Multi-scale watershed segmentation of high-resolution multi-spectral remote sensing image using wavelet transform

CHEN Jie[1], DENG Min[1], XIAO Pengfeng[2], YANG Minhua[1], MEI Xiaoming[1], LIU Huimin[1]

1. *Department of Surveying and Geo-informatics*, *Central South University*, *Hunan Changsha* 410083, *China*;
2. *Department of Geographical Information Science*, *Nanjing University*, *Nanjing* 210093, *China*

Abstract: In order to reduce over segmentation caused by only using watershed algorithm, an efficient multi-scale approach using wavelet transform is presented for the segmentation of the pan-sharpened multi-spectral QuickBird image. The approach toward complete segmentation includes four steps, namely, multi-scale images representation, image segmentation, region merging and result projection. First, the wavelet decomposition is implemented independently for each spectral band to form a number of new images at lower resolutions according to the size of original image. Gradient images are obtained by applying phase congruency model to approximation coefficients, and gradient magnitudes of all bands are fused for each scale. The optimal scale of wavelet decomposition is selected through analysis local gradient variance varying correspond to different scales and varieties of geo-objects. Second, a multi-level marker extraction algorithm is subsequently used to locate regions that are homogeneous, by moving threshold and extended minima transform. A marker driven watershed transform is then used to get segmented image based on gradient reconstruction. Third, a multi-constraint region merging strategy considering spatial adjacency relation, areas, spectral and textural features is proposed to merge the adjacency region pairs by searching the minimum merging cost among the initial segments. Finally, the detail coefficients are updated and the inverse wavelet transform is computed to project the initial segmentation to higher scale images, and pixels at boundaries are processed to keep region contours as original scale is reached. The experimental results demonstrate that the developed method can be applied to the segmentation of high-resolution multi-spectral remote sensing image as well as alleviate over segmentation and get the high accuracy segmentation.

Key words: wavelet transform, watershed transform, optimal scale, marker location, region merging, result projection

CLC number: TP751.1 **Document code:** A

Citation format: Chen J, Deng M, Xiao P F, Yang M H, Mei X M and Liu H M. 2011. Multi-scale watershed segmentation of high-resolution multi-spectral remote sensing image using wavelet transform. *Journal of Remote Sensing*, **15**(5): 908–926

1 INTRODUCTION

Image segmentation is a bridge between a low-level vision subsystem including image processing operations such as noise reduction, edge extraction and a high-level vision subsystem such as object recognition, image analysis and interpretation such advanced processing (Sprikovska, 1993). In the high-resolution remote sensing image processing, the image segmentation is an important prerequisite for object-based analysis including feature extraction, image classification, object detection and recognition. In addition, to obtain the remote sensing image information by segmentation is an effective means of geographic data updating (Hu, *et al*., 2005), while promoting the full integration and rapid development of GIS and Remote Sensing (Blaschke, *et al*., 2000).

Region-based segmentation methods are generally better than thresholding and feature space clustering approaches since they take into account both feature space and the spatial relation of pixels simultaneously (Hu, *et al*., 2005). Watershed transform takes the gradient magnitude map as undulating terrain, which will be split into different sizes of the catchments basin through the establishment of "watershed line". Currently, the watershed transform has a good performance in segmentation of multi-spectral (Sun & He, 2008; Mao, *et al*., 2009; Xiao, *et al*., 2009) and hyperspectral (Angulo, *et al*., 2009; Tarabalka, *et al*., 2010) remote sensing image. For it is prone to over-segmentation problem, there have been many improvements to this method (Nguyen, *et al*., 2003; Kothainachiar, *et al*., 2006), and generally in the gradient computation, marker extraction and region merging. Multi-scale approach can significantly improve segmentation results (Wang, *et al*., 2001), while a multi-scale watershed segmentation method based on discrete wavelet

Received: 2010-08-17; **Accepted**: 2011-03-28

Foundation: National High Technology Research and Development Program of China(No. 2008AA12Z106); National Natural Science Foundation of China(No. 40801166); Open Foundation of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing(No. 09R03)

First author biography: CHEN Jie(1980—), male, Ph.D. candidate, he is interested in intelligent processing of high resolution remotely sensed imagery, and has 6 research articles published in journals, E-mail: cjcsu@163.com

Corresponding author biography: DENG Min(1974—), male, Ph.D., professor, doctoral adviser, His research interests are spatial data mining, reasoning and analysis. E-mail: dengmin208@tom.com

transform has been recognized by researchers (Jung & Scharcanski, 2000; Jung, 2003; Kim & Kim, 2003).

The segmentation method combining watershed transform with discrete wavelet transform can get good results, but there are still some problems. For example, the existing scale selection method for wavelet decomposition does not apply to remote sensing images, the object markers are without consideration of different geo-objects, and the effect of region merging and result projection are not so good, *etc*. Some improvement measures for watershed segmentation of high-resolution multi-spectral remote sensing image are presented in this paper. Firstly, discrete wavelet transform is implemented to produce multi-scale images, and phase congruency model is applied to extract image gray gradient. Then, to select the optimal scale according to the local variance of the gradient, and the segmentation will be taken from the approximate coefficient at this scale. Finally, marker-based watershed segmentation is implemented based on multi-level markers and gradient reconstruction, and segmentation result of the original image is obtained through the multi-constraint region merging and position-keep boundary processing.

2 SEGMENTATION METHOD

This study on the segmentation method will focus on the following four aspects: (1) the optimal decomposition scale must be identified when generating multi-scale remote sensing images by using wavelet transform; (2) the geo-objects with different structure and distribution in remote sensing image should be taken into account, since object marker is very important to watershed transform; (3) since the ground surface is complex, regional boundaries, areas, spectral and texture features such factors need to be considered when region merging; (4) the location shift of the regional boundaries should be avoided when segmentation result is projected to the higher scale image by inverse wavelet transform. The flow chart of this method is shown in Fig. 1, where the input data are the four bands Quickbird multi-spectral remote sensing images which are the near infrared, the red, the green and the blue, arrows indicate the relationship between the steps of the method, the solid line boxes represent data or operations and the dotted line boxes represent the four main links.

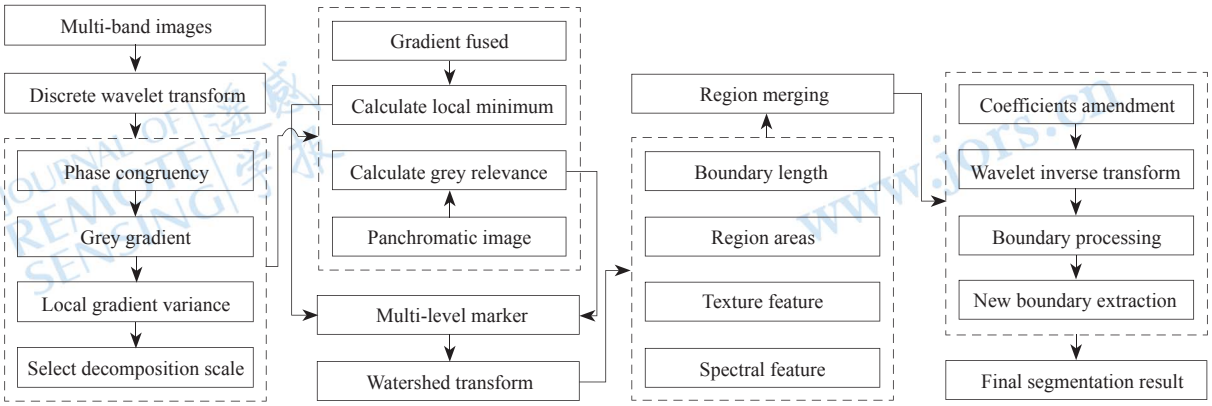

Fig. 1 Flow chart of the method proposed in this paper

2.1 The selection of optimal scale of wavelet decomposition

According to the Matllat algorithm (Mallat, 1989), three high-frequency detail coefficients such as the vertical, the horizontal and the diagonal and a low frequency approximation coefficient are obtained by down sampling using discrete wavelet transform. In this paper, the Haar wavelet transform (Haar, 1911) is used to generate the coefficient sub-images since it has low computational complexity and precise spatial position. Set the decomposition scale as J, the sub-images can be expressed as:

$$(W_J^A, W_J^H, W_J^V, W_J^D, W_{J-1}^H, W_{J-1}^V, W_{J-1}^D, \cdots, W_1^H, W_1^V, W_1^D) \quad (1)$$

where W_J^A is the approximation coefficient; W_J^H is the horizontal coefficient; W_J^V is the vertical coefficient; W_J^D is the diagonal coefficient.

The selection of wavelet decomposition scale J will affect the calculational time and the quantity and quality of the results of watershed transform segmentation. It is generally up to the size and noise level of the image, Jung (2007) proposed a general calculation formula to make a balance between noise reduction and detail maintain, which is expressed as:

$$J = \max\left\{1, 1+\mathrm{rnd}\left(\log 2\left(\frac{\min\{M, N\}}{128}\right)\right)\right\} \quad (2)$$

where rnd( ) is an integral function; M and N are the number of row and column in the image. However, this method can only be applies to the general image while the scale differentiation of various geo-objects in remote sensing image cannot be taken into account. Because the detail information of the geo-objects in different scale images will change accordingly when the remote sensing images are decomposed by using wavelet transform. For example, the geo-object details in the original scale image will disappear and the geo-object outlines are more clearly in the large scale image, and different geo-objects will eventually be mixed up after multiple times wavelet decomposition.

The local variance of the same surface areas is more comparable (Ming, *et al*., 2008), so a wavelet decomposition scale selection method is proposed based on local gradient variance of multi-spectral remote sensing images. Firstly, the original image is of four bands are decomposed by the Haar wavelet to the scale Eq. (2) calculated. For example, if the image size is 1024 × 1024 pixels, then J=4. The scale of the original image is set as 1, thus the wave-

let transform has decomposed 3 times when J=4. Secondly, the phase congruency gradients (Xiao, *et al*., 2007) of 4 bands at each scale are calculated respectively, and the gray gradient amplitudes of different bands are fused by Eq. (3), that is

$$\tilde{\boldsymbol{G}} = \max\left(\boldsymbol{G}_1, \boldsymbol{G}_2, \boldsymbol{G}_3, \boldsymbol{G}_4\right) \tag{3}$$

where $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$, $\boldsymbol{G}_4$ are respectively the phase congruency gradient of 4 bands. In order to make the samples at each scale have the same areas, samplings of these geo-objects are respectively taken with templates of different sizes at the same location in gradient images at each scale. The mean local variances of the geo-objects at each scale are calculated according to the Eq. (4), and the mean local variance of all geo-objects is calculated at each scale from the whole image level. The scale corresponds to minimum mean local variance is taken as the optimal decomposition scale, because the internal homogeneity of the geo-object is the highest at this time, which indicates the gradient magnitude characterize the geo-objects with highest accuracy.

$$\sigma_k^2 = \frac{1}{l}\sum \frac{1}{m \times n}\sum_{x=1}^{m}\sum_{y=1}^{n}[\tilde{\boldsymbol{G}}(x,y) - \bar{\tilde{\boldsymbol{G}}}]^2 \tag{4}$$

where σ_k^2 is the mean local variance of the geo-object k; l is the sample number of the geo-object k; m and n are the length and width of the sampling template; $\tilde{\boldsymbol{G}}(x, y)$ is the gradient amplitude of the pixel at line x and row y calculated by Eq. (3); $\bar{\tilde{\boldsymbol{G}}}$ is the mean gradient in the template.

2.2 Extraction of multi-level markers

The homogeneous region is extracted as the marker from the gradient amplitude image at the optimal wavelet decomposition scale. Based on extended minima transform (Soille, 2003), and combining with moving threshold and regional growth (Hill, *et al*., 2003), multi-level marker extraction method guided by gray relevant is proposed.

Firstly, between the minimum and the maximum of the gradient amplitude image, the threshold sequence $\boldsymbol{H}$ about height is set based on mean and variance, the calculation formula is

$$\begin{cases} \boldsymbol{H} = \text{mean}(\tilde{\boldsymbol{G}}) + \boldsymbol{T}(i) \cdot \text{std}(\tilde{\boldsymbol{G}}) \\ \min(\tilde{\boldsymbol{G}}) < \boldsymbol{H} < \max(\tilde{\boldsymbol{G}}) \end{cases} \tag{5}$$

where mean and std are the mean and variance of the gradient image $\tilde{\boldsymbol{G}}$; $\boldsymbol{T}(i)$is the modulation of the threshold sequence $\boldsymbol{H}$. The size of modulation is set according to the complexity of geo-objects, that is to say, if the geo-objects are of few class and simple structure, a larger move interval is applied, otherwise a smaller move interval applied. To mark all geo-objects as far as possible, maximum marker threshold is set. This threshold takes the areas of a planar geo-object with smooth texture.

Then, markers are extracted from the gradient image successively based on $\boldsymbol{H}$ to get multi-level marker image. All markers in each level is labeled as $\boldsymbol{M}_{q,p}$(q and p are respectively the level and number of $\boldsymbol{M}$). If the central coordinate of a lower level marker is among the coordinate range of an upper level marker, the two markers are "child-parent" related. In this way, an upper marker may be related to several lower markers, but a lower marker can only belong to a certain upper marker or not at all (limited by the maximum marker threshold).

Finally, the relevance of the upper and lower markers is calculated according to the panchromatic image. If the "child-parent" relationship is set up, the gray histogram of the regions corresponding to the two markers in panchromatic image are respectively calculated, and their relevance is

$$R = \frac{n\sum(h_p h_c) - \sum h_p \sum h_c}{\sqrt{n\sum h_p^2 - (\sum h_p)^2}\sqrt{n\sum h_c^2 - (\sum h_c)^2}} \tag{6}$$

where h_p and h_c are respectively the histogram of the two markers; n is the gray level of the histogram. To set the relevance threshold as T_R, and analyze the relevance of the two markers. If the relevance is greater than T_R, the parent marker can represent the child marker; otherwise the child marker is maintained. The markers of all levels which meet above condition are united to get the final marker image.

2.3 Multiple constraints region merging

Gradient image is modified by using minima imposition technology based on multi-level marker image, and the watershed method based on the FIFO queue (Vincent & Soille, 1991) is used to get segmentation. From the result, we can find that the adjacent homogeneous regions may belong to a same class and they are similar in spectral and texture to some extent. Therefore, some rules need to be established for merging these regions. Since the geo-objects in high-resolution remote sensing images are of clear feature and complex relationship, the spectral and texture features should be taken into account in region merging. The texture features are extracted by using Gabor wavelet, which is expressed as (Petkov & Kruizinga, 1997)

$$h_{\theta,\lambda,\varphi,\sigma,\gamma}(x,y) = \exp\left(-\frac{1}{2}\frac{(x')^2 + (\gamma \cdot y')^2}{\sigma^2}\right) \cdot \cos\left(2\pi\frac{x'}{\lambda} + \varphi\right) \tag{7}$$

where $x' = x\cdot\cos\theta + y\cdot\sin\theta$, $y' = -x\cdot\sin\theta + y\cdot\cos\theta$; x and y are coordinates of pixel location in spatial domain; λ and θ are the scale and direction of Gabor function; φ is the phase angle; σ is the Gaussian radius; γ is the Gaussian envelope aspect ratio.

Set A, B as two adjacent regions, merger cost function in following form is used

$$\xi = \frac{L_A \cdot L_B}{L_{AB}^2}\frac{S_A \cdot S_B}{S_A + S_B}\left[\alpha\Delta\left\{F_A^S, F_B^S\right\} + (1-\alpha)\Delta\left\{F_A^T, F_B^T\right\}\right] \tag{8}$$

where L_A, L_B and L_{AB} are respectively the perimeter and common boundary length of region A, B; S_A and S_B are the areas of region A, B; α is the weight of spectral feature; F^S and F^T are the mean value of spectral and texture feature of region A or B ; $\Delta\{\cdot,\cdot\}$ is the Euclidean distance between the mean values of region A and B. Eq. (8) shows that, when the common boundary between the region A and region B counts for a larger proportion of their respective perimeter, their respective areas is small and their spectral and texture feature are to be similar, the merging cost would be relatively small and the merging would be more reasonable. The merging of adjacent regions will take following steps.

Step 1 To label all regions with serial numbers and calculate the mean value of feature vector consist of the spectral and Gabor texture for every region.

Step 2 To make regional adjacency graph and obtain the adjacent regions pairs, and extract the areas, perimeter of all regions and the length of common boundary between the adjacent regions.

Step 3 To calculate the merging cost of all adjacent regions and sort the costs to find the two regions corresponding to the minimum value.

Step 4 Let region A is labeled behind region B, to make region merging and data update as follows:

- Add up the perimeter of region A and B and minus twice the length of the common boundary to get a new perimeter of region B.
- Add the areas of region A and B with the length of the common boundary to get the new areas of region B.
- Turn the adjacent relationship of region A to region B.
- Turn all common boundaries relevant to region A to region B.
- Merge up region A and B and the serial number of region B maintained.

Step 5 To make cycle calculation following step 3 and step 4 until the following two stopping threshold are satisfied: the maximum region areas and the minimum feature Euclidean distance. The former is for region merging costs reduction, the latter highlights the importance of spectral and texture feature.

2.4 Projection of segmentation result

Since the wavelet decomposition is down sampling based on a multiple of 2, the segmentation results after merging need to be stepwise projected to the original image by the wavelet inverse transform. Jung (2003) processed the boundary in projection according to the smallest gray difference between boundary pixels and their neighbor pixels. However, the calculated amount of this method is large, and the boundary tends to get a sharp displacement after several times of projection. In order to keep the border position of the region in the projection, a sliding template process approach is proposed in the following (just one band image needed). The wavelet decomposition coefficients at the scale corresponding to the segmentation results are amended as

$$NW_j^l(n,m)=\begin{cases}0 & if\ (n,m)\ \ belongs\ to\ a\ region\ border\\ W_j^l(n,m) & otherwise\end{cases} \tag{9}$$

where l is the approximate and detail coefficients. Eq. (9) shows that the pixels of the approximation and detail coefficients at boundary are 0, but the other pixels retain the original value.

Fig. 2 shows the boundary processing. First, the boundaries are projected to the approximate coefficient at higher scale by the wavelet inverse transform, and their width are of 2 × 2 as a result of the Haar wavelet (Fig. 2(a)). Second, every region is labeled and sorted (Fig. 2(b)), then a 3 × 3 template (the shadow in Fig. 2(b)) slides along the boundary (the bold 0 in Fig. 2(b)), and the mean value of pixels in the template is taken into the corresponding position in a memory. Finally, a new region (Fig. 2(c)) is get after adding up Fig. 2b and the memory. It should be noted not the approximate coefficient after calculating the mean value of every region but the labeled image is used in this boundary processing method, because the mean value of adjacent region could be the same in the former case. Fig. 2c shows that the pixels in boundaries are labeled with the value of the adjacent regions, so the differences between adjacent regions is more direct and simple than they are in Fig. 2a.

Roberts operator is used to generate new boundaries. Because 2 × 2 convolution template is used in the operator to take the adjacent pixels differences in the diagonal direction to detect the edge, the gradient detection effect is better and the positioning accuracy is higher in the horizontal and vertical direction, and there is little noise impact in boundary processing. The new boundaries will appear in zero tracks as shown in Fig. 2b, and their routes coincide with the initial boundaries. In addition, single pixel fracture will appear in common boundary of three or more regions, which can be linked through the process as shown in Fig. 3. Fig. 3a is the endpoint detection template with width of 3 × 3 in which the two double solid line circles are the existing pixels representing the broken line with zero direction, dotted line circle indicates there is a boundary pixel in front of the broken line, the single solid line circle is the filled boundary pixel, the blank boxes indicate that there is no boundary pixel there. The template can be adjusted to deal with the broken lines with 90°, 180°, 270° direction. The boundary as shown in Fig. 3b is extracted from Fig. 2c by using Roberts operator, and the circles is the filled pixels by this way.

91	86	82	84	90	86	0	0	101	103
88	87	87	88	86	87	0	0	111	109
86	87	0	0	0	0	0	0	108	105
82	88	0	0	0	0	0	0	114	110
93	87	0	0	108	112	108	102	112	109
88	84	0	0	114	110	107	105	112	110
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
47	44	0	0	212	218	208	213	215	220
45	48	0	0	215	217	216	221	209	211

(a)

1	1	1	1	1	1	0	0	2	2
1	1	1	1	1	1	0	0	2	2
1	1	**0**	0	0	0	0	0	2	2
1	1	0	0	0	0	0	0	2	2
1	1	0	0	2	2	2	2	2	2
1	1	0	0	2	2	2	2	2	2
0	0	0	0	**0**	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
3	3	0	0	4	4	4	4	4	4
3	3	0	0	4	4	4	4	4	4

(b)

1	1	1	1	1	1	1	2	2	2
1	1	1	1	1	1	1	2	2	2
1	1	1	1	1	1	1	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
3	3	3	4	4	4	4	4	4	4
3	3	3	4	4	4	4	4	4	4
3	3	3	4	4	4	4	4	4	4

(c)

Fig. 2 Processing of boundary

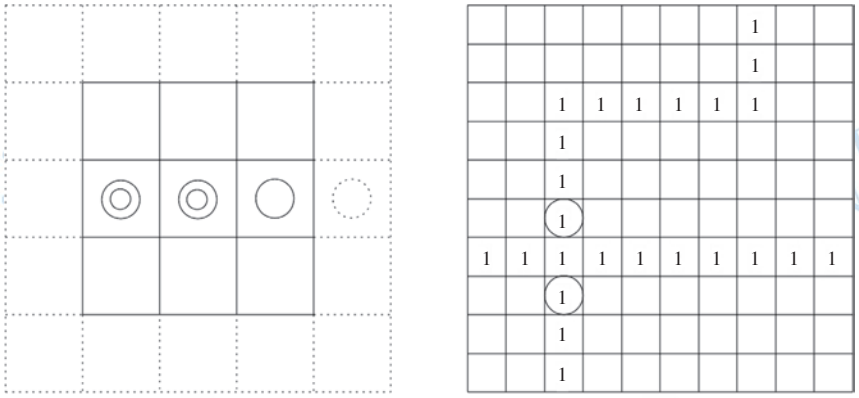

Fig. 3 Linking of broken boundary
(a) The end point detection template; (b) The boundary

3 EXPERIMENT AND RESULT ANALYSIS

The Experimental data is the Quickbird image has been geometric and ortho corrected with the time of November 21, 2004, the location of the Jiangning District of Nanjing city Jiangsu province China, the size of 1024 × 1024 pixels. There are blue, green, red, near infrared spectral bands and panchromatic data, and their spatial resolution is respectively 2.44 m and 0.61 m. To make the resolution of multi-spectral bands data become 0.61 m after fused with panchromatic data by using PANSHARP method (Zhang, 2002) respectively. Fig. 4 is a false color image combined of the infrared, red and green three bands, in which some typical geo-objects, including road (C1), stadium (C2), water (C3, C4), housing (C5, C9), farmland (C6, C7, C10, C11, C12), tree (C8) are labeled with circles and numbers. As the crops have been harvested, the most widely distributed farmland is of diverse spectral and texture feature. Housing is concentrated but of various shapes and complex textures. The up part of road is of outstanding spectral and smooth texture, but the other part of road and the house roadside are almost of the same spectral. Water is of monotonic texture but different spectral feature because of the different substances contained in it. Trees, houses and water are distributed on the bottom of the image. Experimental platform is the Windows XP operating system, CPU clocked at 2.4 GHz, memory is 2 G, and Matlab programming is used.

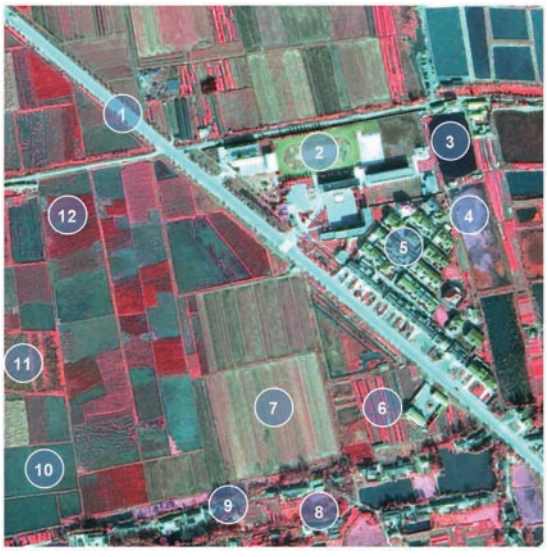

Fig. 4 False color image

3.1 Segmentation using the proposed method

(1) Select the optimal wavelet decomposition scale

Five samples of every geo-object corresponding to twelve classes are used to determine the optimal wavelet decomposition scale. Since the maximum decomposition scale J equals 4 (calculation result of the Eq. (2)), the samples are respectively selected by adopting templates with width of 4×4, 8×8, 16×16, 32×32 from small scale to large one. Phase congruency gradient for approximate coefficient of 4 bands at four scales are calculated. The parameters are set as: wavelet scale n is 5; orientation o is 6 $\left(\text{viz } 0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}\right)$; wavelength of smallest scale filter is three pixels; scaling factor between successive filters is 2.1; noise threshold point T is 0.2; cutoff frequency c is 0.05; index for controlling the sharpness of the transition in the sigmoid function g is 15; ε equals to 0.0001. Fig. 5 is the local gradient variance at the four scales. It shows that class C7 and C8 are of sustained decline; class C1, C10 and C11 are of sustained increase; class C2, C3 and C5 are of first increases and then decreases; class C4, C6, C9 and C12 are of first decreases and then increases. Fig. 6 is the average local variance of gradient at four scales. The "V" curve indicates the minimum value is got at scale 2 which is the optimal wavelet decomposition scale.

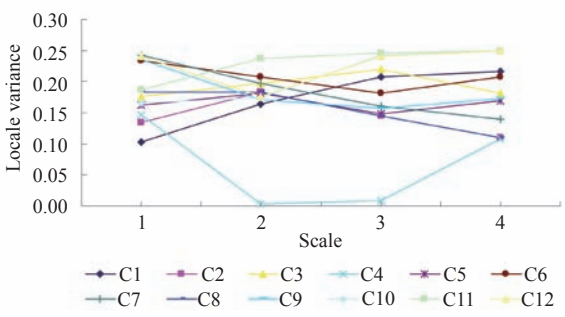

Fig. 5 Local variance of gradient

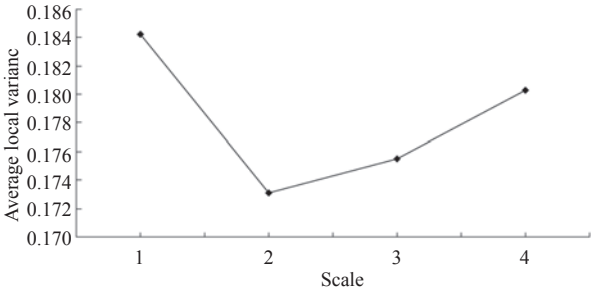

Fig. 6 Average variance of gradient

(2) Multi-level marker extraction

Gradient images of the four bands at scale 2 are fused using Eq. (3). As shown in Fig. 7, all geo-objects are clearly outlined. The mean and variance of this gradient image are 0.2474 and 0.1721 respectively, and the modulation value is set as –0.5, –0.4, –0.3, –0.2, –0.1, 0.0, 0.1, 0.2, 0.3, 0.4 and 0.5, thus the threshold sequence 0.1614, 0.1786, 0.1958, 0.2130, 0.2302, 0.2474, 0.2646, 0.2818, 0.2990, 0.3162 and 0.3334 is calculated using Eq. (5). It can be seen from Fig. 4 that the road is of the largest areas and smooth texture, but the spectral of lower part of road and the roadside house are almost the same, so the areas of up part of road is 3521 which set as the maximum marker threshold. Since only the maximum marker threshold is set, details of geo-objects can appear in the initial segmentation results, especially the geo-object with complex texture and surroundings. To establish the "child-parent" relationship and set the correlation threshold T_R as 0.85, then compose all markers meeting the condition to get the marker image shown in Fig. 8. It can be seen from Fig. 8 all different sizes geo-objects are taken into account by multi-level markers. For example, road, water and farmland which are of monotonic texture and housing, tree and open space which are of complex texture are all reflected in marker image. Initial segmentation result of 577 regions at scale 2 is obtained by using watershed transform. The Fig. 9 shows that watershed lines are congruent with all geo-objects, especially the geo-objects which are of uniform texture edges such as farmland, road and water, and the geo-objects with complex texture are represented by the smaller regions.

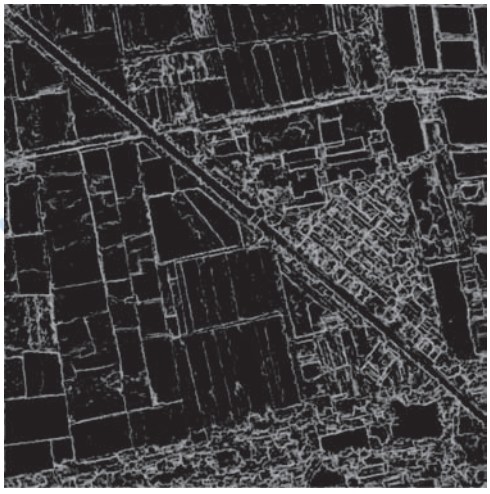

Fig. 7 Gradient image after fusion

Fig. 8 Marker image

(3) Region merging and result projection

To calculate the boundaries, areas, normalized spectral and texture feature of regions and implement region merging in accordance with the steps described in Section 2.3. Parameters of Gabor wavelet are set for the texture feature extraction as γ=1, b=1.3, φ=0, $\theta=\left\{\frac{\pi}{6},\frac{\pi}{3},\frac{\pi}{2},\frac{2\pi}{3},\frac{5\pi}{6},\pi\right\}$, and λ taking the first four scales. Considering the space constraints, the basis of the above parameters setting please refer to the literature (Chen, *et al*., 2010). Region merging costs of all adjacent regions are calculated using Eq. (8). The maximum region areas T_M is set as 380 and the minimum feature Euclidean distance as 0.35. The former should be set to be small and the latter to be large in order to promote the adjacent regions with small areas can be put into merging queue preferentially. As the texture feature is better than the spectral feature for geo-objects separating, the weight for spectral feature distance is set as 0.4 and the weight for texture feature distance as 0.6. Fig. 10 shows the merging result of 310 regions. From it, we can see that the small regions corresponding to the geo-object details almost have been merged correctly. Despite the adjacent regions merging of the same class (for example, farmland in the lower part of the image and housing patch above the road) are hindered by maximum region areas threshold, the result can meet the requirements of image segmentation. Then, according to the method described in Section 2.4, the merging result in Fig. 10 is projected onto the original image. As the optimal decomposition scale is 2, only one time projection is needed. When Roberts operator is used to produce new boundaries, the threshold value is set as 0.000001. The result is shown in Fig. 11 with the boundaries consistent with Fig. 10.

(4) Computational time analysis

In this proposed method, remote sensing image segmentation is carried out based on multi-scale decomposition of wavelet transform, including the multi-scale images representation, image segmentation, region merging and result projection. Since it is a supervised segmentation method, the whole process requires a small amount of human-machine interactivity. For instance, the selection of wavelet decomposition scale requires samples selection in pertinence, the extraction of multi-level marker needs to consider the structure and distribution characteristics of geo-objects to set the modulation and the maximum marker threshold. In the segmen-

tation process, the amount of computing will be affected by the size of image, number of bands and the complexity of image content. Among them, image size and the number of bands can impact the computational time for phase gradient and gradient fuse, content of image can affect the size and number of markers and the number of adjacent region pairs such several important factors determining the computation. In addition, the region merging takes up the longest computational time relatively, but it will be reduced in some extent by setting the maximum region areas threshold.

3.2 Segmentation based on other method

Jung (2007) proposed an unsupervised multi-scale segmentation (UMS) method combining the discrete wavelet with the watershed transform. Segmentation experiments of several natural color images showed that the results of this method were better than Statistical Region Merging (SRM), JSEG, Mean-shift and other methods. Now, UMS method is used for segmentation of the same remote sensing image. After several tests, the best segmentation result is obtained with 285 regions when the wavelet decomposition scale is set as 4 and merger threshold is set as 0.05. As shown in Fig. 12, geo-objects of concentrated distribution and smooth texture such as stadium, water, housing, farmland and road are all well segmented. However, adjacent regions with similar color can not be distinguished due to spectral is considered only, so large areas farmlands and broken road appear in the result. In addition, as there is no restriction on the maximum region areas threshold, one region may cover several geo-objects in the merging results. The biggest limitation of the method is, the minimum distance discrimination method is used to process boundaries for segmentation results projection, which consequently make the regional boundaries direction random after times projection. For example, the initial segmentation boundaries of the water, road and farmland are not maintained regular in the final results. On the other hand, this method is better than the proposed one in shape integrity of geo-objects (such as the patch of houses is totally segmented out), and less computing time due to the unsupervised segmentation process.

3.3 Quantitative comparative analysis of segmentation

Object-level consistency error (OCE) is introduced as the image segmentation criteria (Polak, *et al*., 2009) to evaluate the segmentation results of the two methods in quantity. This criterion can detect the shape, location, size and existence of the segment region, is

Fig. 9 Initial segmentation

Fig. 10 Merged result

Fig. 11 The result by our proposed method

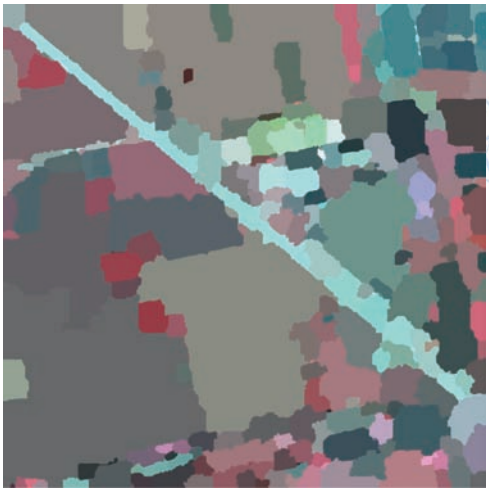

Fig. 12 The result by unsupervised multi-scale segmentation

sensitive to both over-segmentation and under-segmentation, and is of symmetry and scale invariance. OCE meet $0 \leqslant \mathrm{OCE}(I_g, I_s) \leqslant 1$, where I_g is the reference segmentation, I_s is the actual segmentation. If OCE=0, the actual segmentation is the same with the reference segmentation. Since the remote sensing image used in this paper is different with its actual ground situation and the image can reflect the details of most geo-objects with the spatial resolution of 0.61 m, the reference segmentation image for accuracy evaluation will be obtained by the visual interpretation. First, reference segmentation is got by original image interpretation, then the reference segmentation is rasterized, the image with consecutive numbered regions is obtained finally (Fig. 13). The high-resolution remote sensing images is of rich details, complex geo-objects relationships, diverse texture, and the same segmentation method and parameters will lead different effects for segmenting the different geo-objects, so differentiated evaluation is necessary. The OCE is based on the object, so it is feasible to organize the regions into a number of patches according to the concentrative degree and similarity of spectral and texture features of geo-objects. It can be seen form Fig. 13

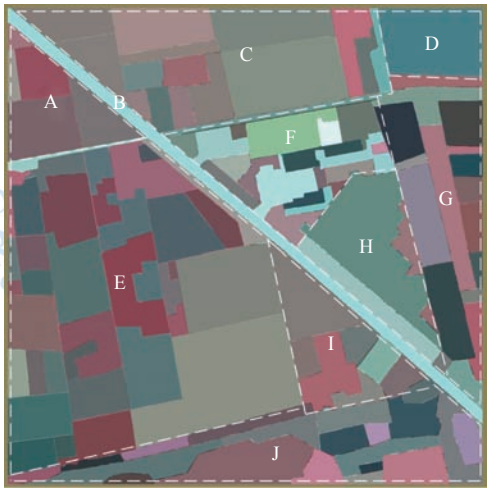

Fig. 13 Reference image for evaluating segmentation

that A, C, E are the concentrated farmland areas with clear outline, B is the road with clear shape, D is the water, F is the housing, stadiums, open space, trees and other geo-objects, G is mainly the water with houses and farmland interspersed, H is the concentrated housing, I is mainly the farmland with some separate buildings, J is the diverse geo-objects with complicated relationship.

Table 1 makes comparisons of the two segmentation methods in the number of regions and the OCE value according to the patches in the reference segmentation. Since there are segmented regions across two or more patches, the region number corresponding to the proposed method is 103 more than its total segmented regions and UMS is 105 more, which indicates the two methods are of similar regional duplication degree. Specifically, region number in the proposed method is less than UMS in the A, B, D, F, G, I and J patches and more in the C, E and H, which indicates the region dispersion of UMS is over it in the proposed method. On the other hand, OCE value in the proposed method is smaller than UMS in the A, B, E, F, H, I and J patches and lager in the C, D and G patches, which indicates the segmented regions of the proposed method is more coincidence with the reference segmentation. For the road patch B, the segmented regions of UMS are more discrete than the proposed method. For the water patch D, the segmented regions of UMS are more concentrated than the proposed method in which there are segmented regions across two or more patches. It can be found by combining the region number and OCE of the F, G, H, I and J patches that although a small region number correspond to a small OCE value, the region number is not positively corresponding to OCE value. On the whole, the OCE value of the proposed method is 4.8445 which is 0.2568 less than the value of UMS 5.1013, indicating that the former can get a better segmentation. It can be known from the comparison of the advantages and disadvantages of the two methods that small regional dispersion, centralized distribution and the exact boundaries is the outward manifestation of the high segmentation accuracy, while it is difficult for few segmentation rules to take the geo-objects of different spectral, texture, shape, size, class in remote sensing images all into accounts.

Table 1 Comparison of two segmentation methods

		A	B	C	D	E	F	G	H	I	J
The proposed method	Region number	14	26	55	8	82	46	48	46	27	61
	OCE	0.4810	0.4590	0.5231	0.5160	0.4090	0.4605	0.5217	0.4877	0.4733	0.5132
UMS	Region number	20	36	31	21	41	62	52	28	29	70
	OCE	0.5009	0.5289	0.5018	0.5014	0.4971	0.5100	0.5060	0.5087	0.4987	0.5478

4 CONCLUSION AND FUTURE WORKS

Watershed transform is combined with discrete wavelet transform in this paper to develop the advantages of multi-scale segmentation, and improvement methods are proposed from different angles. It is an important prerequisite for the watershed transform using phase congruency gradient to capture the boundaries accurately. Multi-band information is fully used by fusing the gray gradient of different bands images and multi-scale images are produced more purposefully by analyzing the local gradient variance to select the optimal wavelet decomposition scale. When the minimum local variance is taken as a basis for the optimal scale selection, the spectral and texture feature at this scale can be more separable, thus a support for the region merging. To establish the relationship between the upper and lower markers guided by the gray correlation is reasonable and the multi-level markers can take different geo-objects into account, thereby the watershed transform segmentation effect is improved. The Gabor texture feature, region areas and spatial adjacency relation of regions are taken as the con-

straints to improve the regions merging accuracy. The inverse wavelet transform has provided a reliable guarantee for the results projection, and the proposed boundary processing method makes the segmentation results projected to the original image successfully.

Although the experiment results show that the proposed method can achieve good segmentation results, there are still some inadequacies. For example, many parameters are involved in the gradient and texture extraction and thresholds of markers extraction and region merging require many times debugging. Further work will be the analysis of the impact of different parameters on the segmentation results, and the introduction of other technical methods to further improve this method.

REFERENCES

Angulo J, Velasco-Forero S and Chanussot J. 2009. Multiscale stochastic watershed for unsupervised hyperspectral image segmentation. IEEE International Geoscience and Remote Sensing Symposium: 93–96

Blaschke T, Lang S, Lorup E, Strobl J and Zeil P. 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. *Environmental Information for Planning*, **2**: 555–570

Chen J, Deng M, Xiao P F and Yang M H. 2010. Texture segmentation based on Gabor filters and ant colony optimization algorithm. *Geomatics and Information Science of Wuhan University*, **35**(11): 1271–1274, 1334

Haar A. 1911. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, **71**(1): 38–53 DOI: 10.1007/BF01456927

Hill P R, Canagarajah C N and Bull D R. 2003. Image segmentation using a texture gradient based watershed transform. *IEEE Transactions on Image Processing*, **12**(12): 1618–1633 DOI: 10.1109/TIP.2003.819311

Hu X Y, Tao C V and Prenzel B. 2005. Automatic segmentation of high-resolution satellite imagery by integrating texture, intensity, and color features. *Photogrammetric Engineering and Remote Sensing*, **71**: 1399–1406

Jung C R. 2003. Multiscale image segmentation using wavelets and watersheds. Proceedings of the XVI Brazilian Symposium on Computer Graphics and Image Processing: 278–286

Jung C R. 2007. Unsupervised multiscale segmentation of color images. *Pattern Recognition Letters*, **28**(4): 523–533 DOI: 10.1016/j.patrec.2006.10.001

Jung C R and Scharcanski J. 2000. Robust watershed segmentation using the wavelet transform. Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing: 131–137

Kim J B and Kim H J. 2003. Multiresolution-based watersheds for efficient image segmentation. *Pattern Recognition Letters*, **24**(1-3): 473–488 DOI: 10.1016/S0167-8655(02)00270-2

Kothainachiar S, Wahita Banu R S D and Saravanan A. 2006. Unsupervised morphological segmentation for textured and non-textured images. *The International Journal on Graphics, Vision and Image Processing*, **6**(2): 33–39

Mallat S G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7): 674–693 DOI: 10.1109/34.192463

Mao F, Liu Z, Zhou W S and Li Q. 2009. Extracting of urban features from high resolution remote sensing data based on multiscale segmentation. Proceedings of the 2009 Urban Remote Sensing Joint Event

Ming D P, Wang Q and Yang J Y. 2008. Spatial scale of remote sensing image and selection of optimal spatial resolution. *Journal of Remote Sensing*, **12**(4): 529–537

Nguyen H T, Worring M and van den Boomgard R. 2003. Watersnakes: energy-driven watershed segmentation. *IEEE transactions on Pattern Analysis and machine Intelligence*, **25**(3): 330–342 DOI: 10.1109/TPAMI.2003.1182096

Petkov N and Kruizinga P. 1997. Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. Biological Cybernetics: 83–96

Polak M, Zhang H and Pi M H. 2009. An evaluation metric for image segmentation of multiple objects. *Image and Vision Computing*, **27**(8): 1223–1227. DOI: 10.1016/j.imavis.2008.09.008

Soille P. 2003. Morphological Image Analysis: Principles and Applications. 2nd ed. New York: Springer Verlag

Spirkovska L. 1993. A summary of image segmentation techniques. NASA Technical Memorandum: 1–11

Sun Y and He G J. 2008. Segmentation of high-resolution remote sensing image based on marker-based watershed algorithm. The 5th International Conference on Fuzzy Systems and Knowledge Discovery: 271–276

Tarabalka Y, Chanussot J, Benediktsson J A, Angulo J and Fauvel M. 2010. Segmentation and classification of hyperspectral data using watershed. *Pattern Recognition*, **43**(7): 2367–2379

Vincent L and Soille P. 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**(6): 583–598 DOI: 10.1109/34.87344

Wang J Z, Li J, Gray R M and Wiederhold G. 2001. Unsupervised multiresolution segmentation for images with low depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(1): 85–90 DOI: 10.1109/34.899949

Xiao P F, Feng X Z and Li H. 2009. Multispectral remotely sensed imagery segmentation based on first fundermental form. Proceedings of the 2009 Urban Remote Sensing Joint Event

Xiao P F, Feng X Z, Zhao S H, Deng M and She J F. 2007. Feature detection from high-resolution remotely sensed imagery based on phase congruency. *Journal of Remote Sensing*, **11**(3): 303–310

Zhang Y. 2002. Problems in the fusion of commercial high-resolution satellite images as well as Landsat7 images and initial solutions. *International Archives of Photogrammetry and Remote Sensing*, **34**: 236–242

利用小波变换的高分辨率多光谱遥感图像多尺度分水岭分割

陈杰[1]，邓敏[1]，肖鹏峰[2]，杨敏华[1]，梅小明[1]，刘慧敏[1]

1. 中南大学 测绘与国土信息工程系，湖南 长沙，410083；
2. 南京大学 地理信息科学系，南京，210093

摘　要：为了减少仅用分水岭变换而导致的过分割问题，本文提出利用小波变换的多尺度处理方式用于融合后多光谱QuickBird图像的分割。整个分割过程包括多尺度图像表示、图像分割、区域合并和结果映射等过程。首先，依据原始图像的大小确定分解尺度并用小波变换产生各波段的低尺度图像。采用相位一致模型提取各近似系数的梯度，并逐尺度地融合各梯度图。分析不同尺度下的不同地物的局部梯度方差，以选择最佳的小波分解尺度。然后，通过移动阈值与扩展最小变换，利用多层次标记提取方法标记均质区域。进而，在梯度重建的基础上利用标记分水岭变换得到分割图像。其次，采取空间相邻关系、面积、光谱与纹理等多约束策略，以搜索最小合并代价的方式合并最初分割区域中的邻接区域对。最后，修改细节子图并进行小波逆变换将最初分割结果投影到更高尺度图像，同时处理边界上的像元以保持区域边界直至原始图像。实验结果表明本文方法不仅能够用于高分辨率多光谱遥感图像的分割，而且缓解了过分割问题且取得了较准确的分割效果。

关键词：小波变换，分水岭变换，最佳尺度，标记提取，区域合并，结果投影

中图分类号：TP751.1　　**文献标志码**：A

引用格式：陈杰，邓敏，肖鹏峰，杨敏华，梅小明，刘慧敏. 2011. 利用小波变换的高分辨率多光谱遥感图像多尺度分水岭分割. 遥感学报，**15**(5): 908–926

Chen J, Deng M, Xiao P F, Yang M H, Mei X M and Liu H M. 2011. Multi-scale watershed segmentation of high-resolution multi-spectral remote sensing image using wavelet transform. *Journal of Remote Sensing*, **15**(5): 908–926

1　引　言

分割是连接去噪、边缘检测等低级图像视觉处理与目标识别、图像分析和解译等高级图像视觉处理的桥梁(Sprikovska，1993)。在高分辨率遥感图像处理中，图像分割是包括特征提取、图像分类、地物探测与识别等在内的面向对象分析的重要前提。此外，通过分割获取遥感图像信息是地理数据更新的有效手段(Hu 等，2005)，同时促进了GIS与遥感全面结合和快速发展(Blaschke 等，2000)。

基于区域的分割方法兼顾了像元的特征相似性与空间相关性，其效果优于阈值法与聚类法(Hu等，2005)。分水岭变换将梯度幅值图视作高低起伏的地形，通过建立“分水脊线”将地形表面分割成大小不同的集水盆地。目前，分水岭变换在多光谱遥感图像(Sun和He，2008；Mao 等，2009；Xiao 等，2009)与高光谱遥感图像(Angulo 等，2009；Tarabalka 等，2010)的分割中有着出色的表现。针对其容易出现过分割的问题，已出现较多改进的方法(Nguyen 等，2003；Kothainachiar 等，2006)，且一般是在梯度计算、标记提取与区域合并等环节进行改善。多尺度处理方式能够明显改善图像分割的效果

收稿日期：2010-08-17；**修订日期**：2011-03-28

基金项目：国家高技术研究发展计划(863计划)(编号：2008AA12Z106)；国家自然科学基金(编号：40801166)；测绘遥感信息工程国家重点实验室开放基金项目(编号：09R03)

第一作者简介：陈杰(1980—)，男，博士研究生，主要从事遥感影像智能处理方面的研究，已发表论文6篇。E-mail: cjcsu@163.com。

通信作者简介：邓敏(1974—)，男，博士，教授，博士生导师，现从事空间数据挖掘、推理与分析方面的研究。E-mail: dengmin208@tom.com。

(Wang 等，2001)，基于离散小波变换的多尺度分水岭分割方法已得到研究者们的认可(Jung和Scharcanski，2000；Jung，2003；Kim和Kim，2003)。

结合分水岭变换与离散小波变换的分割方法虽然可取得较好结果，但仍然存在一些问题。例如，现有小波变换的分解尺度选择方法并不适用于遥感图像，目标标记没有兼顾不同地物，区域合并与结果投影效果不佳等。本文提出一些改进措施用于高分辨率多光谱遥感图像的分水岭分割。首先，用离散小波变换产生多尺度图像，利用相位一致算子提取图像的灰度梯度。然后，根据图像梯度的局部方差选择最佳分解尺度，对该尺度下的近似系数进行图像分割。最后，基于多层次目标标识和梯度重建进行分水岭分割，通过多约束的区域合并与位置保持的边界处理得到原始图像的分割结果。

2 分割方法

本文分割方法将围绕以下4个方面的问题进行研究：(1)利用小波变换产生多尺度遥感图像时需要确认哪个尺度为最佳分解尺度；(2)目标标识对于分水岭变换十分重要，应综合考虑遥感图像上的结构与分布各异的地物目标；(3)鉴于地表的复杂性，合并区域时需要顾及区域的边界、面积以及光谱与纹理特征等因素；(4)由逆小波变换向上级尺度图像作分割结果投影时，需尽量避免区域边界位置的偏移。本文方法的框架如图1所示。其中，输入数据为近红外、红、绿、蓝4个波段的Quickbird多光谱遥感图像，箭头表示步骤间的关系，实线方框表示数据或者操作，虚线框表示4个主要的中间环节。

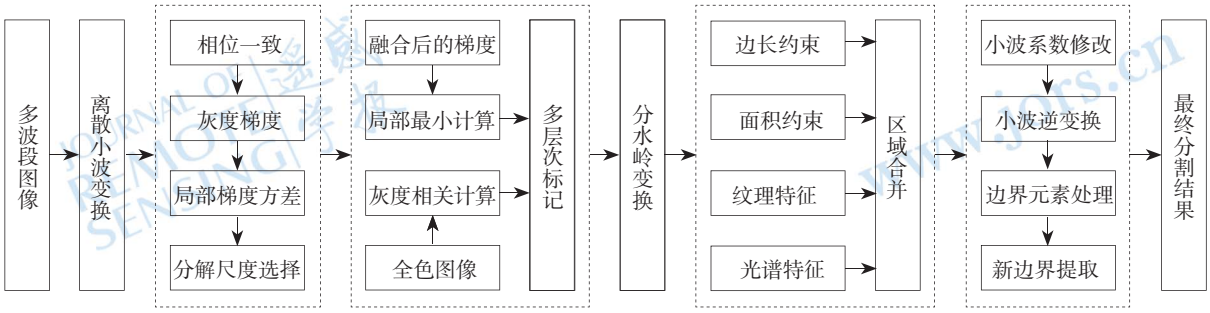

图1 本文方法框架

2.1 最佳小波分解尺度选择

根据Matllat算法(Mallat，1989)，利用离散小波变换向下采样得到垂直系数、水平系数、对角系数等3个高频细节系数与1个低频近似系数。本文采用计算复杂度较低且空间定位准确的Haar小波变换(Haar，1911)产生系数子图。设分解尺度为J，各子图可表示为：

$$(W_J^A, W_J^H, W_J^V, W_J^D, W_{J-1}^H, W_{J-1}^V, W_{J-1}^D, \cdots, W_1^H, W_1^V, W_1^D) \quad (1)$$

式中，W_J^A为近似系数；W_J^H为水平系数；W_J^V为垂直系数；W_J^D为对角系数。

小波分解尺度J的选择会影响分水岭变换分割的计算时间与所得区域的数量和质量。一般可根据图像的大小与噪声程度来确定，Jung(2007)提出了能兼顾噪声去除与细节保持的通用计算式，表达为：

$$J = \max\left\{1, 1+\mathrm{rnd}\left(\log 2\left(\frac{\min\{M, N\}}{128}\right)\right)\right\} \quad (2)$$

式中，rnd( )为取整函数；M与N为图像的行列数。但是，该方法仅适用于一般图像而无法顾及遥感图像中各类地物的尺度差异性。因为在用小波变换分解遥感图像后，不同尺度图像中地物信息的详细程度会发生相应变化。比如，原始尺度图像中的地物细节会在大尺度下消失而地物轮廓却更加清晰，不同地物在小波的多次分解后最终可能会彼此混淆。

由于相同地面面积的局部方差更具可比性(明冬萍等，2008)，本文基于多光谱遥感图像的局部梯度方差提出一种小波分解尺度选择方法。首先，用Haar小波分解4个波段的原始图像直至满足式(2)计算的尺度。比如，若图像大小为1024 × 1024像元，则J=4。

本文设原始影像对应的尺度为1，那么J=4时小波变换进行了3次分解。然后，分别计算4个波段在各尺度下的相位一致梯度(肖鹏峰 等，2007)，采用式(3)融合不同波段的灰度梯度幅值，即：

$$\tilde{\boldsymbol{G}}=\max\left(\boldsymbol{G}_1,\boldsymbol{G}_2,\boldsymbol{G}_3,\boldsymbol{G}_4\right) \tag{3}$$

式中，$\boldsymbol{G}_1$，$\boldsymbol{G}_2$，$\boldsymbol{G}_3$，$\boldsymbol{G}_4$分别为4个波段的相位一致梯度。为使各尺度下的样本具有相同的地面面积，在各尺度梯度图的相同位置用不同大小的模板对各类地物分别采样。用式(4)计算各尺度下地物的局部方差均值，并从整幅图像的层面计算各尺度下所有地类的局部方差均值。取最小局部方差均值对应的尺度为最佳分解尺度，因为此时的地物内部均质度最高，表示梯度幅值对该地物的刻画最为准确。

$$\sigma_k^2=\frac{1}{l}\sum\frac{1}{m\times n}\sum_{x=1}^{m}\sum_{y=1}^{n}[\tilde{\boldsymbol{G}}(x,y)-\bar{\tilde{\boldsymbol{G}}}]^2 \tag{4}$$

式中，σ_k^2为第k种地物的局部方差均值；l为第k种地物的样本数量；m与n为采样模板的长宽；$\tilde{\boldsymbol{G}}(x,y)$为由式(3)计算得到的第$x$行、第$y$列像元的梯度值；$\bar{\tilde{\boldsymbol{G}}}$为模板内的梯度均值。

2.2 多层次标记提取

从最佳小波分解尺度对应的灰度梯度幅值图，提取同质区域作为目标标记。基于扩展最小变换(Soille，2003)，并结合移动阈值与区域增长(Hill等，2003)，本文提出灰度相关性引导的多层标记提取方法。

首先，在梯度幅值图的最小值与最大值间的范围内，基于均值与方差设定高度阈值序列$\boldsymbol{H}$，计算式为：

$$\begin{cases}\boldsymbol{H}=\text{mean}(\tilde{\boldsymbol{G}})+\boldsymbol{T}(i)\cdot\text{std}(\tilde{\boldsymbol{G}})\\ \min(\tilde{\boldsymbol{G}})<\boldsymbol{H}<\max(\tilde{\boldsymbol{G}})\end{cases} \tag{5}$$

式中，mean与std分别为梯度图$\tilde{\boldsymbol{G}}$的均值与方差；$\boldsymbol{T}(i)$为高度阈值序列$\boldsymbol{H}$的调制。调制的大小依据地物的复杂程度设置，若地物类型较少且结构简单则采用较大移动间隔，若地物类型较多且结构复杂则采用较小移动间隔。为尽量标记出所有地物，仅设置最大标记阈值。该阈值取某光滑纹理面状地物的面积。

然后，根据高度阈值序列$\boldsymbol{H}$，依次从梯度图像中提取标记以得到多层次标记图像。并对各层中的标记进行编号，记为$\boldsymbol{M}_{q,p}$(q与p分别标识M的层次与编号)。进而，判断下层标记的质心坐标是否在上层某个标记的坐标范围内。如果是，则在下、上层标记间建立“child-parent”关系。在这种情况下，上层标记可以包含多个下层标记，但下层标记或属于某个上层标记或没有上层标记(受最大标记阈值限制)。

最后，根据全色图像计算上下层标记的相关程度。如果子、父标记间的“child-parent”关系成立，则分别取得它们在全色图像对应区域的灰度直方图，并用式(6)计算其相关性：

$$R=\frac{n\sum(h_p h_c)-\sum h_p\sum h_c}{\sqrt{n\sum h_p^2-(\sum h_p)^2}\sqrt{n\sum h_c^2-(\sum h_c)^2}} \tag{6}$$

式中，h_p与h_c分别为父、子标记的直方图；n为直方图的灰度级。设置相关性阈值T_R，并分析满足“child-parent”关系的标记之间的灰度相关性。若高于T_R，则认为父标记能够代表其子标记，否则保留子标记。将各层次所有满足上述条件的标记统一起来，得到最终标记图。

2.3 多约束区域合并

在多层次标记图的基础上用强制最小技术修改梯度图，利用基于FIFO队列的分水岭方法(Vincent和Soille，1991)对图像进行分割。在得到的分割结果中，相邻同质区域可能属于同类地物，即它们的光谱与纹理等特征存在一定程度的相似。因此，需要通过设置一些规则对它们进行合并处理。高分辨率遥感图像上地物特征明显、关系复杂，在进行区域合并时应考虑光谱与纹理特征。本文的纹理特征采用Gabor小波进行提取，其表达形式为(Petkov和Kruizinga，1997)：

$$h_{\theta,\lambda,\varphi,\sigma,\gamma}(x,y)=\exp\left(-\frac{1}{2}\frac{(x')^2+(\gamma\cdot y')^2}{\sigma^2}\right)\cdot\cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \tag{7}$$

式中，$x'=x\cdot\cos\theta+y\cdot\sin\theta$，$y'=-x\cdot\sin\theta+y\cdot\cos\theta$；$x$、$y$为空间域中的像元位置坐标；$\lambda$、$\theta$分别为Gabor函数的尺度与方向；$\varphi$为相角；$\sigma$为高斯半径；$\gamma$为高斯包络线的纵横比。

设A、B为两个相邻的区域，合并代价函数的形式如下：

$$\xi=\frac{L_A\cdot L_B}{L_{AB}^2}\frac{S_A\cdot S_B}{S_A+S_B}\left[\alpha\Delta\left\{F_A^S,F_B^S\right\}+(1-\alpha)\Delta\left\{F_A^T,F_B^T\right\}\right] \tag{8}$$

式中，L_A、L_B与L_{AB}分别为区域A、B的周长及其公共边界长度；S_A与S_B分别为区域A、B的面积；α为光谱特征距离的权重；F^S与F^T分别为区域A或区域B的光谱与纹理特征均值；$\Delta\{\cdot,\cdot\}$为区域A与区域B特征均值之间的欧氏距离。由式(8)可知，当区域A、B之间的公共边界在彼此的周长中所占比例较大、各自面积较小且光谱与纹理特征较一致时，它们合并的代价就会比较小且更具有合理性。相邻区域的合并将按如下步骤进行。

步骤1 对所有区域进行序号标记，并计算各区域由光谱与Gabor纹理所构成特征向量的均值。

步骤2 建立区域邻接图并取得相邻区域对，提取所有区域的面积、周长与相邻区域的公共边界长度。

步骤3 计算所有相邻区域间的合并代价，排序后找到最小值对应的两个区域。

步骤4 假设区域A的序号在B之后，按如下过程进行区域合并和数据更新：

- 将区域A与区域B的周长相加，减去两倍公共边界长度，作为区域B的新周长；
- 将区域A、B的面积与公共边界的长度3者相加后作为区域B的新面积；
- 将区域A的相邻关系转向区域B；
- 将所有与区域A相关的公共边界转至B；
- 将区域A合并到区域B，使两者序号同为B的序号。

步骤5 循环进行步骤3与步骤4的计算，直到满足以下两个停止阈值：最大区域面积与最小特征欧氏距离。其中，前者用于降低区域合并的代价，后者突出光谱与纹理特征的重要性。

2.4 分割结果投影

由于小波分解是以2的倍数向下采样，合并后的分割结果需要通过小波逆变换逐级投影到原始图像。Jung(2003)依据边界像元与其邻域像元的最小灰度差，进行投影过程中的边界处理。但该方法计算量较大，且多次投影后的边界容易出现大幅位移。为了在投影过程中尽量保持区域边界的位置，下面提出滑动模板边界处理方法(只需一个波段图像)。将分割结果对应尺度下的小波分解系数修改为：

$$NW_J^l(n,m)=\begin{cases}0 & \text{如果 } (n,m) \text{ 属于区域边界}\\ W_J^l(n,m) & \text{其他}\end{cases} \tag{9}$$

式中，l为近似与细节系数。式(9)表示近似与细节系数边界位置的像元为0，其他像元保留原值。

图2所示为边界处理的过程。首先，由小波逆变换将边界投影到上级尺度的近似系数，得到宽度为2×2(因采用Haar小波)的边界(图2(a))。然后，对各区域进行标记排序(图2(b))，并用3×3的模板(图2b中的阴影)沿着边界(图2(b)中加粗的0)进行滑动，取模板内的像元均值放入存储器的对应位置。最后，将图2(b)与存储器相加得到新的区域(图2(c))。需要说明的是，这种边界处理方法不是用计算区域均值后的近似系数而是用标记图，是因为前者可能会出现相邻区域的均值相等的情况。从图2(c)可见，由于边界线像元都被赋予了相邻区域的标记值，相邻区域的差异性比图2(a)更加直接和简单。

进而，利用Roberts算子产生新边界。因为，该算子通过2×2的卷积模板计算对角线方向相邻像元之差来检测边缘，在水平与垂直方向的梯度检测的效

91	86	82	84	90	86	0	0	101	103
88	87	87	88	86	87	0	0	111	109
86	87	0	0	0	0	0	0	108	105
82	88	0	0	0	0	0	0	114	110
93	87	0	0	108	112	108	102	112	109
88	84	0	0	114	110	107	105	112	110
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
47	44	0	0	212	218	208	213	215	220
45	48	0	0	215	217	216	221	209	211

(a)

1	1	1	1	1	1	0	0	2	2
1	1	1	1	1	1	0	0	2	2
1	1	**0**	0	0	0	0	0	2	2
1	1	0	0	0	0	0	0	2	2
1	1	0	0	2	2	2	2	2	2
1	1	0	0	2	2	2	2	2	2
0	0	0	0	**0**	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
3	3	0	0	4	4	4	4	4	4
3	3	0	0	4	4	4	4	4	4

(b)

1	1	1	1	1	1	1	2	2	2
1	1	1	1	1	1	1	2	2	2
1	1	1	1	1	1	1	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
3	3	3	4	4	4	4	4	4	4
3	3	3	4	4	4	4	4	4	4
3	3	3	4	4	4	4	4	4	4

(c)

图2 边界处理

果较好、定位精度较高，且在处理边界时受噪声的影响很小。新边界将出现在如图2b所示的零值轨道中，这样其走向便与初始边界一致。另外，3个以上区域的公共边界线可能会出现单像元的断裂。本文将通过图3所示的过程进行连接。图3a为3×3的端点像元检测模板。其中，两个实线双圆为已有像元，它们表示0°方向的断线，虚线圆圈表示断线前方有边界像元，实线的单圆为填入的边界像元，空白方格表示该位置没有边界像元。通过调节方向，使该模板能够处理90°、180°、270°的断线。用Roberts算子提取图2c的边界线，得到如图3b所示的边界，而其中的圆圈为经上述方式填入的像元。

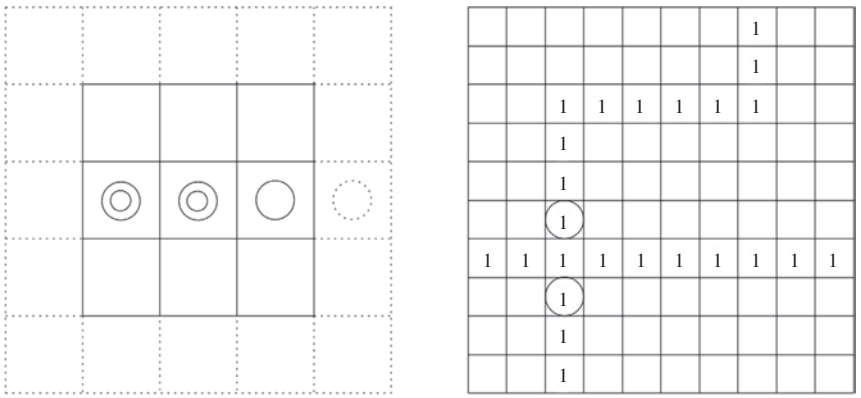

图3 断边连接

(a) 3×3的端点像元检测模板；(b) 用Roberts算子提取的边界线

3 实验及结果分析

实验数据采用经几何校正与正射校正的Quickbird影像，其获取时间为2004年11月21日，地理位置为南京市江宁区，大小为1024×1024像元。它包括蓝、绿、红、近红外等多光谱波段与全色数据，其空间分辨率分别为2.44 m与0.61 m。将多光谱数据分别与全色数据进行PANSHARP融合(Zhang，2002),使它们的空间分辨率变为0.61 m。图4为红外、红、绿三个波段组合而成的假彩色图像，其中圆圈与数字标示出一些典型地物，包括道路(C1)、体育场(C2)、水面(C3、C4)、房屋(C5、C9)、农田(C6、C7、C10、C11、C12)、树木(C8)。由于庄稼已经收割，分布最广的农田有着多样的光谱与纹理特征。房屋较为集中，但形态各异、纹理复杂。道路的上半段光谱突出、纹理光滑，但下半段与路边房屋的光谱接近。水面的纹理单一，却因所含物质不同而表现出不同光谱特征。图像下部零星分布着树木、房屋与水面。实验平台为Windows XP操作系统，CPU主频为2.4 GHz，内存为2 G，用Matlab编程实现。

3.1 基于本文方法的分割

(1)选择最佳小波分解尺度

针对所选12种地物各取5个样本用于判断最佳小波分解尺度。由于最大分解尺度J等于4(由式(2)可得)，按尺度从小到大分别用4×4、8×8、16×16、32×32的模板选择样本。计算4个波段近似系数在4个尺度下的相位一致梯度。参数设置如下：尺度n=5，方向o=6$\left(\text{即}0,\frac{\pi}{6},\frac{\pi}{3},\frac{\pi}{2},\frac{2\pi}{3},\frac{5\pi}{6}\right)$，最小尺度滤波器带宽为3个像元，相邻尺度间带宽比为2.1，噪声估计T=0.2，截止频率c=0.05，增益因子g=15，ε=0.0001。图5为4个尺度下局部梯度方差。从中可以看出，持续减少的有C7、C8，持续增加的有C1、C10、C11，先增加后减少的有C2、C3、C5，先减少后增加的有C4、C6、C9、C12。图6为4个尺度下的平均局部梯度方差。“V”形曲线指出在尺度2取得最小值。此尺度即为最佳小波分解尺度。

(2)提取多层次标记

通过式(3)融合尺度2下4个波段的梯度图。如图7

所示，各类地物均具有清晰的轮廓。该梯度图像的均值与方差分别为0.2474与0.1721，设置调制值为–0.5、–0.4、–0.3、–0.2、–0.1、0.0、0.1、0.2、0.3、0.4、0.5，进而根据式(5)计算得到高度阈值序列为0.1614、0.1786、0.1958、0.2130、0.2302、0.2474、0.2646、0.2818、0.2990、0.3162、0.3334。从图4可以发现，面积最大且纹理光滑的地物为道路，但其下半段与路边房屋的光谱接近，故取上半段道路面积为最大标记阈值，该值为3521。由于只设置最大标记阈值，地物细节都能够出现在初始分割结果中，特别是纹理比较复杂的地物及其周边。建立“child-parent”关系，设置相关性阈值T_R为0.85，组合所有满足条件的标记得到如图8所示的标记图。从图8可以看到，多层次标记能够兼顾大小不同的地物。比如，道路、水面、农田等纹理单一的地物与房屋、树木、空地等纹理复杂的地物都在标记图上反映出来。利用分水岭变换得到尺度2下的初始分割结果，区域数为577。如图9所示，分水岭脊线准确地包络了图像上的各类地物，特别是农田、道路与水面等纹理相对均匀的地类，而复杂纹理的地物则由更小的区域体现。

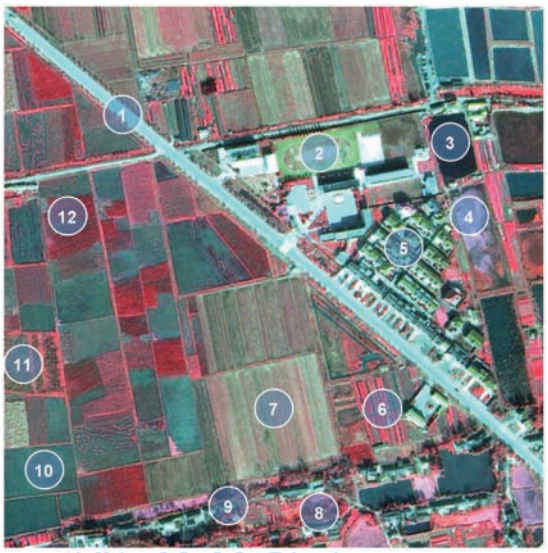

图4 假彩色图像

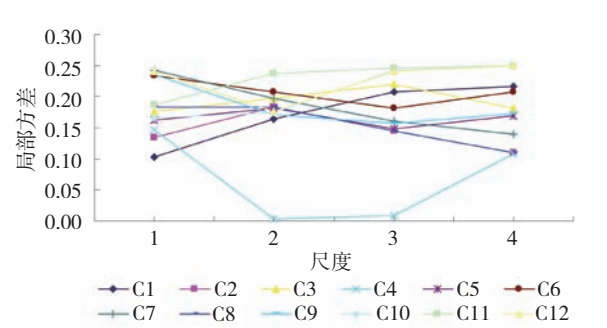

图5 局部梯度方差

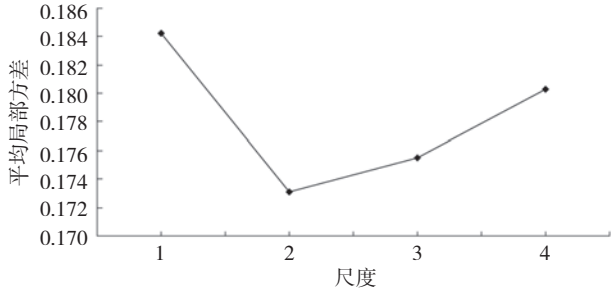

图6 平均梯度方差

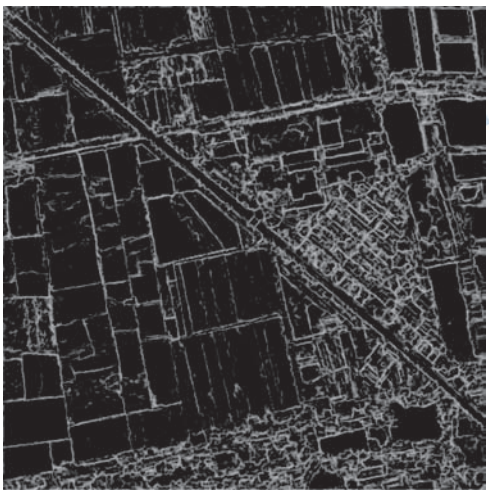

图7 融合后的梯度图

图8 标记图

图9 初始分割结果

(3)区域合并与结果投影

计算得到区域的边界、面积、归一化的光谱与纹理特征，并按照2.3节描述的步骤进行区域合并。在提取纹理特征时，Gabor小波的参数设置为：γ=1，b=1.3，φ=0，$\theta=\left\{0,\frac{\pi}{6},\frac{\pi}{3},\frac{\pi}{2},\frac{2\pi}{3},\frac{5\pi}{6}\right\}$，$\lambda$取前4个尺度。考虑篇幅的限制，上述参数设置的依据请参见文献(陈杰 等，2010)。利用式(8)计算相邻区域的合并代价。设置最大区域面积阈值T_M为380，最小特征欧氏距离阈值为0.35。设置阈值时注意让前者偏小后者偏大，目的是促使相邻的小面积区域能够优先进入合并队列。由于纹理特征的地物区分能力较光谱特征更强，设光谱特征距离权重为0.4，纹理特征距离权重为0.6。图10为合并结果(区域数为310)，从中可以看出，对应地物细节的小区域基本上都得到了正确合并。尽管最大区域面积阈值阻碍了部分属于相同地类的相邻区域的合并(比如，位于图下端的农田与道路上侧成片的房屋)，但所得结果依然能够满足图像分割的要求。进而，依照2.4节介绍的方法，将图10的合并结果投影到原始图像。由于最佳分解尺度为2，只作一次投影。用Roberts算子产生新边界，并设置阈值为0.000001。所得结果如图11所示，其边界线位置与图10基本一致。

(4)运算速度分析

本文方法是在小波变换的多尺度分解基础上实现遥感图像的分割，包括多尺度图像生成、图像分割、区域合并和结果映射等4个关键步骤。它属于监督分割方法，整个过程需要少量人机互动操作。比如，在选择小波分解尺度时需要有针对性地选择各地物样本，在多层次标记提取设置调制值和最大标记阈值时需考虑地物的结构与分布特点。其次，遥感图像的大小、波段数与图像内容的复杂程度，会直接影响分割过程的计算机运算量。其中，图像的大小和波段数影响相位一致梯度计算以及梯度融合的计算时间，图像内容则影响标记的大小与数目、相邻区域对的数量等决定计算量的几个重要因素。另外，区域合并所需时间相对最长，但通过设置最大区域面积阈值可以一定程度地减少其运算时间。

图10 合并结果

图11 本文方法的最终结果

3.2 基于其他方法的分割

Jung(2007)将离散小波与分水岭变换相结合提出非监督多尺度分割(UMS)方法。通过多幅自然彩色图

像的分割实验表明，该方法的分割效果较Statistical Region Merging(SRM)、JSEG与Mean-shift等方法更好。下面用UMS方法进行同样遥感图像的分割。经多次测试后，在设置小波分解尺度与合并阈值分别为4、0.05时可取得最佳的分割结果，其区域数为285。如图12所示，该方法对于纹理光滑与分布集中的地物分割效果较好，如体育场、水面、房屋、农田与道路。但由于合并区域时仅考虑光谱因素，色彩相似的邻接区域无法区分，出现了大面积的农田与破碎的道路。另外，因为没有最大面积阈值的限制，合并结果中的1个区域可能包含多个地物。该方法的最大不足是，在分割结果投影时采用了最近距离判别方法处理边界线，因此增加了区域边界在多次投影后的方向随机性。比如，具有规则边界的水面、道路与农田都未能在最后的结果中保持初始分割的边界线。但另一方面，该方法得到的地物形状完整性优于本文方法(如成片房屋被整体分割出来)，而且非监督的分割过程使其计算时间较少。

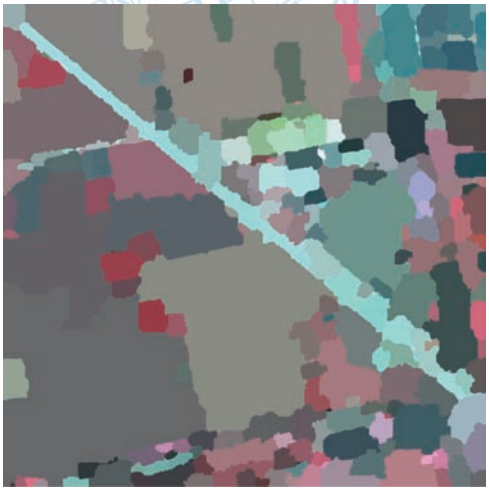

图12 非监督多尺度分割结果

3.3 分割效果定量对比分析

引入对象一致性误差(OCE)分割评判标准(Polak等，2009)定量评价上述两种方法的分割效果。该标准能够检测到分割区域的形状、位置、大小及存在性，对过分割与欠分割一样敏感，且具备对称性与尺度不变性。OCE满足条件$0 \leqslant \mathrm{OCE}(I_g, I_s) \leqslant 1$，其中$I_g$为参考分割，$I_s$为实际分割。如果OCE=0，则表示实际分割与参考分割相同。由于本文所用遥感图像与其地面现状有较大不同，而且0.61 m 的空间分辨率使图像能够体现大部分的地物细节，参考分割图将由目视解译获得。首先判读原图获得参考图，然后将参考图栅格化，最后得到按序编号的区域图像(图13)。高分辨率遥感图像上的地物细节丰富、地物关系复杂、纹理结构多样，同样的分割方法及参数对不同地物会有不同的效果，因此有必要进行区别评价。OCE是以分割区域为计算对象的，因此可以依据地物的集中程度及光谱与纹理特征的相似度将区域组织成若干片区。如图13所示：A、C、E区为农田集中区且轮廓清晰；B区为形状突出的道路；D区为水面；F区分布有房屋、体育场、空地与树木等地物；G区以水面为主，房屋与农田穿插其间；H区为集中分布的房屋；I片区以农田为主并有零散建筑；J区地物类型多样且关系复杂。

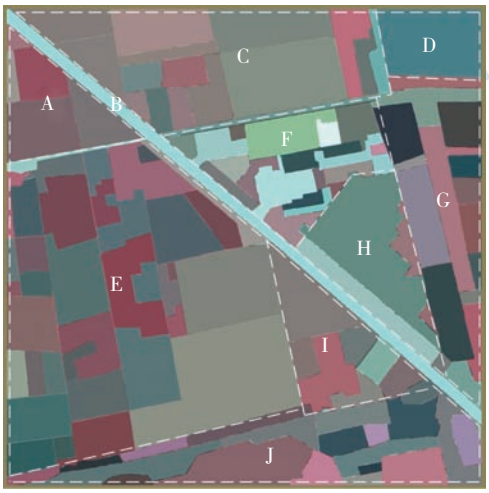

图13 分割评价的参考图

表1按照参考图的片区从区域数与OCE值两个方面对比两种分割方法的差别。从区域数来看，由于存在分割区域跨越两个以上片区的情况，本文方法与UMS方法分别比它们的分割区域总数多103与105，该数字表明两种方法的区域重复规模相当。具体上，本文方法在A、B、D、F、G、I与J片区的区域数少于UMS，C、E与H片区多于UMS，表示UMS的区域离散度高于本文方法。从OCE值来看，本文方法在A、B、E、F、H、I与J片区小于UMS，C、D与G片区大于UMS，表示本文方法的分割区域更加符合参考图像。对于道路片区B，UMS不如本文方法是因为前者分割区域更加离散。对于水面片区D，UMS优于

本文方法是因为前者的区域更加集中，后者存在跨越两个以上片区的区域。而结合F、G、H、I与J片区的区域数与OCE值可以发现，虽然区域数量少可以得到更小的OCE值，但区域数量并非与OCE值成正对应关系。总体来看，本文方法的OCE值为4.8445比UMS的5.1013少0.2568，说明前者分割效果更佳。从两种方法的优劣对比中可知，区域离散度小、分布集中以及边界准确是较高分割精度的外在表现。而有限的分割规则一般难以兼顾遥感图像上不同光谱、纹理、形状、大小与类型的地物。

表1 两种分割方法的对比

		A	B	C	D	E	F	G	H	I	J
本文方法	区域数	14	26	55	8	82	46	48	46	27	61
	OCE值	0.4810	0.4590	0.5231	0.5160	0.4090	0.4605	0.5217	0.4877	0.4733	0.5132
UMS	区域数	20	36	31	21	41	62	52	28	29	70
	OCE值	0.5009	0.5289	0.5018	0.5014	0.4971	0.5100	0.5060	0.5087	0.4987	0.5478

4 结论与展望

本文将分水岭变换与离散小波变换相结合以发挥多尺度分割的优势，并从不同角度提出改进方法。利用相位一致梯度准确捕捉地物边界是分水岭变换的重要前提。融合不同波段图像的灰度梯度以充分利用多波段信息，通过分析局部梯度方差选择小波分解尺度更有目的性地产生多尺度图像。以最小局部方差作为最佳尺度选择的依据，使在该尺度下获取的光谱与纹理特征更具可分性，从而为区域的合并提供支持。由灰度相关程度引导上下层标记关系的建立具有合理性，且多层标记可以兼顾不同地物，有利于提高分水岭变换的分割效果。以Gabor纹理特征、区域面积和区域空间相邻关系为约束因素，促使区域合并更加准确。小波逆变换为结果投影提供了可靠保障，且文中提出边界处理方法使分割结果能够很好地反映到原始图像。

虽然实验结果显示本文方法可取得较好分割效果，但也存在一些不足。比如，提取梯度与纹理特征时涉及较多参数，标记获取与区域合并的阈值需要多次调试。进一步的工作将是分析不同参数对结果造成的影响，以及引入其他技术方法来对本文方法作进一步的完善。

REFERENCES

Angulo J, Velasco-Forero S and Chanussot J. 2009. Multiscale stochastic watershed for unsupervised hyperspectral image segmentation. IEEE International Geoscience and Remote Sensing Symposium: 93–96

Blaschke T, Lang S, Lorup E, Strobl J and Zeil P. 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. *Environmental Information for Planning*, **2**: 555–570

Chen J, Deng M, Xiao P F and Yang M H. 2010. Texture segmentation based on Gabor filters and ant colony optimization algorithm. *Geomatics and Information Science of Wuhan University*, **35**(11): 1271–1274, 1334

Haar A. 1911. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, **71**(1): 38–53 DOI: 10.1007/BF01456927

Hill P R, Canagarajah C N and Bull D R. 2003. Image segmentation using a texture gradient based watershed transform. *IEEE Transactions on Image Processing*, **12**(12): 1618–1633 DOI: 10.1109/TIP.2003.819311

Hu X Y, Tao C V and Prenzel B. 2005. Automatic segmentation of high-resolution satellite imagery by integrating texture, intensity, and color features. *Photogrammetric Engineering and Remote Sensing*, **71**: 1399–1406

Jung C R. 2003. Multiscale image segmentation using wavelets and watersheds. Proceedings of the XVI Brazilian Symposium on Computer Graphics and Image Processing: 278–286

Jung C R. 2007. Unsupervised multiscale segmentation of color images. *Pattern Recognition Letters*, **28**(4): 523–533 DOI: 10.1016/j.patrec.2006.10.001

Jung C R and Scharcanski J. 2000. Robust watershed segmentation using the wavelet transform. Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing: 131–137

Kim J B and Kim H J. 2003. Multiresolution-based watersheds for efficient image segmentation. *Pattern Recognition Letters*, **24**(1-3): 473–488 DOI: 10.1016/S0167-8655(02)00270-2

Kothainachiar S, Wahita Banu R S D and Saravanan A. 2006. Unsupervised morphological segmentation for textured and non-textured images. *The International Journal on Graphics, Vision and Image Processing*, **6**(2): 33–39

Mallat S G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pat-*

tern Analysis and Machine Intelligence, **11**(7): 674–693 DOI: 10.1109/34.192463

Mao F, Liu Z, Zhou W S and Li Q. 2009. Extracting of urban features from high resolution remote sensing data based on multiscale segmentation. Proceedings of the 2009 Urban Remote Sensing Joint Event

Ming D P, Wang Q and Yang J Y. 2008. Spatial scale of remote sensing image and selection of optimal spatial resolution. *Journal of Remote Sensing*, **12**(4): 529–537

Nguyen H T, Worring M and van den Boomgard R. 2003. Watersnakes: energy-driven watershed segmentation. *IEEE transactions on Pattern Analysis and machine Intelligence*, **25**(3): 330–342 DOI: 10.1109/TPAMI.2003.1182096

Petkov N and Kruizinga P. 1997. Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. Biological Cybernetics: 83–96

Polak M, Zhang H and Pi M H. 2009. An evaluation metric for image segmentation of multiple objects. *Image and Vision Computing*, **27**(8): 1223–1227. DOI: 10.1016/j.imavis.2008.09.008

Soille P. 2003. Morphological Image Analysis: Principles and Applications. 2nd ed. New York: Springer Verlag

Spirkovska L. 1993. A summary of image segmentation techniques. NASA Technical Memorandum: 1–11

Sun Y and He G J. 2008. Segmentation of high-resolution remote sensing image based on marker-based watershed algorithm. The 5th International Conference on Fuzzy Systems and Knowledge Discovery: 271–276

Tarabalka Y, Chanussot J, Benediktsson J A, Angulo J and Fauvel M. 2010. Segmentation and classification of hyperspectral data using watershed. *Pattern Recognition*, **43**(7): 2367–2379

Vincent L and Soille P. 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**(6): 583–598 DOI: 10.1109/34.87344

Wang J Z, Li J, Gray R M and Wiederhold G. 2001. Unsupervised multiresolution segmentation for images with low depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(1): 85–90 DOI: 10.1109/34.899949

Xiao P F, Feng X Z and Li H. 2009. Multispectral remotely sensed imagery segmentation based on first fundermental form. Proceedings of the 2009 Urban Remote Sensing Joint Event

Xiao P F, Feng X Z, Zhao S H, Deng M and She J F. 2007. Feature detection from high-resolution remotely sensed imagery based on phase congruency. *Journal of Remote Sensing*, **11**(3): 303–310

Zhang Y. 2002. Problems in the fusion of commercial high-resolution satellite images as well as Landsat7 images and initial solutions. *International Archives of Photogrammetry and Remote Sensing*, **34**: 236–242

附中文参考文献

陈杰，邓敏，肖鹏峰，杨敏华. 2010. 利用Gabor滤波器与蚁群算法进行纹理分割. 武汉大学学报(信息科学版)，**35**(11): 1271–1274，1334

明冬萍，王群，杨建宇. 2008. 遥感影像空间尺度特性与最佳空间分辨率选择. 遥感学报，**12**(4): 529–537

肖鹏峰，冯学智，赵书河，邓敏，佘江峰. 2007. 一种基于相位一致的高分辨率遥感图像特征检测方法. 遥感学报，**11**(3): 303–310