International Journal of Minerals, Metallurgy and Materials
Volume 18, Number 4, August 2011, Page 385
DOI: 10.1007/s12613-011-0451-8

Segmentation algorithm of complex ore images based on templates transformation and reconstruction

Guo-ying Zhang, Guan-zhou Liu, and Hong Zhu

School of Mechanical-Electrical and Information, China University of Mining & Technology, Beijing 100083, China
(Received: 25 February 2011; revised: 22 March 2010; accepted: 5 April 2010)

Abstract: Lots of noises and heterogeneous objects with various sizes coexist in a complex image, such as an ore image; the classical image thresholding method cannot effectively distinguish between ores. To segment ore objects with various sizes simultaneously, two adaptive windows in the image were chosen for each pixel; the gray value of windows was calculated by Otsu's threshold method. To extract the object skeleton, the definition principle of distance transformation templates was proposed. The ores linked together in a binary image were separated by distance transformation and gray reconstruction. The seed region of each object was picked up from the local maximum gray region of the reconstruction image. Starting from these seed regions, the watershed method was used to segment ore object effectively. The proposed algorithm marks and segments most objects from complex images precisely.

Keywords: ores; image analysis; image segmentation; morphological transformation; algorithms

[*The work was financially supported by the National Key Technologies R & D Program of China (No.2009BAB48B02) and the National High-Tech Research and Development Program of China (Nos.2010AA060278600 and 2008AA062101).*]

1. Introduction

Getting the physical characteristics of conveyor belt ores by capturing and segmenting images is an effective method for crusher automatic control. However, the marking and segmenting of ore objects in a complex image are very challenging [1-2]. Some heterogeneous ore objects with various sizes are coexisting, and a number of small ores embedded into big ores bring about noise. A part of boundaries of conglutinating ores are blurred.

Pixel feature analysis, such as color and texture analysis, is the classical method for image object segmentation and recognition [3-4]. The segmentation technique that generates binary object shapes through learning of pixel and shape features is proposed by Mukherjee *et al*. [5]. Pixel analysis is ineffective for complex ore segmentation, because some pixels within an ore object have different properties. Kittler *et al*. [6] used a gray threshold method to extract the highlight region of objects, but some of object shapes were destroyed. Otsu's method is a global threshold method, which selects the threshold based on the maximum of between-class variances [7]. Otsu's approach may miss some lower brightness objects in a multi-object complex image.

The dynamic threshold method is a suitable selection for complex image segmentation. Simphiwe [8] used two windows to segment ore images. The larger window treats large ores more precisely and the smaller window finds out small ores more completely. The image is thresholded by the mean gray in windows. However, ore and background information within windows is uneven and some big noise areas within the ores are diffused.

An adaptive Otsu threshold method [9] combining bi-window with Otsu threshold method was used in this paper. The sizes of the two windows were determined by ore object geometrical features. The gray thresholds of the two windows were calculated using Otsu's optical threshold

Corresponding author: Guo-ying Zhang E-mail: zhangguoying1101@163.com

method. Each pixel was binarized according to the smaller gray value of the two windows. By the method adaptive to threshold complex ore images, more edges between objects could be preserved and more noise areas within the ores could be removed.

Some linking ore objects in a binary image resulted from adaptive Otsu thresholds may be segmented into a whole object. Just extracting a seed region for each ore object can guarantee that most of ores are identified and segmented precisely. The thinning algorithm [10] can separate conglutinated ore objects through extracting the ore skeleton, but many ore objects are out of their original shapes. Distance transformation [11-12] of the binary image can segregate these linked objects and ore shapes is well maintained. The definition principle of distance transformation templates was proposed in the paper. Reconstruction [13] after the distance image can obtain high gray regions of each ore, and the highest gray region of each ore is marked as its seed region. By the marker-based watershed algorithm [14-15], seed regions grow outward continuously in parallelism until the boundary condition of segmented ore objects is satisfied.

2. Adaptive Otsu threshold

The threshold method is an image segmentation method, and gray thresholds are selected for discriminating two or more classes of gray levels in an image. An adaptive Otsu algorithm with the advantages of both bi-windows and Otsu threshold can treat a gray image into a high-performance binary image.

For each pixel, its adaptive threshold is the maximum of intensity variances between neighborhood object pixels and neighborhood background pixels. The step of adaptive Otsu thresholding is as follows.

(1) Select two neighborhoods for every pixel. For each pixel p in image $f(x,y)$, select two windows with the sizes of $n\times n$ and $m\times m$ respectively, which are determined by image objects (n and m are odd integers). The inspected pixel p is located at the center of windows.

(2) Calculate the variance of neighborhood images. For a given neighborhood image $s(x,y)$ with h gray level, all pixels in image $s(x,y)$ are divided into two groups by the threshold j ($0\le j \le h-1$), and. Group A includes all pixels whose gray values are less than or equal to j, and group B includes all pixels whose gray values are greater than j. ($w_A(j)$, $M_A(j)$) and ($w_B(j)$, $M_B(j)$) denote the number of pixels and the average gray level in groups A and B, respectively. M_T is the average gray level value of all the pixels in image $s(x, y)$, and $\sigma^2(j)$ is the variance between the two groups.

(3) Calculate the optical window threshold. The variance of each gray level j ($0 < j < h-1$) is calculated by Eq. (1). The gray value corresponding to the greatest variance is the resulting threshold T of image $s(x, y)$. $T_{\text{Otsu}}(N_p^n, j)$ and $T_{\text{Otsu}}(N_p^m, j)$, determined by maximal between-class variance $\sigma^2(j)$, denote the two optimal window thresholds of pixel p.

$$\begin{cases} T_{\text{Otsu}}(N_p^n, j) = \arg\max\limits_{0\le j<l}\{\sigma^2(j)\} \\ T_{\text{Otsu}}(N_p^m, j) = \arg\max\limits_{0\le j<l}\{\sigma^2(j)\} \end{cases} \tag{1}$$

(4) Binarize image. For each pixel p, the minimum gray value between $T_{\text{Otsu}}(N_p^n, j)$ and $T_{\text{Otsu}}(N_p^m, j)$ is chosen and pixel p is thresholded by Eq. (2). Finally, the binary image $f_b(x,y)$ for all indices x and y is obtained.

$$f_b(x,y) = \begin{cases} 0, & f(x,y) < \min\{T_{\text{Otsu}}(N_p^n, j), T_{\text{Otsu}}(N_p^m, j)\} \\ 1, & \text{otherwise} \end{cases} \quad \text{for each } p \in f(x,y) \tag{2}$$

3. Distance transformation based on templates

A binary image is expressed by $f_b(x,y) = \{X, X^C\}$, where X is the subset of all ore pixels p_o and X^C is the subset of all background pixels p_b. By assigning each ore pixel to a corresponding gray value, which is the shortest distance to its nearest background pixel, $f_b(x,y)$ is transformed into distance image $f_d(x,y)$ according to Eq. (3).

$$f_d(x,y) = \begin{cases} \min\{\text{dis}(p_o, p_b) \mid p_o \in X \,\&\, p_b \in X^C\} \\ 0, \; p_b \in X^C \end{cases} \tag{3}$$

Function dis(p_o, p_b) denotes the distance of each ore pixel p_o within X to its neighborhood background pixel p_b within X^C. Distance image $f_d(x,y)$ is resulted by a pair of templates; the gray value of each background pixel p_b in templates is zero and the gray value of each ore pixel p_o in templates is the shortest distance to its nearest background pixel p_b. Distance transformation can be implemented by definitions 1 and 2.

Definition 1 The size of a pair of templates M^+ and M^- is $N\times N$. N is equal to 3 in Fig. 1(a) and equal to 5 in Fig. 1(b). N varies with the sizes of objects detected in the image. The value of a template element denotes the distance of the element to the template center.

Starting from the upper left pixel of image $f_b(x,y)$, a sub-image $\boldsymbol{N}_{p_o}$ is chosen in $f_b(x,y)$ for each object pixel p_o.

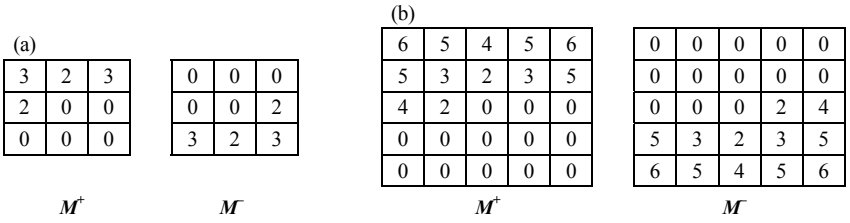

Fig. 1. Two pairs of templates with different sizes: (a) 3×3; (b) 5×5.

The size of $\boldsymbol{N}_{p_o}$ is the same as that of template M^+ and pixel p_o is located in its center. Elements of corresponding position in M^+ and $\boldsymbol{N}_{p_o}$ are added together.

Definition 2 $\oplus$ is an operator. $\boldsymbol{N}_{p_o} \oplus \boldsymbol{M}$ denotes the adding relationship of elements in $\boldsymbol{N}_{p_o}$ and M.

$$\boldsymbol{N}_{p_o} \oplus \boldsymbol{M} = \begin{cases} N^i_{p_o} + M^i, (N^i_{p_o} \in \boldsymbol{N}_{p_o} \,\&\, N^i_{p_o} > 0) \,\& \\ \quad (M^i \in \boldsymbol{M} \,\&\, M^i > 0) \\ 0, \quad (N^i_{p_o} \in \boldsymbol{N}_{p_o} \,\&\, N^i_{p_o} = 0) \,\| \\ \quad (M^i \in \boldsymbol{M} \,\&\, M^i = 0) \end{cases} \tag{4}$$

Index i denotes the i-th element in template M^+ and M. Only when both $N^i_{p_o}$ and M^i are non-zero, $N^i_{p_o}$ and M^i are added together; otherwise, the sum of $N^i_{p_o} + M^i$ is equal to zero. The gray value of each object pixel is the minimum of $\boldsymbol{N}_{p_o} \oplus \boldsymbol{M}^+$. The upper left skeleton image $f^+_{\mathrm{d}}(x, y)$ is obtained by M^+.

$$f^+_{\mathrm{d}}(x, y) = \begin{cases} \min\{\boldsymbol{N}_{p_o} \oplus \boldsymbol{M}^+\}, p_{\mathrm{o}} \in X \\ 0, p_{\mathrm{b}} \in X^{\mathrm{C}} \end{cases} \tag{5}$$

Scanning an image from the lower right pixel, the lower right skeleton image $f^-_{\mathrm{d}}(x, y)$ is obtained by module M^-. Distance image $f_{\mathrm{d}}(x,y)$ is obtained by

$$f_{\mathrm{d}}(x, y) = f^+_{\mathrm{d}}(x, y) + f^-_{\mathrm{d}}(x, y). \tag{6}$$

4. Ores segmentation

Distance an image may contain a set of local maximum regions whose gray values are higher than their neighborhood pixels. Each ore pixel in $f_{\mathrm{d}}(x,y)$ degrades a constant value h, and ores linked together in $f_{\mathrm{d}}(x,y)$ could be separated [13].

$$f_h(x,y) = f_{\mathrm{d}}(x,y) - h \tag{7}$$

Ore pixels in $f_h(x,y)$ are the subsets of ore pixels in $f_{\mathrm{d}}(x,y)$. Gray scale reconstruction is defined in terms of geodesic dilation method. An elementary geodesic dilation of gray scale image $f_h(x,y)$ under $f_{\mathrm{d}}(x,y)$ is defined as

$$\delta^{(1)}_{f_{\mathrm{d}}(x,y)}(f_h(x,y)) = (f_h(x,y) \oplus \boldsymbol{B}) \wedge f_{\mathrm{d}}(x,y) \tag{8}$$

where $f_h(x,y) \oplus B$ is the gray scale dilation of $f_h(x,y)$ by a structuring element of size unity, and ^ the point-wise minimum operator. A geodesic dilation of size n is obtained by iterating elementary geodesic dilations n times in the binary case. The gray scale reconstruction $R_{f_{\mathrm{d}}(x,y)}(f_h(x,y))$ of $f_h(x,y)$ originated from $f_{\mathrm{d}}(x,y)$ is obtained by iterating geodesic dilations of $\delta^{(1)}_{f_{\mathrm{d}}(x,y)}(f_h(x,y))$ until stability is reached. For detected ores, gray reconstruction remains a region set whose gray value is higher than neighborhood pixels. The highest gray region within an ore is marked as the ore seed region.

$$\mathrm{Set(regions)} = f_{\mathrm{d}}(x, y) - R_{f_{\mathrm{d}}(x,y)}(f_h(x,y)) \tag{9}$$

A classical method for image segmentation application is the watershed transform. An image intensity map is viewed as a topographical landscape where intensity maxima are catchment basins and the ridges are the watersheds [16]. For these image objects whose color and texture information are unreliable and inconsistent, their general shortcomings are over-segmentation and sensitive to noise. In morphological difference image, all seed regions grow outward in parallelism according to the marked watershed method until the boundary conditions of segmented ores are satisfied.

5. Experiment result and analysis

Thresholding is important for object segmentation of complex image, and a lot of different objects are transformed into foreground of binary image. Complex ore image in conveyor belt shown in Fig. 2(a) includes a lot of noises and ores with all sizes. The performance of adaptive Otsu threshold is validated by binarizing the ore image and being compared with Otsu's method and the bi-window threshold method. The thresholding results of the Otsu, bi-window, and adaptive Otsu method are shown in Fig. 2(b)-2(d).

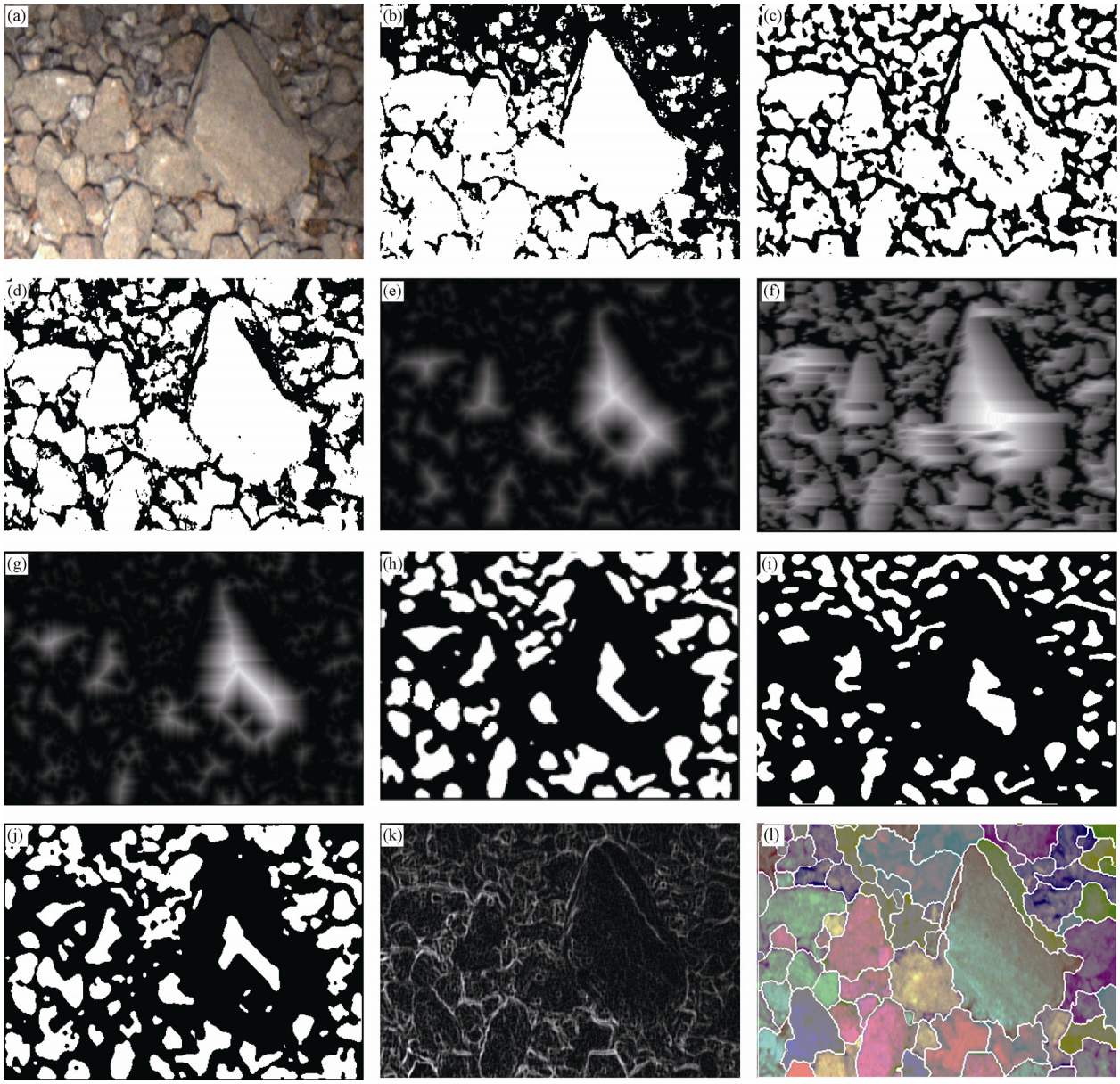

Fig. 2. Original image and the results: (a) original complex ore image with a size of 408×289; (b) binary image using Otsu's method; (c) binary image of the bi-windows technique; (d) binary image of the adaptive Otsu the technique; (e) distance image of the ore image by templates (a); (f) distance image of the ore image by templates (b); (g) thinning result of the ore image; (h) seed image of (e); (i) seed image of (f); (j) seed image of (g); (k) morphological gradient image of the original ore image; (l) image segmentation result with the watershed algorithm from (h).

The Otsu threshold method does not simultaneously find objects with different sizes. In Fig. 2(b), more than 60% of smaller ores are thresholded into background and almost all of adjacent larger ores are conglutinated together.

The smaller window with a size of 25×25, which is close to the mean size of small ores in Fig. 2(a), is used for detecting small ore objects, and the larger one with a size of 195×195, which is close to the size of the largest ore in Fig. 2(a), is used for detecting larger ore objects. In Fig. 2(c), the bi-window method can detect more than 80% of small ore objects. However, the noise areas within ores are diffused by the bi-window method, and 7.1% pixels within ores are in the form of holes.

The threshold result of the adaptive Otsu threshold in Fig. 2(d) is superior to that of the Otsu threshold and bi-window threshold. In Fig. 2(d), most of the ores are well-segmented, and almost no noise exists in ore objects.

Using templates in Fig. 1(a) and 1(b), the distance trans-

formation results of Fig. 2(d) are shown in Fig. 2(e) and 2(f). Thinning results of Fig. 2(d) is shown in Fig. 2(g). Reconstruction parameter h is 25. For Fig. 2(e)-2(g), their seed images are shown in Fig. 2(h)-2(j), respectively. The morphological gradient result of Fig. 2(a) is shown in Fig. 2(k).

The segmentation result in Fig. 2(h) is almost the same as that in Fig. 2(i). Eight additional small ores are found and two ores are over-segmented in Fig. 2(i). The segmentation results in Fig. 2(h) and 2(i) are better than that in Fig. 2(j); the largest ore region is segmented into four different seeds and more than 10% ore seed areas are linked together in Fig. 2(j).

Each seed region is growing outward in a morphological gradient image to search the boundary of ore objects. Ore segmentation is dealt with the traditional watershed algorithm, and the segmentation result of the seed image in Fig. 2(h) is shown in Fig. 2(l). Most of large ores are segmented precisely. However, some small ores are aggregated together. The templates proposed in this paper are very adaptive for complex ore image segmentation. The under-segmentation problem can be overcome by marking each ore object in Fig. 2(f) with the edge tracing method.

6. Conclusion

An adaptive Otsu threshold method was used to segment a complex ore image. The resulting binary image has fewer conglutinated ores than Otsu threshold' result and has less noise area than the bi-windows thresholding result. The principle and templates of distance transformation are proposed; a reconstruction technique based on distance transformation benefits seed region extraction of ore objects. Seed regions grow outward to the ore boundary in a morphological gradient image. The adaptive Otsu threshold method and template transformation have a good adaptability for object segmentation in a complex image.

References

[1] J.A. Sanchidrián, P. Segarra1, F. Ouchterlony, *et al.*, On the accuracy of fragment size measurement by image analysis in combination with some distribution functions, *Rock Mech. Rock Eng.*, 42(2009), p.95.

[2] J. Tessier, C. Duchesne, and G. Bartolacci, A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts, *Min. Eng*., 20(2007), p.1129.

[3] T. Mäenpää and M. Pietiäinen, Classification with color and texture: jointly or separately, *Pattern Recognit.*, 37(2004), p.1629.

[4] H. Stephen, *Texture Measures for Segmentation* [Dissertation], University of Cape Town, Cape Town, 2007, p.30.

[5] D.P. Mukherjee, Y. Potapovich, I. Levner, *et al.*, Ore image segmentation by learning image and shape features, *Pattern Recognit. Lett.*, 30(2009), p.615.

[6] J. Kittler and J. Illingworth, Minimum error thresholding, *Pattern Recognit.*, 19(1986), No.1, p.41.

[7] N. Otsu, A threshold selection method from gray level histograms, *IEEE Trans. Syst. Man Cybern.*, 9(1995), No.1, p.62.

[8] M. Simphiwe, *A Machine Vision-based Approach to Measuring the Size Distribution of Rocks on a Conveyor Belt* [Dissertation], University of Cape Town, Cape Town, 2004, p.23.

[9] G.Y. Zhang, G.Z. Liu, H. Zhu, and B. Qiu, Ore image thresholding using bi-neighborhood Otsu's approach, *Electron. Lett.*, 46(2010), p.1666.

[10] E.R. Davies and A.P.N. Plummer, Thinning algorithms: a critique and a new methodology, *Pattern Recognit.*, 14(1981), p.53.

[11] P.J. Toivanen, New geodesic distance transforms for gray-scale images, *Pattern Recognit. Lett.*, 17(1996), No.5, p.437.

[12] S. Svensson and G.Sanniti Di Baja, Using distance transforms to decompose 3D discrete objects, *Image Vision Comput.*, 20(2002), No.8, p.529.

[13] G.Y. Zhang and Y. Sha, *Object Segmentation and Recognition of Mining*, Petroleum Industry Press, Beijing, 2010, p.69.

[14] C. Snehamoy, B. Ashis, S. Biswajit, *et al.*, Rock-type classification of an iron ore deposit using digital image analysis technique, *Int. J. Min. Miner. Eng.*, 1(2008), No.1, p.22.

[15] I. Levner and H. Zhang, Classification-driven watershed segmentation, *IEEE Trans. Image Process.*, 16(2007), No.5, p.1437.

[16] L. Vincent and P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(1991), No.6, p.583.