 4]College of Mechanical and Electrical Engineering, China University of Mining and Technology, Xuzhou, Jiangsu, China, 221008
[3]College of Electronic Engineering, Xidian University, Xi'an, Shanxi, China, 710071
[1]gaokuidong@cumt.edu.cn, [2]luck_honey@163.com,
[3]wanghaoxiang1102@gmail.com, [4]wh-zhangshibo@163.com

Abstract

The separation for coal and gangue is an important process in mine production. In order to automatically select the gangue, this study obtained the difference and distribution regularity of the grayscale distribution by analyzing a large number of image data. By improving Bayesian Decision theory, identifiably character Bayesian Discriminant algorithm was proposed to get grayscale division threshold of coal and gangue. Aimed at the problem that impurities gangue and vitrinite in coal affect the accuracy of recognition result, mean smoothing filter algorithm was used to pre-process image and related neighborhood pixels recognition algorithm was proposed for recognizing the coal and gangue. The recognition system was tested on-line with a large number of random selected materials for many times, the average correct recognition was 96.8%. The test results indicated that the algorithm is stable and robust and the recognition system has a great potential in automatic selecting of gangue.

Keyword: *coal and gangue, grayscale intensity, bayesian decision theory, threshold, image processing*

1. Introduction

Initial separation for coal and gangue is an indispensable process of mine production. With the widely use of coal mine production's mechanization and the increasing extraction of thin seam, the output of gangue increases greatly, manual separation for coal and gangue is inefficiency and less reliable, so it cannot meet the requirements. Recent years, according to the difference of physical properties between coal and gangue, various separation methods have been proposed. Double photon absorptionmetry sorting machine for gangue can be used in the separation of the materials ranging from 50 to 150 mm in diameter, based on the difference of low energy γ-ray decay coefficient between coal and gangue, but its widely application is limited for γ-ray is harmful to human health [1]. Impact crushing and hydraulic break methods can be used in separating the gangue from coal by using the hardness difference between coal and gangue, but low efficiency and accuracy limit its application [2, 3].

Image recognition technology is an integrated technology which using the computer image processing technology, analysis technology and understanding technology to identify the target of various patterns, and its application field is aerospace. with recent advanced in computer technology, video-signal on-line transition and processing technology has been perfected and raised, computer-based vision system have great success in more and more fields. Bressan, *et al.*, used this technology to identify the

Corresponding author: Changlong Du
This work is supported by the industrialization of research results program of Jiangsu province under Grand (No.JHB2011-31).

different color object by their color histogram[4]; Museleta and Macaireb eliminated the influence of illumination on the recognition results and provided a new recognition method which combined the color composition and spatial configuration in recognition process[5]; Liu applied recognition technology in automatic classification of raisins which was based on raisins color and shape[6]; Zhi, *et al*., introduced the segmentation method of binary image and grey level image into the design of simulated pattern painting[7]; Chen, *et al*., proposed a novel discrete wavelet transform domain image watermark scheme to meet the watermarking properties[8]; Sun, *et al*. proposed a novel image recognition based on subspace and SIFT, which could provide a recognition from global features to minutiae features[9]; Gong, *et al*., applied a three-dimensional canonical correlation analysis method to feature fusion for image recognition[10]; Guo and Zhao applied recognition technology in character recognition of NaXi pictograph which was based on the feature extraction[11]; Deng, *et al*., illustrated a real-time hand gesture recognition system by using shape context matching and cost matrix[12]; Duan, *et al*., applied the wavelet analysis technology to recognize the vehicle license plate in low light level and achieved good results[13]; Lei, *et al*., proposed an extraction algorithms for geometric features of 3D face and recognition of the face by principal component analysis[14]; Xu and Ding proposed a method of processing and linear discriminant analysis of palmprint image which focused on the location and segmentation[15].Computer-based vision recognition technology provides high level of reliability and accuracy for recognition of materials, so it also had been widely used in other fields [16, 17], besides quality monitoring for mining industry [18, 19]. In the meantime, many scholars have turned their attentions to the study of recognition for coal and gangue based on computer vision. The difference of grayscale distribution between coal and gangue was found by analyzing the image grayscale histogram, and can be treated as the recognition feature for the on-line recognition system [20, 21]. According to character that the standard deviation of coal's diffuse reflectance light intensity is obviously higher than gangue, the discriminating photoelectric system was used in the separation for coal and gangue [22]. But because many technological problems have not been solved, these methods have not been applied extensively.

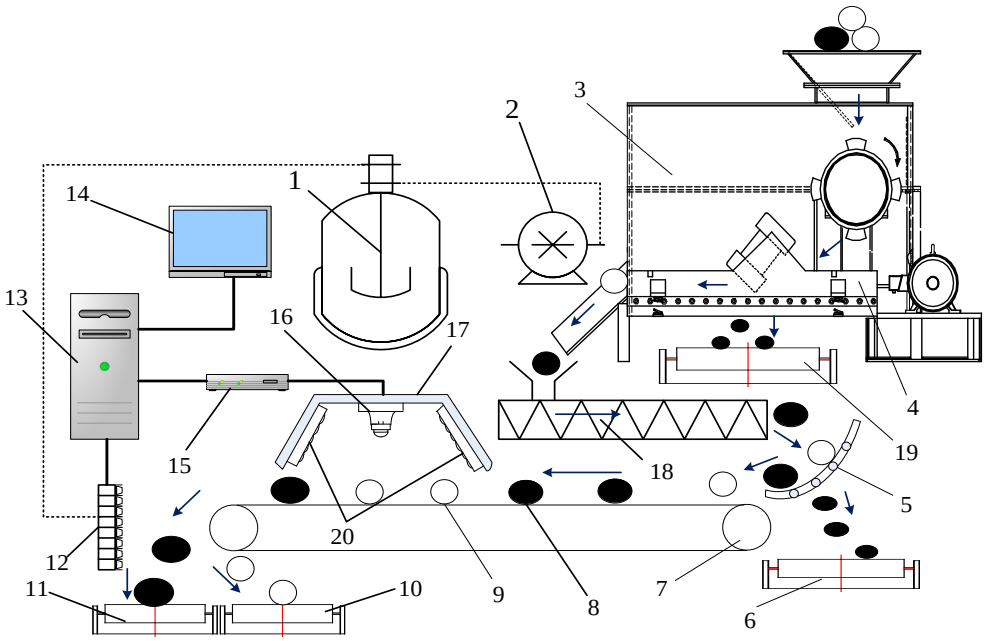
2. Materials and Recognizable Feature

2.1. Coal and Gangue

As the research objects, the coals and gangues were randomly selected from two mines (Jiahe colliery, Jiangsu, China), with a total number of 70 coals and 23 gangues, with totally different size and shape. By observing the texture and color character, gangue can be divided into three kinds: limestone, mudstone and sandstone, coal is only bituminous coal, the samples image are shown in Figure 1. Coal is mainly black and has low brightness in the whole, but reflection rate exists considerable increasing in partial area where high reflection rate vitrinite exists. Gangue surface is mainly gray and has higher brightness than coal. The main purpose of this paper is to find the recognizable feature between coal and gangue and design the algorithm to realize the recognition of gangue.

1-Limestone; 2-Coal; 3-Mudstone; 4-Sandstone

Figure 1. Samples Image



1-High pressure gasholder; 2-Air compressor; 3-Separation equipment; 4-Vibrating sieve; 5-Sieve plate; 6,11,19-Coal conveyer belt;7-Converyer belt for coal and gangue recognition; 8-Coal;
9-Gangue; 10-Gangue conveyer belt; 12-Nozzle array; 13-Control equipment; 14-Computer;
15-Image acquisition equipment; 16-CCD camera; 17-Camera box; 18-Spiral sort equipment;
20-LED light array

Figure 2. Separation System of Coal and Gangue

2.2. Acquisition of Image

For ensuring the accuracy of results and precise controllable of experimental condition, the images in this study were gotten from the machine vision system combined by CCD camera, Camera box, LED light array and conveyer belt, which also was an important part of separation system of coal and gangue as shown in Figure 2. The belt conveyer was used for transmitting samples into lighting chamber, lighting chamber was a closed stainless steel box with dimension of $170\text{cm}\times120\text{cm}\times140\text{cm}$, inside this chamber, the camera was mounted on the top of chamber to capture image of samples, six 10W LED flexible light bends were fixed to both sides and the top of chamber with an angle of 45° to provide homogenous lighting on the samples as they passed below the camera.

2.3. Grayscale Distribution Feature of Coal and Gangue

The distribution regularity of image grayscale is one of recognizable features; it can be obtained by analyzing grayscale histogram, grayscale histogram expresses statistical relation between every gray level and its frequency [23], it can be calculated by applying (1):

$$P_f(f_k)=\frac{n_k}{n}, k=0,1,\cdots,L-1 \tag{1}$$

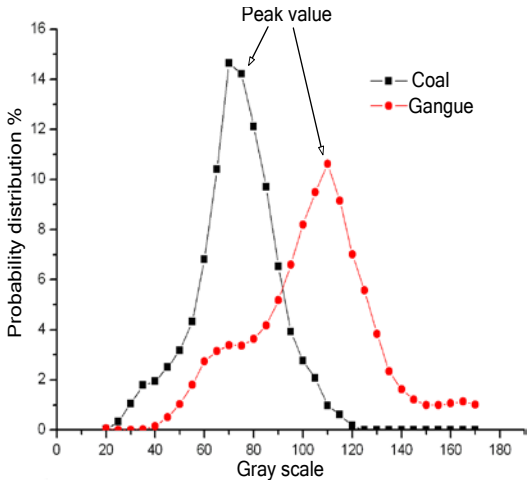
Where f_k is grayscale value of level k, n_k is the mount of pixels that the intensity is f_k, n is pixel number, L is grayscale level number, in addition $\sum_{k=0}^{L-1}P_f(f_k)=1$.

In this work, according to the statistics of RGB three components' grayscale

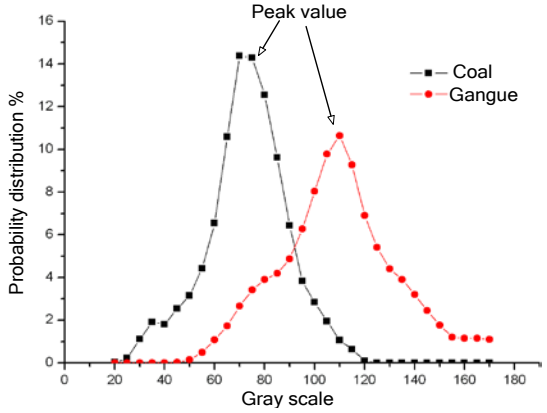
distribution in coal and gangue images, the grayscale histogram was built as shown in Figure 3, in the figure, grayscale as abscissa (range from 0 to 255) against distribution of grayscale as ordinate. By analyzing the distribution of grayscale, we can get the conclusions: Grayscale distribution features of RGB components are similar, so we chose the mean value as the research objects; Grayscale distributions of coal and gangue are concentrated at two areas and have small variance, the difference is obvious; The overlap of grayscale distributions between coal and gangue will affect the accuracy of result, but the area of intersection region is small, which can be eliminated by image pre-processing.

In this work, according to the statistics of RGB three components' grayscale distribution in coal and gangue images, the grayscale histogram was built as shown in Figure 3, in the figure, grayscale as abscissa (range from 0 to 255) against distribution of grayscale as ordinate. By analyzing the distribution of grayscale, we can get the conclusions: Grayscale distribution features of RGB components are similar, so we chose the mean value as the research objects; Grayscale distributions of coal and gangue are concentrated at two areas and have small variance, the difference is obvious; The overlap of grayscale distributions between coal and gangue will affect the accuracy of result, but the area of intersection region is small, which can be eliminated by image pre-processing.

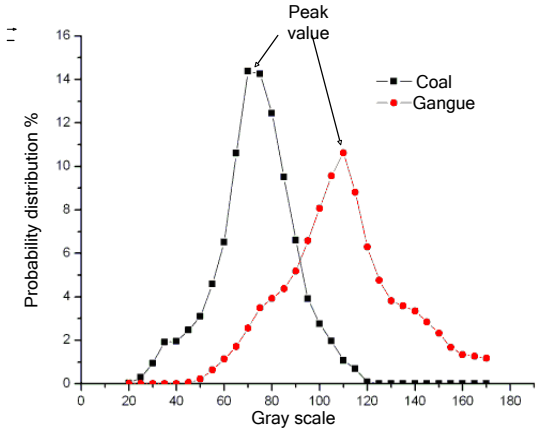
According to above conclusions, this study used the distribution feature of mean grayscale value of RGB components as research object. The identifying process had various steps. The first step was designing the algorithm to calculate the threshold to segment the pixels of gangue from coals'. The second step was pre-processing the image by mean filter algorithm and designing neighborhood association identify algorithm to reduce the influence on the accuracy of result of overlap of grayscale distributions.



(a) Grayscale histogram of red component



(b) Grayscale histogram of green component



(c) Grayscale histogram of green component

Figure 3. Grayscale Histograms of RGB Components

3. Image Feature Extraction and Pre-processing

3.1. Threshold Selection Algorithm

Aiming at the condition that gangue and coal's grayscale feature cannot be perfectly separated, which caused by intersection region in grayscale histogram, the statistical decision theory was applied to select threshold. Bayesian decision theory is a basic method of statistical pattern recognition [24, 25], at present, there are two major decision algorithms based on Bayesian decision theory including minimum error Bayesian decision and minimum risk Bayesian decision, but in application process, two algorithms did not have good processing result, the calculation had big error.

For the difference of grayscale distribution between coal and gangue image was the difference of peak values in grayscale histogram, recognizable decision algorithm was appropriated, by bringing in recognizable degree variable $\varphi(\omega_i)$ as the influence factor for the decision result, which was calculated using (2):

$$P(x \mid \omega_i) = \frac{p(\omega_i \mid x)\varphi(\omega_i)}{\sum_{j=1}^{2} p(\omega_j \mid x)\varphi(\omega_j)} \tag{2}$$

Where ω is the actual category, includes coal and gangue two categories; x is object property, represents the grayscale value of pixel; $\varphi(\omega_i)$ is the influence factor of decision result, called prior probability; $p(\omega_i \mid x)$ is the probability of x under the condition of ω_i ,called class conditional probability; $p(x \mid \omega_i)$ is the probability of ω_i under the condition of x ,called posteriori probability. In addition $\sum_{j=1}^{n} \varphi(\omega_j) = 1$.

The calculation process of threshold is as follows: firstly, get the distribution of class probability density distribution by counting up the sample images data; secondly, get the posteriori probability distribution $p(x \mid \omega_i)$ by substituting class conditional probability into Eq.(2); Finally, choose the grayscale value, whose posteriori probability equal 0.5, as the divided threshold.

The obtaining value method of $\varphi(\omega_i)$ in (2) is unknown, in this paper by comparing the effect of relationship between coal's recognizable degree variable $\varphi(\omega_1)$ and difference of peak values to gangue recognition results and got the obtaining value method distribution curve as shown in Figure 4, when the difference of peak value is less than 40, the obtaining value of $\varphi(\omega_1)$ is small and changes smoothly, because the intersection area

of grayscale distribution in this region is big, the recognition result focuses on wiping out the coal's pixels; when the difference of peak value is more than 45, the obtaining value of $\varphi(\omega_1)$ increased rapidly, because the intersection area of grayscale distribution in this region is small, the recognition result is assuring the clearness of gangue shape.

Took coal and gangues in Figure 1 as an example, the difference of grayscale peak values was 50, the obtaining value of recognizable degree $\varphi(\omega_1)$ was 0.19 in terms of the curve given in Figure 4, and got the posteriori probability by (2) as shown in Figure 5, the threshold was 78 by searching the pot with value to be 0.5 in the posteriori probability curve, the two-valued processing result of Figure 1 is shown in Figure 6.

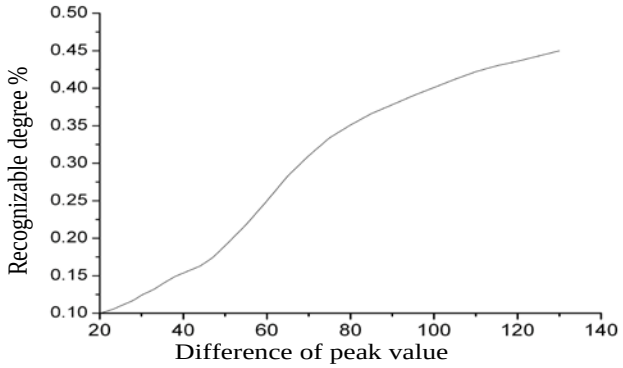


Figure 4. Obtaining Value Method of $\varphi(\omega_i)$

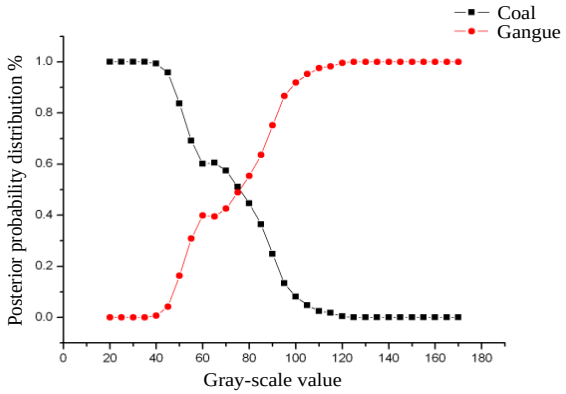


Figure 5. Posteriori Probability of Coal and Gangue Gray-scale Value

Figure 6. Two-valued Processing Result

In Figure 6, the threshold obtained by recognizable decision algorithm can separate the most pixels of coal and gangue, but the existing miscalculations will affect the recognition result. By analyzing the color and composition of coal and gangue, the impurities in gangue and vitrinite in coal are the main influence factors of miscalculations, the

impurities in gangue have several colors and vitrinite in coal has high reflectance rate.

3.2. Image Pre-processing

For improving the accuracy of recognition and eliminating the image noise, mean filter algorithm was used in pre-processing the image [26]. The calculating process of mean filter algorithm is replacing the value of pixel by mean value of the value of pixels in a region of $N \times N$. A image could be pre-processed by using (3):

$$f(x,y) = \frac{1}{N^2} \times \sum_{i=(-N/2)}^{N/2} \sum_{j=(-N/2)}^{N/2} f(x+i, y+j) \tag{3}$$

Where (x,y) is calculating pixel, $N \times N$ is a region centered in (x,y), $f(x,y)$ is the grayscale value of (x,y).

Pre-processed Figure 1 with different template size, counted the erroneous judgment pixels of recognition result, which pre-processed by mean filter algorithm with different size templates, is shown in Table 1. Comparing the quantity of erroneous judgment pixels, the mean filter algorithm can filter noise pixels of gangue image, the processing efficiency increases with the increasing of the template size, but has a poor effect on coal images.

Table 1. Processing Effect of Mean Filter Algorithm

		Non preprocessed	Pre-processed		
			3×3	5×5	10×10
Quantity of erroneous judgment pixels	Coal	5819	5751	5300	4700
	Gangue	3486	2862	2800	2100
Processing efficiency for coal			1.16 %	8.92%	18.23 %
Processing efficiency for gangue			17.9 %	19.67 %	39.75 %

4. Related Neighborhood Pixels Recognition Algorithm

Recognition algorithm determines the accuracy of recognition results, which is the most important part of process. In this work, the Related neighborhood pixels recognition algorithm (RNRA) was established based on erosion algorithm [27], the calculate process of RNRA is as Eq(4):

$$G(j,k) = F(j,k) \otimes H(j,k) \tag{4}$$

Where $F(j,k)$ is target image; $H(j,k)$ is structure element.

The character of RNRA is that $H(j,k)$ includes six kinds of structure element models as shown in Figure 7, and they can keep the boundary pixels and eliminate noise pixels. The calculate process is as follows: firstly, threshold process the image; secondly, scan image data space $F(j,k)$ by six kind of models, and calculate the intersection of matrix, if the intensity of all the pixels in $F(j,k)$ corresponded with the location that the pixels in $H(j,k)$ whose intensity is1,then keep the pixel that corresponds the centre point of $H(j,k)$, if not clean the pixel that corresponds the centre point of $H(j,k)$; Finally, combine six processing results together.

The RNRA was tested for processing the image obtained under the experiment conditions as shown in Figure 8(a), the size of image is 640×480 pixels. Firstly, pre-process the image by mean filtering method; Secondly, to fit the requirement of separation machine, reduce the image into 64×48 pixels(Figure 8(b)); Finally, process the image by erosion algorithm and RNRA, the processing results are shown as Figure 8(c)

and Figure 8(d). In the image (Figure 8(c)) processed by erosion algorithm exists the condition that one piece of gangue was separated into two pieces. In the image (Figure 8(d)) processed by RNRA, the boundary of gangue is completed.

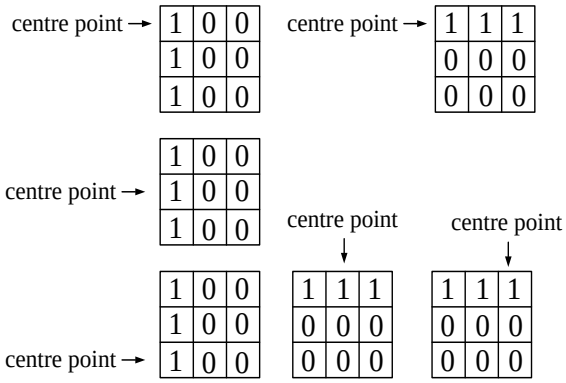


Figure 7. Six Kinds of Structure Element Models

5. Test Verification

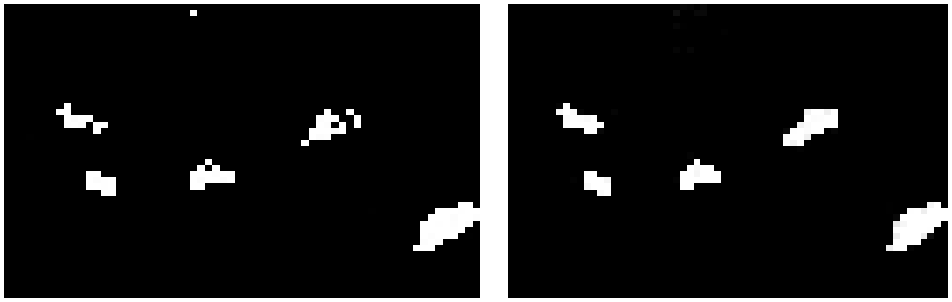
Programmed the algorithm by Visual C++, ported the program to the gangue separator which is based on computer, verified the algorithm.

Selected 70 gangues and 23 coals as the testing samples, performed tests in the machine vision system as shown in Figure 2.Firstly, calculated threshold by analyzing the grayscale distribution of images; Secondly, pre-processed the image and recognize the gangue by RNRA. In order to verify the accuracy and robustness of algorithm, a lot of tests have been done under different light intensity and sample arrangement, fewer disturbances existed in the recognition results, and the efficiency of gangue recognition is shown in Table 2.

Table 2. Efficiency of Gangue Recognition

Quantity of samples	Quantity of accurately recognition	Quantity of no recognition	Quantity of error recognition	Accuracy rate /%	Work efficiency /(kg.s^{-1})
265	256	9	0	96.8%	13

(a) original image (b) processing image



(c) image processed by erosion algorithm (d) image processed by RNRA

Figure 8. The Processing Results of Image

6. Conclusions

As a part of whole machine-vision system, an image processing technique was developed to recognize the gangue based on grayscale feature. The experimental machine vision system was designed and used in testing the accuracy and robustness of algorithm. Recognizable decision algorithm based on Bayesian decision theory was developed to calculate the threshold of the grayscale distribution histogram. Mean filter pre-processing algorithm was used to eliminate the noise pixels. RNRA was developed to recognize the gangue while they travelling over a belt conveyer. In addition, the overall accuracy of system for recognizing gangue was 96.8%, in short, it is a reassuring way to recognize gangue from coal.

Acknowledgements

This work is supported by the industrialization of research results program of Jiangsu province under Grand (No.JHB2011-31).

References

[1] Y. Jin, J.-H. Zhang and G.-Z. Chen, “Nuclear Electronics and Detection Technology”, vol. 14, no. 5, **(1997)**.
[2] J. Li, C. Du and L. Xu, Journal of China Coal Society, vol. 36, no. 4, **(2011)**.
[3] C. Dong and Z. Liu, “Coal Science and Technology”, vol. 35, no. 3, **(2007)**.
[4] M. Bressan, D. Guillamet and J. Vitria, “Pattern Recognition”, vol. 36, no. 3, **(2003)**.
[5] D. Muselet and L. Macaire, “Pattern Recognition Letters”, vol. 28, no. 10, **(2007)**.
[6] X. Liu, “Study on raisin grading technology based on image analysis”, Master Thesis, Northwest Agriculture & Forestry University, **(2006)**.
[7] J. Zhi, J.-Y. Liu, H. Yuan and G.-H Wu, Journal of Software, vol. 6, no. 11, **(2011)**.
[8] C. Yongqiang, Z. Yanqing, H. Hanpin and L. Hefei, Journal of Software, vol. 6, no. 5, **(2011)**.
[9] T. Sun, S. Ding and Z. Ren, Journal of Software, vol. 8, no. 5, **(2013)**.
[10] X. Gong, J. Zhou, H. Wu, G. Lei and X. Li, Journal of Computers, vol. 6, no. 11, **(2011)**.
[11] H. Guo and J.-Y. Zhao, Journal of Computers, vol. 6, no. 5, **(2011)**.
[12] L. Y. Deng, J. C. Hung, H.-C. Keh, K.-Y. Lin, Y.-J. Liu and N.-C. Huang, Journal of Networks, vol. 6, no. 5, **(2011)**.
[13] P. Duan, K. Xie, N. Song and Q. Duan, Journal of Networks, vol. 5, no. 12, **(2010)**.
[14] L. Yunqi, L. Haibin and L. Qingmin, Journal of Multimedia, vol. 6, no. 2, **(2011)**.
[15] S. Xu and J. Ding, Journal of Multimedia, vol. 7, no. 3, **(2012)**.
[16] J. Tessier, C. Duchesne and G. Bartolacci, Mining Engineering, vol. 20, no. 12, **(2007)**.
[17] G. Elmasry, S. cubero, E. Molto and J. Blasco, “Journal of Food Engineering”, vol. 112, no. 1, **(2012)**.
[18] V. Singh and S. M. Rao, Mineral Engineering, vol. 18, no. 15, **(2005)**.
[19] S. Chatterjee, A. Bhattacherjee, B. Samanta and S. Kumar, Computers in Industry, vol. 61, no. 5, **(2010)**.
[20] F. Liu, J. Qian, X. Wang and J. Song, Journal of China Coal Society, vol. 25, no. 10, **(2000)**.
[21] D. M. Hobson, R. M. Carter, Y. Yan and Z. Lv, International Workshop on Imaging Systems and Techniques, vol. 5, **(2007)**.
[22] W. Li, J. Yang and J. Wen, China Journal of Scientific Instrument, vol. 22, no. 3, **(2001)**.
[23] J. Zhao, Z. Wei, Y. Wang and S. Gao, Transactions of the Chinese Society of Agricultural Engineering, vol. 27, no. 4, **(2011)**.

[24] K. Fukunaga, "Introduction to Statistical Pattern Recognition", Academic Press, New York, **(1990)**.
[25] Z. Bian and X. Zhang, "Pattern Recognition", Tsinghua Press, Beijing, **(2000)**.
[26] M. S. Moore and S. K. Mitra, Nonlinear Image Processing X. 3646, **(1999)**.
[27] J. Serra, "Image Analysis and Mathematical Morphology", Academic Press, London, **(1989)**.

Authors

Kuidong Gao, was born in Shandong Province of China in 1986. He received his Bachelor degree in mechanical design and theory from Central South University, China, in 2008. Currently he is a PhD student at Department of Mechanical Design and Theory, College of Mechanical and Electrical Engineering, China University of Mining and Technology, pursuing his PhD degree. His research interests include rock broken theory, separation theory of coal and gangue and pattern recognition.

Changlong Du, PhD, Professor. He was born in Jiangsu province of China in 1958. He received his master degree in mining machinery from China University of Mining and Technology in 1987, PhD degree in mining from China University of Mining and Technology in 1994. From 1987 to 1990, he was a Lecturer at China University of Mining and Technology, and was an Associate Professor at China University of Mining and Technology from 1990 to 2001. At present, he is a professor at China University of Mining and Technology. His current research interests include rock broken theory, mining machine design, separation theory of coal and gangue and pattern recognition.

Haoxiang Wang was born in Shaanxi province of China, in 1991. He is currently an undergraduate student in School of Electronic Engineering in Xidian University(XDU). His recent research mainly focus on pattern recognition, image understanding and changing detection (especially for the VHR images).

Shibo Zhang was born in Shandong Province of China in 1988 He received his Bachelor degree in applied physics from China University of Mining and Technology, China, in 2010. Currently he is a postgraduate student at Department of Mechanical Design and Theory, College of Mechanical and Electrical Engineering, China University of Mining and Technology, pursuing his master degree. His research interests include separation theory of coal and gangue and pattern recognition.