A Hybrid Multi-Scale Segmentation Approach for Remotely Sensed Imagery

Qiu-Xiao Chen[1,2], Jian-Cheng Luo[1], Cheng-hu Zhou[1], Tao Pei[1]

(1. The State Key Laboratory of Resources & Environmental Information System, Institute of Geographical Sciences and Natural Resources, Chinese Academy of Sciences, Building 917, Datun Road, Anwai, Beijing, 100101, P.R.China

2. Department of Regional and Urban Planning, Zhejiang University, Hangzhou, 310028)

Tel: (86)-10-64889633 Fax: (86)-10-64889630 Email: chenqx@lreis.ac.cn

***Abstract*-The general image segmentation approach used in other domains may not be applicable to the remote sensing field, which is due to the following factors: remotely sensed data is multi-spectral, always very large in size, and in multi-scale as well. How to quickly and efficiently segment remotely sensed imagery is still a big issue to be solved. Based on human vision mechanism, a new hybrid multi-scale segmentation approach is presented, which is implemented at three coarse-to-fine scale levels. First, remotely sensed imagery is segmented at a coarse scale, and image regions (segments) are produced. Then, the corresponding regions in the original image are segmented by another segmentation approach one by one at the fine scale. From the experiment results, we found the approach is rather promising. However, there still exists some problems to be settled, and further researches should be conducted in the future.**

***Keywords*-scale; top-down strategy; segmentation; remotely sensed imagery**

Ⅰ. BACKGROUND

In most cases, information needed for image analysis and understanding is not represented in pixels but in meaningful image objects and their mutual relations. Therefore, to partition images into sets of useful image objects is a fundamental procedure for successful image analysis or automatic image interpretation (Gorte 1998, Baatz and Schäpe 2000, Blaschke et al. 2001). In this sense, image segmentation is critical for subsequent image analysis, and even image understanding further.

From the perspective of image engineering, conventional per-pixel approach for remotely sensed imagery is just at its low level -- image processing level. To extend and expand the application field of remote sensing, the shift in this field from per-pixel image processing to object-based image analysis taking segmentation as its initial procedure is recommended. In the remote sensing field, one strong motivation for segmentation of remotely sensed imagery stems from the fact that most image data exhibit characteristic texture which is always neglected in common classifications.

However, segmentation is by no means a new research field. On the contrary, it has been a widely concerned research topic by scientists and specialists in several domains like computer vision, artificial intelligence, signal processing, etc., and also many application fields for years. However, to segment images successfully still remains a difficult task to be settled. As to the remote sensing field, corresponding researches are rather insufficient compared with that of other domains, which will prevent many potential application fields of remote sensing considerably.

So far, there are over 1000 kinds of segmentation approaches developed (Zhang 2001). However, these segmentation approaches may not applicable in the remote sensing field due to the following facts: (1) Remotely sensed imagery are multi-spectral and in multi-scale, so both the complexity and redundancy are increased obviously; (2) In contrast to other applications, various objects of heterogeneous properties with respect to size, form, spectral behaviour, etc. have to be considered; (3) In contrast to other applications a model-based interpretation is much more difficult due to the heterogeneity of the inherent object classes; (4) Sub-optimal solutions will probably not be considered for remote sensing applications because there is no need for real-time or dynamic evaluations (Blaschke 2000). Hence, segmentation algorithms have been introduced relatively late for the analysis of remotely sensed data (e.g., Ryherd and Woodcock 1996).

Generally, segmentation of remote-sensing images can only provide initial information for the following processing to reveal spatial pattern in a certain scale. Segments produced by the fine segmentation may represent nothing meaningful at the coarse scale, that is to say, segments in an image will never represent meaningful objects at all scales. Therefore, to fully grasp the spatial pattern and its dynamic, multiscale segmentation of remotely sensed imagery is needed. On the other hand, there exists a dilemma that fine segmentation approach can reveal detail features like fine texture, however it is always unable to provide macro structural feature. Vice versa, coarse segmentation can extract macro feature but fail to provide detail feature. To fully exploiting the features and information hidden in the images, multiscale segmentation is in need.

In this study, a hybrid multiscale segmentation taking top-down strategy is proposed. First, using wavelet transform, remotely sensed images are decomposed. Second, by clustering the feature space formed with the subimages, a coarse segmentation result is obtained. Then, by some templates, another feature space of the original image is produced. Each region in original image, corresponding to the coarse scale segmentation results (coarse segments), is segmented into fine segments under the newly established feature space. At last, boundary pixels of coarse scale are relabled to refine and correct the blocked boundary.

II. Methods

Human vision mechanism lays a good foundation for multiscale segmentation, which always tends to generalize images into homogeneous areas first, and then characterizes these areas more carefully later (Gorte 1998). From the legible extent to the blurring extent, human being's perception of objects in the scene will be shifted from detail features to macro features which is a process when human vision system is relaxing gradually (Luo, 1999); while from the blurring extent to the legible extent, human vision's focus is changed from macro features to detail feature, which is a process when human vision system is continuously focusing on the scene. According to this, two strategies for multiscale segmentation can be taken: (1) Bottom-up strategy – from fine to coarse scale, imitating human vision's relaxing process; (2) Top-down strategy – from coarse scale to fine scale, simulating human vision's focusing process.

Multiscale segmentation always takes bottom-up strategy (e. g., Luo 1999, Tilton 2000, Lakshmanan et al. 2000, Vanhamel and Sahli 2001, etc.). However, segmentation using top-down strategy is seldom developed. This study attempts to contribute a little in this regard.

A. Segmentation at the Coarse Scale

The wavelet approach provides us a framework for multiscale analysis. In this study, we use wavelet decomposition to downscale the image. For a 2-D image, wavelet transforms can produce a family of hierarchically organised sub-images. For example, after each transform, a $2^N \times 2^N$ input image is decomposed into four $2^{N-1} \times 2^{N-1}$ sub-images: approximate, horizontal, vertical, and diagonal. Each has distinctive spatial and frequency characteristics. The last three images are also known as "detail images". Further information on wavelet analysis can be found in [10]. In this study, a Daubechies's wavelet with Rank 2 and Genus 2 is used. After the decomposition, segmentation results in a coarse scale can be obtained by clustering the sub-images. Meanwhile, this segmentation result will be recorded into a new channel Z and written back to the original image.

Due to the multispectral feature of remotely sensed data, information lost will be inevitable if only one channel is utilized. Therefore, before wavelet transform, the principle component analysis is conducted.

B. Segmentation at the fine Scale

In a fine scale, both grey information and texture information should be considered due to the fact that texture feature may not so obvious and human vision is sensitive to grey information. On the other hand, general texture segmentation is always very complicated, and its implementation is generally time consuming. Bearing these in mind and also in consideration that remotely sensed imagery is very large in size, we seek to develop a relatively simple approach to represent the grey and texture features in order to improve the implementation effectiveness. In this study, we utilize Sobel 3x3 filters to extract horizontal and vertical gradient features from the image, Kirsch 3x3 filters to extract diagonal gradient and antidiagonal gradient features. At he same time, a 3×3 mean filter and variance filter is utilized to extract local mean and variance. Using these filters, we can construct a 6×m dimension feature space from a remotely sensed image with m bands. To reduce the computing cost, only the main component of the image is used. Thus, six features are extracted in the fine scale.

Recent works have shown that region based approaches outperform other methods in terms of segmentation accuracy and satisfactory results have been presented by many researchers as in [12], [13], [14], [15] and many other literatures. As a typical region based approach, region growing algorithms cluster pixels with seed points which is growing into regions until a certain threshold is reached. This threshold is normally a homogeneity criterion or a combination of size and homogeneity. A region grows until no more pixels can be merged into, and new seeds are placed and the process is repeated. This continues until the whole image is segmented.

In this study, a region growing approach is proposed based on the feature space or feature vector mentioned above. During the growing process, neighbour pixels with the minimum feature distance to the feature center of the region are merged with higher priority so that the final partition contains the smallest possible population of separate and homogeneous areas. Distinctive from other regional growing methods, the growing is confined to the range of coarse segments in each time. To improve the implementation speed, a channel containing distance between nieghbour pixels is calculated in advance.

In detail, a fine segmentation approach is implemented as follows,

(1) Calculate between each pixel and its four-neighbour pixels the feature distance

$$D = \sum_{h=1}^{r} \left(\frac{P_{i,h} - P_{j,h}}{P_{i,h} + P_{j,h}} \right)^2 , \qquad (1)$$

Where, $P_{i,h}$ and $P_{j,h}$ are the *h*th feature value of the *i*th and *j*th pixel , h denotes feature, and r is the feature number. These distances are then written in a matrix named Y. If the image is in m× n size, then Y is in (2m-1) × n size. For a random pixel like P_{ij}, the distance between P_{ij} and its four-neighbour pixels, namely $P_{i-1,j}$, $P_{i,j-1}$, $P_{i,j+1}$, $P_{i+1,j}$, are recorded as $Y_{2(i-1),j}$, $Y_{2(i-1)+1,j-1}$, $Y_{2(i-1)+1,j}$, $Y_{2i,j}$ respectively.

(2) Set R = P_{ij}, here P_{ij} being a seed pixel, and R being the initial segment containing only one pixel so far. We also define a region R_n, which is comprised with all the pixels in R and their four-neighbour pixels, and set $R_n' = R_n - R$, put the distances between P_{ij} and its four-neighbour pixel into a set named Y'. Obviously, Y'= $\{Y_{2(i-1),j}, Y_{2(i-1)+1,j-1}, Y_{2(i-1)+1,j}, Y_{2i,j}\}$.

(3) Verify whether each pixel in R_n' belongs to the same coarse scale segment or not, which can be judged by its digital number in channel Z. If not, then this neighbour pixel should be removed from R_n and R_n'. Thus, for every segment at the coarse scale, its counterpart region in the origional image can be segmented by a fine segmentation approach one by one.

(4) Arrange all the elements in Y' in an ascending order, find out Y1's counterpart pixel (P_{min}) in image by utilizing

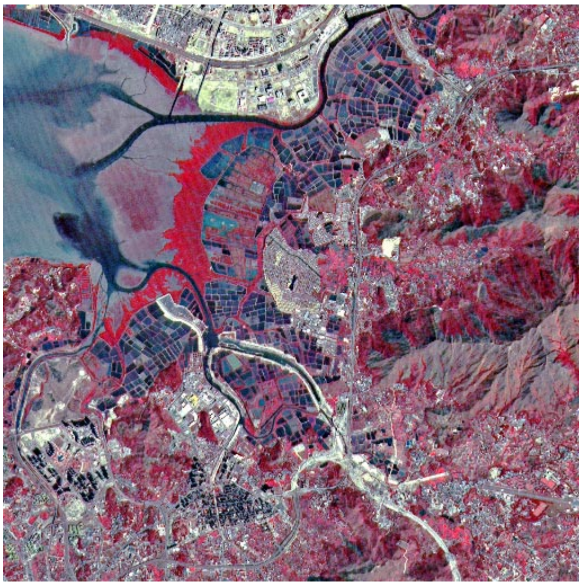

Figure 1. A false colour composite Sample image

footnote as its clues, merge P_{min} into R, calculate the heterogeneity (H) of the region R after merge. If H < T (threshold), replace Y_1 with a large enough value, remove Y_1' from Y', then go to step (5); otherwise, remove P_{min} from R, stop the growing of R. H can be calculated as follows,

$$H = \sum_{h \subset R} \left(\frac{P_{i,h} - \overline{F_i}}{\overline{F_i}} \right)^2 , \tag{2}$$

Where $\overline{F_i}$ is the ith feature value of segment's feature center, $P_{i,h}$ is the ith feature value of the hth in the segment, R stands for the segment.

(5) Retrieve the feature distances between P_{min} and its four-neighbour pixel from Y, and add these elements into Y'.

(6) Select a pixel which is not in the already formed fine segment but still in the same region at the coarse scale, repeat step 2-5.

(7) Repeat step 6 until no pixel is left in the same region at the coarse scale.

(8) Repeat step 2-7 until all the segments at coarse scale are segmented by the fine segmentation approach.

C. Boundary pixels Adjusting

Due to the wavelet decomposition, blocked boundary is inevitable. So, the boundary of segments at the coarse scale level may not fit the reality well. Therefore, to adjust the boundaries is also very important to improve the segmentation result. As to every boundary pixel of coarse segments, we calculate its distance to the feature center of the fine segment the pixel belongs to, and also the distances to its neighbour fine segment. By comparing the distance, we can finally label the "boundary pixels" and fine-tune the boundary of coarse segments.

Assume that there exist two fine segments named B1 and B2, which belong to the coarse segment A1 and A2 respectively. As to a pixel -- C from B1, and with B2 as its neighbour segment, if the feature distance to B1 is smaller than that to B2, the label of C should not be adjusted. Otherwise, C should be merged into B2. Here, a normalized Euclidean distance is utilized.

III. EXPERIMENTAL DATA

To evaluate the performance of the proposed segmentation approach, a *SPOT-HRV* data with three spectral bands (*CH1*: Green Band, 0.50-0.59μm; *CH2*: Red Band 0.61-0.68μm; *CH3*: Near Infrared Band, 0.79-0.89μm) is used. This remotely sensed image was acquired on *Feb* 3, 1999 covering the Hong Kong island. The size of the sub-image cut down from the whole image is 1024 rows by 1024 columns with a spatial resolution of 20m by 20m (Figure 1).

IV. RESULTS and DISCUSSION

In the coarse segmentation period, principle component analysis is conducted based on three bands of the raw image, the first component is extracted. Then the first component image undergoes two levels of wavelet decomposition, producing two sets of sub-images with the following sizes: 512 × 512 (resolution 40 m), 256 × 256 (80m). As the level increases, the spatial resolution for ground objects becomes coarser. According to knowledge acquired from supporting materials and surveys, and with visual interpretation of the corresponding remote-sensing data, seven main types of land covers are identified, they are residential areas, cropland, hilly woodland, beach, mangrove areas and grass land. So in the following FCM (Fuzzy Center Means) clustering rountine, seven classes are used.

From the figure 2-3, we can see that segmentation of residential areas (mainly located in the upper-middle corner of the image), crop lands (always with a regular shape), mangrove areas, beach (in the upper-left corner) and water after boundary pixels' adjusting is rather satisfying. As to the residential areas, the approach proposed here shows some advantages because both patches with obvious texture feature and those with predominately grey feature can all be segmented rather successfully. In the study, we utilize a supervised approach to obtain the threshold of region growing, that is to say, some typical region are selected based on visual interpretation, and the heterogeneity of these region are calculated subsequently which can be referred as threshold. This simple method is rather feasible according to the successful segmentation of residential areas, crop lands, mangrove areas and beach.

Nevertheless, this method has demonstrated some weaknesses such as the oversegmentation effect that is a little bit obvious in the hilly woodland areas where the highly mixture of house and the vegetations does exist due to the shortage of construction land in Hong Kong. Also influenced by the sunshine and shadow, the segment in this area is much fragmented, which is a bit far from the reality.

As well known to us, to segment successfully the remotely sensed imagery is really a challenge due to the fact mentioned above. To us, this study is only an initial test, many problems still remain unsettled. For example, how to develop an

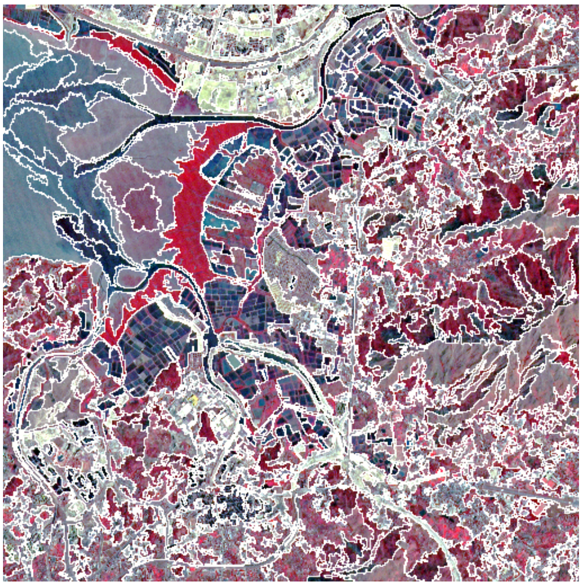

Figure 3. Fine segmentation results

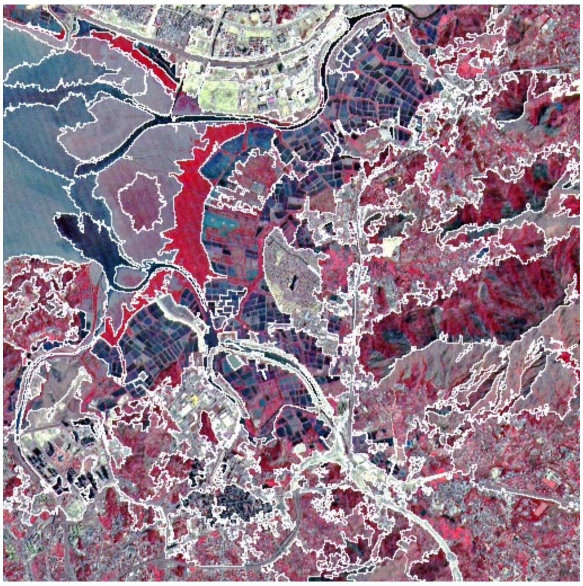

Figure 2. Coarse segmentation results after adjusting boundary pixels

efficient coarse segmentation approach and how to represent macro image feature are something deserve our attention in the further study. At the fine scale, computing efficiency should be improved further as well as the representation of texture feature in this scale. What's more, boundary pixel adjusting is time consuming and should be bettered.

V. CONCLUSIONS

Multiscale segmentation of remotely sensed images is a very important theme since scale effect does exist both in the nature and the society. As one of main information sources, remotely sensed images are playing an increasingly important role in many fields (Marceau and Hay 1999). Therefore, multiscale segmentation of remotely sensed images is a very important research field. In this study, a hybrid multiscale segmentation approach by taking the top-down strategy is proposed. From the experiment results, we found the approach is a bit promising. However, there still exists some problems to be settled, and further researches should be conducted in the future.

REFERENCES

[1] B. Gorte, *Probabilistic Segmentation of Remotely Sensed Images* (ITC Publication Series No. 63). Enschede: ITC Publication, 1998.

[2] M. Baatz and A. Schäpe, "Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation," in *Angewandte Geographische Informationsverarbeitung*, vol. xii, Strobl, Blaschke, and Griesebner, Eds. Heidelberg: Wichmann-Verlag, 12-23, 2000.

[3] T. Blaschke, and J. Strobl, "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS," *GIS – Zeitschrift für Geoinformationssysteme*, no.6, pp.12-17, 2001.

[4] Y. J. Zhang, "Image segmentation." Peking: Since Press, pp. 2, 2001.

[5] T. Blaschke, S. Lang, E. Lorup, et al., "Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications," in *Environmental Information for Planning, Politics and the Public*, A. Cremers, and K. Greve, Marburg: Metropolis-Verlag, pp.555-570, 2000,

[6] S. Ryherd, and C. Woodcock, "Combining spectral and texture data in the segmentation of remotely sensed images," *Photogrammetric. Engineering & Remote Sensing,* vol. 62, no. 2, pp. 181–194, 1996.

[7] J.C. Luo, Y. Leung and C.H. Zhou, "Scale Space based Hierarchical Clustering Method and its Application to Remotely Sensed Data Classification," *Acta Geodaetica et Cartographica Sinica*, vol. 28, no. 4, pp. 319-324, 1999. (in Chinese)

[8] J.C. Tilton, "Interactive analysis of hierarchical image segmentation," Proc. of IGARSS' 00, Honolulu, July 24-28, 2000.

[9] V. Lakshmanan, R. Rabin, V. DeBrunner, "Texture-Based Segmentation of Satellite Weather Imagery," *Proc. of ICIP 2000*, Vancouver, September 10-13, 2000.

[10] S. M. I. Vanhamel and H. Sahli, "Scale space segmentation of color images", *IRIS-Technical Report*, no. 0076, 2001.

[11] H.L. Resnikoff, and R.O. Wells, "Wavelet analysis: the scalable structure of information." New York: Springer-Verlag, 1998.

[12] [19] A. Moghaddamzadeh and N. Bourbakis, "A Fuzzy Region Growing Approach Segmentation of Color Images," Pattern Recognition, vol. 30, pp. 867-881, 1997.

[13] [2] J. M. Gauch, "Image Segmentation and Analysis via Multiscale Gradient Watershed Hierarchies," IEEE Trans. Image Processing, vol. 8, no. 1, pp. 69-79, 1999.

[14] [22]S. E. Hernandez and K. E. Barner, "Joint Region Merging Criteria for Watershed-Based Image Segmentation," Proc. International Conference on Image Processing, vol. 2, pp. 108- 111, 2000.

[15] S. Makrogiannis, G. Economou and S. Fotopoulos, "A Fuzzy Dissimilarity Function for Region Based Segmentation of Color Images," Int. Journal of Pattern Recognition and Artificial Intelligence, vol. 15, pp. 255-267, 2001.

[16] D. J. Marceau and G. J. Hay, "Remote Sensing Contributions to the Scale Isuue, Canadian Journal of Remote sensing," vol. 25, no. 4, pp. 357-366, 1999.