Multi-granularity synthesis segmentation for high spatial resolution Remote sensing images

Lina Yi, Pengfei Liu, Xiaojun Qiao, Xiaoning Zhang, Yuan Gao, Boyan Feng

College of GeoScience and Surveying Engineering, China University of Mining and Technology, Beijing, People’s Republic of China, 100083

Email: lina1986350@163.com

Abstract. Traditional segmentation method can only partition an image in a single granularity space, with segmentation accuracy limited to the single granularity space. This paper proposes a multi-granularity synthesis segmentation method for high spatial resolution remote sensing images based on a quotient space model. Firstly, we divide the whole image area into multiple granules (regions), each region is consisted of ground objects that have similar optimal segmentation scale, and then select and synthesize the sub-optimal segmentations of each region to get the final segmentation result. To validate this method, the land cover category map is used to guide the scale synthesis of multi-scale image segmentations for Quickbird image land use classification. Firstly, the image is coarsely divided into multiple regions, each region belongs to a certain land cover category. Then multi-scale segmentation results are generated by the Mumford-Shah function based region merging method. For each land cover category, the optimal segmentation scale is selected by the supervised segmentation accuracy assessment method. Finally, the optimal scales of segmentation results are synthesized under the guide of land cover category. Experiments show that the multi-granularity synthesis segmentation can produce more accurate segmentation than that of a single granularity space and benefit the classification.

1. Introduction

High spatial resolution remote sensing image (HSRI) segmentation is the first crucial step in object-based remote sensing image analysis, classification and object recognition. The intent of HSRI segmentation is to produce meaningful objects that match objects on the ground [1]. While landscapes are composed of layers of patterns that occur from a variety of natural and human-driven process operating over vastly different spatial and temporal frequencies [2], there often does not exist a single scale of segmentation that could be deemed appropriate for analysis of the entire image. Multi-scale segmentation therefore becomes necessary to get meaningful objects for complex landscape analysis, ecological system analysis [3] etc. Though multi-scale segmentation techniques can provide interesting and useful results over a single scale or narrow range of scales for certain object extraction, it is always not enough for analysis of the scene due to the complexity of natural landscapes. Intelligently combine objects at different optimal scales together is another crucial issue to get the appropriate segmentation result that can accurately characterize ground objects of varying scale properties. With respect to this issue, how to link the multi-scale objects and combine sub-optimal scales of segmentation together are two key problems that should be solved.

Quotient space granular theory which comes from the human cognition science has been regarded as a unified framework for theories, methodologies and techniques that make use of granules in the...
process of problem solving. Here a granule can be seen as a clump of objects (points), in the universe of discourse, drawn together by indistinguishability, similarity, proximity, or functionality. Granulation leads to information compression/summarization. In situations involving incomplete, uncertain, or vague information, it may be difficult to differentiate different elements and instead it is convenient to consider granules, i.e., clump or group of indiscernible elements, for performing operations [4]. Based on granular theory, we can solve the complex HSRI multi-scale segmentation problem in an intelligent way. The HSRI image can be firstly divided into different ‘granules’ (regions) based on the quotient space model of multi-granularity image segmentation, each granule is assumed to be consisted of ground objects that have similar optimal segmentation scale. And then the sub-optimal segmentations of each region can be selected and synthesized together by the granular synthesis technique to get the final segmentation result. The result is the combination of sub-optimal scales of objects and is therefore more coherent to ground objects.

To validate this method, the land cover category map is used to guide the scale synthesis of multi-scale image segmentations for Quickbird image land use classification. Experiments show that the multi-granularity synthesis segmentation can produce more accurate segmentation than that of a single granularity space and benefit the classification.

In the following text, we firstly illustrate the Quotient space model of multi-granularity image segmentation. Next, the granular synthesis technique for HSRI segmentation is illustrated. Then, we show the object-oriented land use classification experiments of Quickbird images to validate the methods. Finally, we summarize the results and draw conclusions.

2. Quotient space model of multi-granularity image segmentation

The aim of image segmentation is to divide the image into non-overlapping regions, each region is homogenous inside according to certain homogeneity rule defined based on the pixel attributes, and the neighboring regions are of heterogeneity. Suppose X is the pixel set of an image, R is the homogeneity rule, image segmentation is to divide X into many subsets (regions) \{X_1,X_2,…X_n\}, each of this subset satisfies the following four conditions:

1. \(X = \bigcup_{i=1}^{n} X_i\)

2. If \(i \neq j\), \(X_i \cap X_j = \emptyset\)

3. For each subset \(X_i\), \(R(X_i) = true\)

4. If \(i \neq j\), and \(X_i\) is the neighbor of \(X_j\), \(R(X_i \cup X_j) = false\)

It can be concluded that though different image segmentation algorithms may have distinct characteristics that can be useful for segmenting different kinds of images for various tasks, they all have the common nature of dividing the pixel set into different regions according to certain homogeneous rules in respect with certain features or neighboring relationships of pixels.

Image segmentation can be seen as a special granular computing problem that can be solved based on the Quotient space granular theory. It can be done by defining an image as a collection of pixels and the equivalence relation, that defines the homogenous rules, induced partition as pixels lying within each non-overlapping segmented object over the image. Accordingly, it is possible to describe the multi-granularity image segmentation problem by a quotient space model, in which the granular computing principle and methods construct a unified theory scheme for image segmentation.

In the quotient space model, if we use a three element set \((X, f, \Gamma)\) to define the original image, \(X\) is the domain of discourse, \(f\) is the attribute function of each pixel in \(X\), \(\Gamma\) is the structural relationships among pixels in \(X\). Then, after using a predefined equivalence relationship to segment the image, the segmented image is defined as \((\{X\},\{f\},\{\Gamma\}\), \([X]\) represents the segmented object set, \([f]\) represents the attributes of the objects in \([X]\), and \([\Gamma]\) represents the structural relationships among the objects in \([X]\).
In this way, before segmentation, the single pixel is the basic element in $X$, and after segmentation, the pixel set is the basic element in $\{X\}$. In fact, multi-scale image segmentation is a process of producing multi-granular space $\langle\{X\},\{f\},\{\Gamma\}\rangle$ by defining different equivalence relationships and segments the image.

After get the multi-granularity segmentations based on the above quotient space model, we can use the granule synthesis principle to synthesize the sub-optimal granules $\{X\}$ (segmented object sets) to get the final segmentation result which can take into account the distinct scale properties of different ground objects which are characteristic with distinct heterogeneity, aggregation or amplitude. The main idea is to firstly segment the image into multi-granules, represented as $(X_1,f_1,\Gamma_1),$ $(X_2,f_2,\Gamma_2),$ \cdots $(X_m,f_m,\Gamma_m),$ and then select the sub-optimal granules and synthesize the selected granules to get the final segmentation result which meets the segmentation requirement of different applications.

Here, the multi-granule synthesis is implemented by repeatedly synthesizing two granules many times; the synthesis of two granules is described as follows.

In synthesizing $(X_1,f_1,\Gamma_1)$ and $(X_2,f_2,\Gamma_2)$ to get the result of $(X_3,f_3,\Gamma_3)$, the following conditions are satisfied based on Quient space granule theory:

1. $X_1, X_2$ are the quotient space of $X_3$,
2. $\Gamma_1, \Gamma_2$ are the quotient structures of $\Gamma_3$ in respect with $X_1$ and $X_2$ respectively
3. $f_1, f_2$ are the projections of $f_3$ in $X_1, X_2$ respectively. $(X_3,f_3,\Gamma_3)$ satisfies a certain optimal rule, which is defined based on the overlay operation based on the topological structure of each granule.

To implement the synthesize procedure, we can define the synthesized topological structure $\Gamma_3$ based on $\Gamma_1, \Gamma_2, \Gamma_3$ is the smallest upper bound of $\Gamma_1, \Gamma_2$. The detailed granule synthesis steps are as follows:

Define $B$ as the basic topological structure, then $B$ can be used to construct the topological structure of set $\Gamma_3$ by combining all the elements in $B$ together.

For example, define a remote sensing image as a domain of discourse $X$, then for a segmentation result, $\{X\}$ is the assembled set of the segmented objects. As shown in figure 1, if one segmentation result a is the assembled set of $X_1$ defined as \{1,2,3,4\}, another segmentation result b is the assembled set of $X_2$ defined as \{1,2,3\}, then in the segmentation result a and b, $\Gamma_1$ and $\Gamma_2$ respectively is constructed by the assembled set of the corresponding segmented objects. If synthesize $\Gamma_1$ and $\Gamma_2$, then the basic topological structure $B$ is defined as follows:

$$B = \{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\{7\},\{8\},\{9\},\{10\},\{11\},\emptyset, X\}$$

Where, $\{i\}$ represents region i in figure 1(c), i=1, 2,\ldots 11. So $\Gamma_3$ is all the possible assembled set of the elements in $B$.

In HSRI segmentation, we use the granule synthesis methods to simplifying the segmentation. There are three main procedures: (1) generation of multi-granules by multi-scale segmentation; (2) selection of the sub-optimal granule for segmentation of different ground objects; (3) granule synthesis. The methodology of each step in granule synthesis is detailed in the next section.
3. Experiment

To validate the applicability of the proposed method, the land cover category map is used to guide the scale synthesis of multi-scale image segmentations for Quickbird image land use classification. Firstly, the image is coarsely divided into multiple regions, each region belongs to a certain land cover category. Then multi-scale segmentation results are generated by the Mumford-Shah function based region merging method. For each land cover category, the optimal segmentation scale is selected by the supervised segmentation accuracy assessment method. Finally, the optimal scales of segmentation results are synthesized under the guide of land cover category. For the details of the methodology used in the experiment, please refer to paper [5], in the following, we only describe the main idea of the methods and the experiment results.

3.1. Generation of multi-granules

An edge-embedded marker based watershed segmentation algorithm (EEMW) [6] is used to derive the initial oversegmentation result of HSRI. This method performs well in both retaining the weak boundary and reducing the undesired over-segmentation. In the over-segmentation result, the minimum mapping unit for each class being mapped are well segmented with all the object boundary pixels retained. Then, the Mumford-Shah segmentation model based region merging algorithm [7] is implemented to obtain multi-scale segmentation results. In this method, image segmentation is viewed as a compromise between the shape characteristic of region boundary and the fitting error of the region enclosed by the boundary. A good segmentation therefore is a very efficient description of ground object boundaries with consideration of the associated errors. By setting different threshold, multi-scale of segmentation results can be obtained. As the threshold increase, coarser segmentation result can be generated based on the finest scale of segmentation. This eases the latter granule synthesis procedure.

3.2. Sub-optimal granule selection

The optimal granule (segmentation scales) of different land cover categories should be selected before granule synthesis. We use the supervised image segmentation accuracy assessment measures proposed by Crevier [8] to guide the optimal segmentation scale selection. In this method, the segmentation samples of each land cover category are firstly extracted. Then, the segmentation scale with highest accuracy is selected as the optimal scale for the corresponding land cover category according to the related segmentation samples. In this way, different segmentation requirements can be satisfied by adjusting the segmentation samples accordingly to the application purposes.

3.3. Granule synthesis

Granule synthesis is a process of synthesizing the sub-optimal granules of different land cover categories together. In the process, the LCCM is extracted in advance to guide the latter scale synthesis. To get LCCM, the features of typical ground objects are firstly analyzed to construct classification
rules for classifying the initial over-segmented objects into different land cover categories. Then, a post-processing is implemented to remove the small undesired objects and merge neighboring objects with the same land cover category.

Guided by LCCM, the optimal segmentation results are synthesized by the following two steps:

1. For each optimal granule (segmentation result) of a land cover category, mask the objects that do not belong to this land cover category.

2. Synthesize all the land cover categories’ optimal granules by the granule synthesis method as shown in section 2. In the process, the topological structure construction is implemented by layer overlaying.

3.4. Experiment results

The study area mainly includes ten typical ground objects, including water, forest, residential area, mining area, road, concrete area, paddy field, vegetable plot, bare soil and grass according to the land investigation. These classes can be categorized into four main land cover categories: water, forest, artificial area, and farm field. To derive the land cover category map, the sample features of typical ground objects are extracted and analyzed to establish the classification rule. The samples of each ground object classes are gathered manually based on land survey. These samples are divided into different land cover categories respectively to provide segmentation samples for the segmentation accuracy assessment in selection of sub-optimal granules. For the detail of the experiment steps, please refer to paper [5], here we only presents the synthesized segmentation result and the classification result as shown in figure 2 and 3. The experiment result shows that synthesize the optimal segmentations together can improve the final segmentation accuracy quite well. This will benefit the classification of the study area. The overall accuracy is 77.8% and the Kappa index is 0.727.

4. Conclusion

This paper proposes a quotient space multi-granularity segmentation model to simplify HSRI image segmentation by the granule synthesis method. In this method, an edge-embedded marker-based watershed algorithm is firstly implemented to get an initial over-segmentation result. Then, a bottom-up region merging method is implemented with a Mumford-Shah segmentation model to establish a linking hierarchy multi-scale segmentation network. Among the multi-scale segmentation results, the optimal scale of different land cover category is selected by a supervised segmentation accuracy assessment method. The
final segmentation result is generated by synthesizing the selected segmentation results together. Because the scale synthesis method takes into account the scale characteristics of different land cover categories, it provides high quality segmentation result for latter classification.

5. Acknowledge
The research was supported by the National Undergraduate Innovation Training Programs (No. 201211413030) and the Fundamental Research Funds for the Central Universities (NO. 2011QD03).

References
[1] Woodcock C and Harward V 1992 Nested-hierarchical scene models and image segmentation *Int J Remote Sens* 16 3167–87
[2] Karl J and Maurer B 2010 Multivariate correlations between imagery and field measurements across scales: comparing pixel aggregation and image segmentation *Landscape Ecol* 4 591–605
[3] Burnett C and Blaschke T 2003 A multi-scale segmentation/object relationship modelling methodology for landscape analysis *Ecological Modelling* 3 233–49
[4] Pal S K Shankar B U and Mitra P 2005 Granular computing, rough entropy and object extraction *Pattern Recognition Letters* 26 2509-17
[5] Yi L N Zhang G F and Wu Z C 2012 A scale synthesis method for high spatial resolution remote sensing image segmentation *IEEE Trans GeoSci Remot Sen* 10 4062-70
[6] Li D R Zhang G F and Yi L N 2010 An Edge Embedded Marker-Based Watershed Algorithm for High Spatial Resolution Remote Sensing Image Segmentation *IEEE Trans Image Processing* 10 2781- 87
[7] Robinson D J Redding N J and Crisp D J 2002 Implementation of a fast algorithm for segmenting SAR imagery *DSTO Electronics and Surveillance Research Laboratory Report*
[8] Crevier D 2008 Image segmentation algorithm development using ground truth image data sets *Comput Vis Image Understand* 2 143–59