*Research Article*

# Study on Semantic Contrast Evaluation Based on Vector and Raster Data Patch Generalization

**Jun Yang,**[1,2] **Yuechen Li,**[1] **Jianchao Xi,**[2] **Chuang Li,**[1] **and Fuding Xie**[1]

[1] *Liaoning Normal University, Liaoning Key Laboratory of Physical Geography and Geomatics, Dalian 116029, China*
[2] *Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China*

Correspondence should be addressed to Jun Yang; yangjun@lnnu.edu.cn

We used buffer superposition, Delaunay triangulation skeleton line, and other methods to achieve the aggregation and amalgamation of the vector data, adopted the method of combining mathematical morphology and cellular automata to achieve the patch generalization of the raster data, and selected the two evaluation elements (namely, semantic consistency and semantic completeness) from the semantic perspective to conduct the contrast evaluation study on the generalization results from the two levels, respectively, namely, land type and map. The study results show that: (1) before and after the generalization, it is easier for the vector data to guarantee the area balance of the patch; the raster data's aggregation of the small patch is more obvious. (2) Analyzing from the scale of the land type, most of the land use types of the two kinds of generalization result's semantic consistency is above 0.6; the semantic completeness of all types of land use in raster data is relatively low. (3) Analyzing from the scale of map, the semantic consistency of the generalization results for the two kinds of data is close to 1, while, in the aspect of semantic completeness, the land type deletion situation of the raster data generalization result is more serious.

## 1. Introduction

With the development of computer technology and the constant deepening of the GIS application technology, the traditional paper maps are gradually being replaced by digital maps [1, 2], while as a frontier and hot issue in cartography, cartographic generalization has also faced unprecedented changes [3]. The cartographic generalization refers to the process where, under the premise of maintaining the structures and characteristics of the spatial entities, the extraction and processing are conducted for the map data of the cartographic regions through appropriate selection, generalization, and other operations according to the factors such as scale and use of the map as well as geographical characteristics of cartographic regions, so as to finally achieve the purpose of passing more and more important spatial information on to the limited representation media [4–7].

Due to factors such as the various changes of geographical space itself and the relative uncertainty of generalization results, the cartographic generalization has always been a difficult international problem in the field of geographical science. Up to the early 1920s, many map scholars at home and abroad have been engaged in the theoretical and practical research of cartographic generalization and have achieved fruitful results since M. Echert first put forward the term "cartographic generalization" [4, 8]. The structures of the vector data are intuitive and simple, with each specific target element being directly endowed with spatial position and attribute information, and the vector data having natural advantages in the aspect of calculating the quantitative and qualitative indicators of elements such as distance, area, and topological relations. Therefore, the research emphases of most scholars were mainly focused on the vector data. The "decomposition type" combination method based on Delaunay triangulation and skeleton line structure was proposed for maintaining the area balance of the patch; the automatic generalization method based on genetic algorithm was proposed for the overall optimal configuration of the punctate element annotation; and the Douglas-Peucker algorithm, declination algorithm, and rounding algorithm were proposed for the linear distribution elements [7, 9]. The raster data divide the space into regular grids, with the cell value of each grid

having attribute information, and the row-column number being the position information. From the perspective of data structure, it is easier for the raster data to conduct the simple and fast neighborhood analysis and the construction of the mathematical model. In the early 1980s, Monmonier had applied mathematical morphology to the research on the cartographic generalization of the planar elements and proposed that the raster data structure was more suitable for research on the cartographic generalization of land use [10]; subsequently, Su et al. discussed the processing methods of the raster data such as feature simplification, integration, and displacement using mathematical morphology [11, 12]; based on the mixed data of the raster vector, Huilian et al. proposed the GABP neural network model and achieved the simplification of the buildings [13]; at the beginning of this century, Li et al. of British Kingston University first applied cellular automata to the cartographic generalization, bringing progress and breakthrough to the cartographic generalization based on the raster mode [14].

On the basis of the research results by the above-mentioned scholars, according to the different characteristics of the vector data and raster data, we used the corresponding generalization methods, conducted generalization for the land use patches of the same cartographic region, and carried out the semantic contrast evaluation for the generalization results. Among them, for the vector data, we comprehensively considered the auxiliary spatial topological relations of the land type in order to achieve the amalgamation of the adjacent patches and established the polygon through the buffer intersection nodes to achieve the aggregation of the adjacent patches; for the raster data, we used the closing operation in the mathematical morphology to achieve the aggregation of the patches and added the semantic conception of the cellular automata operation of the mode filtering operation rules to achieve the generalization.

## 2. Vector Data Patch Generalization Algorithm Process

*2.1. Definition of Topological Relation.* Since the spatial distribution of the land use data is characterized by full coverage, no overlap and no gap, the generalization of the land use data inevitably cannot dispense with the aggregation and amalgamation operations of the patch polygon. The multilevel semantic features make the decision-making of the comprehensive guidelines more complex. In this research, we gave comprehensive consideration to the cartographic generalization of the auxiliary topological adjacent patches of the patch category.

In the land use data, we selected the set LandSet$\{F_i\}$ of the polygons whose patch area $S_i$ is less than the smallest epigraph area $S_{\text{Area}}$ and set $B_i$ as the patch boundary. $D_{\text{distance}}$ is the minimum distance between the elements on the map, $D(F_i, F_j)$ is the distance between patch $F_i$ and patch $F_j$, and Area$(F_i)$ is the area of the element $F_i$. The topological relation of $F_i$ and $F_j$ within the set LandSet$\{F_i\}$ is defined as follows:

① if $B_i \cap B_j \neq \emptyset$, $F_i$ and $F_j$ are topological adjacency;

② if $B_i \cap B_j = \emptyset$ and $D(F_i, F_j) \leq D_{\text{distance}}$, and $F_i$ and $F_j$ are topological adjacency.

In the generalization process, the merging processing is only conducted directly for the similar patches which satisfy the condition ①, while further category analysis should be conducted for other conditions, before comprehensive treatment can be continued [15].

*2.2. Aggregation Processing.* A merging operation needs to be conducted for the parted patches on the space, so as to prevent the similar small patches in close proximity to each other from being deleted in the generalization process to result in too large of an area change of the generalization result. This operation process is known as aggregation. The specific process is as follows: as shown in Figure 1, $F_1$ and $F_2$ are the patches of the same land type. First, as shown in Figure 1, draw the buffers of $F_1$ and $F_2$ with the minimum distance between elements ($D_{\text{distance}}$) as the radius, which are, respectively, Buffer$(F_1)$ and Buffer$(F_2)$; obtain the intersection of the two buffers Buffer$(F_1) \cap$ Buffer$(F_2)$ (the grid section in Figure 1(b)); then, extract the nodes of the polygons $F_1$ and $F_2$ in the intersection of the buffers (Buffer$(F_1) \cap$ Buffer$(F_2)$) and establish the polygon $F_{\text{new}}$ according to the nodes (the red area as shown in Figure 1(c)); finally, conduct merging processing for $F_1$, $F_2$, and $F_{\text{new}}$ and delete the overlap area generated in the merging result [16]. The aggregation result is shown in Figure 1(d).

*2.3. Amalgamation Processing.* Aiming at the nonsimilar patch ($F_0$) of topological adjacency whose patch area is less than the smallest epigraph area, the method of extracting the Delaunay triangulation skeleton lines is used to conduct the subdivision processing for $F_0$. The secondary patch $F_0$ is decomposed according to the skeleton lines; after decomposition, the small patches are, respectively, merged into the adjacent main patches, and their original land use type names are changed [17–19]. The specific process is shown in Figure 2. $F_1$, $F_2$, $F_3$, and $F_4$ are the patches of different land types, and the area of $F_1$ is less than the smallest epigraph area. The Delaunay triangulation is made for $F_1$, the subdivision is conducted for $F_1$ according to the skeleton lines (as shown in Figure 2(b)), and the generalization result is shown in Figure 2(c).

## 3. Theoretical Basis and Algorithm Process of Raster Data Patch Generalization

*3.1. Mathematical Morphology.* Mathematical morphology is a discipline with complete mathematical foundations, established on the basis of set theory; its main idea being to use a structural element with certain size to detect the geometrical shapes in the images. The most basic mathematical morphology operators include erosion, dilation, opening operation, and closing operation. In the cartographic generalization, if there are tiny connections between two patches and the structural element is large enough, the erosion operation can be used to separate them; if there are small gaps between two patches and the spacing is less than the structural element,
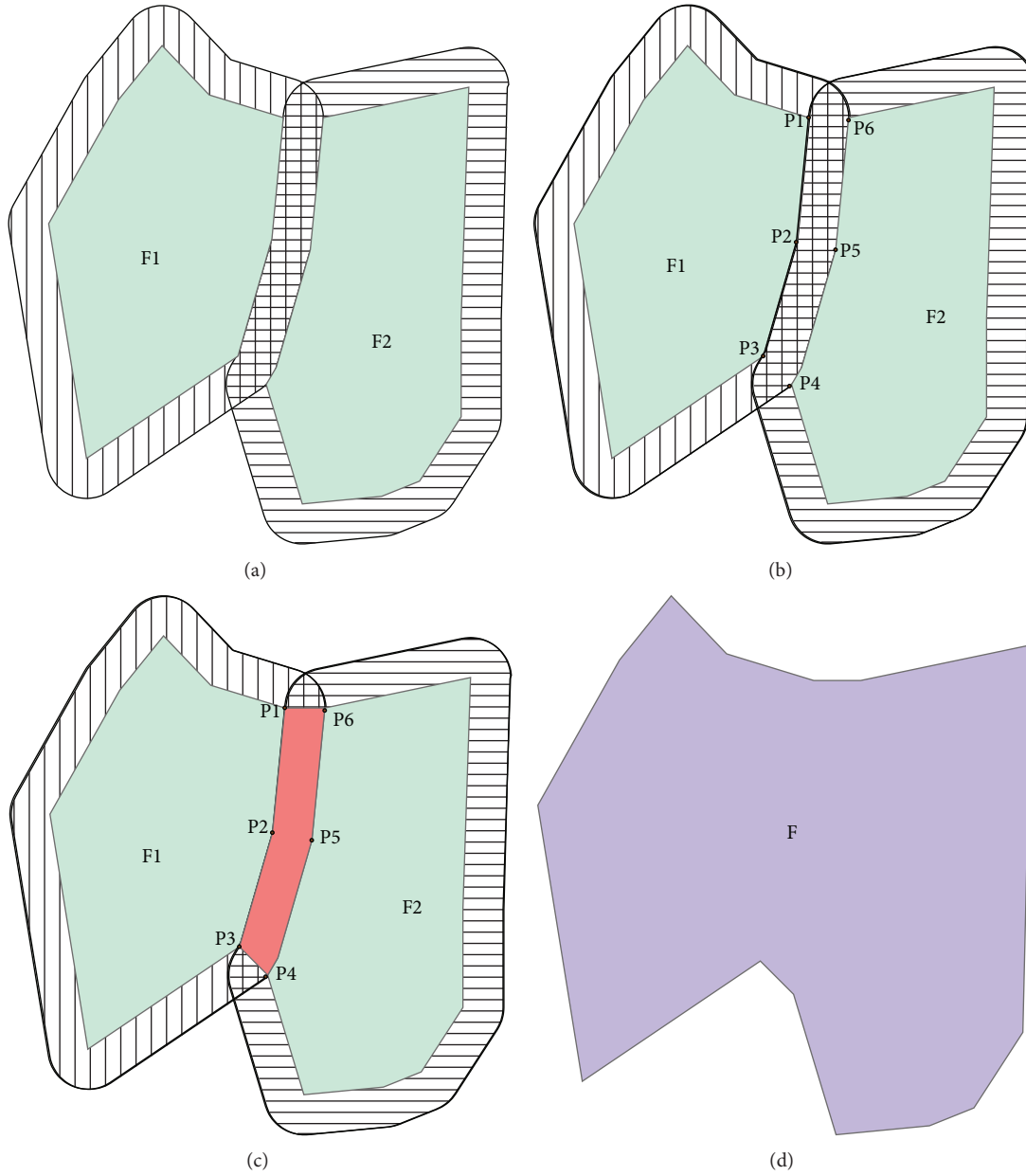
Figure 1: Patch aggregation.

the dilation operation can be used to achieve the connection. Their operational formulas are as follows:

$$A\Theta B = \{a : a + b \in A, b \in B\},$$
$$A \oplus B = \{a + b : a \in A, b \in B\}.$$
(1)

In the formula, $A$ is the set of the pixel points to be processed, $B$ is the structural elements used for detecting the geometrical shapes of the images, $A\Theta B$ represents that $A$ is eroded by the structural element $B$, and $A \oplus B$ represents that the set $A$ is dilated by the structural element $B$. The detailed process is shown in Figure 3. Figure 3(a) shows that there are only the binary images of 0 and 1 in the set elements to be

processed, and Figure 3(b) represents the structural elements of the four neighborhood pixels of the focus. The translation is conducted successively for the structural element $B$ along the row-column of the binary image $A$. When the center pixel of the structural element $B$ overlaps with the pixels whose pixel value is 0 in the binary image $A$, its four neighborhoods will be eroded, the pixel value will become 0, and the eroded pixels are represented by "−" (as shown in Figure 3(c)); on the contrary, when the center pixel of the structural element $B$ overlaps the pixels whose pixel value is 1 in the binary image $A$, its four neighborhoods will be inflated, the pixel value will become 1, and the inflated pixels are represented by "+" (as shown in Figure 3(d)).
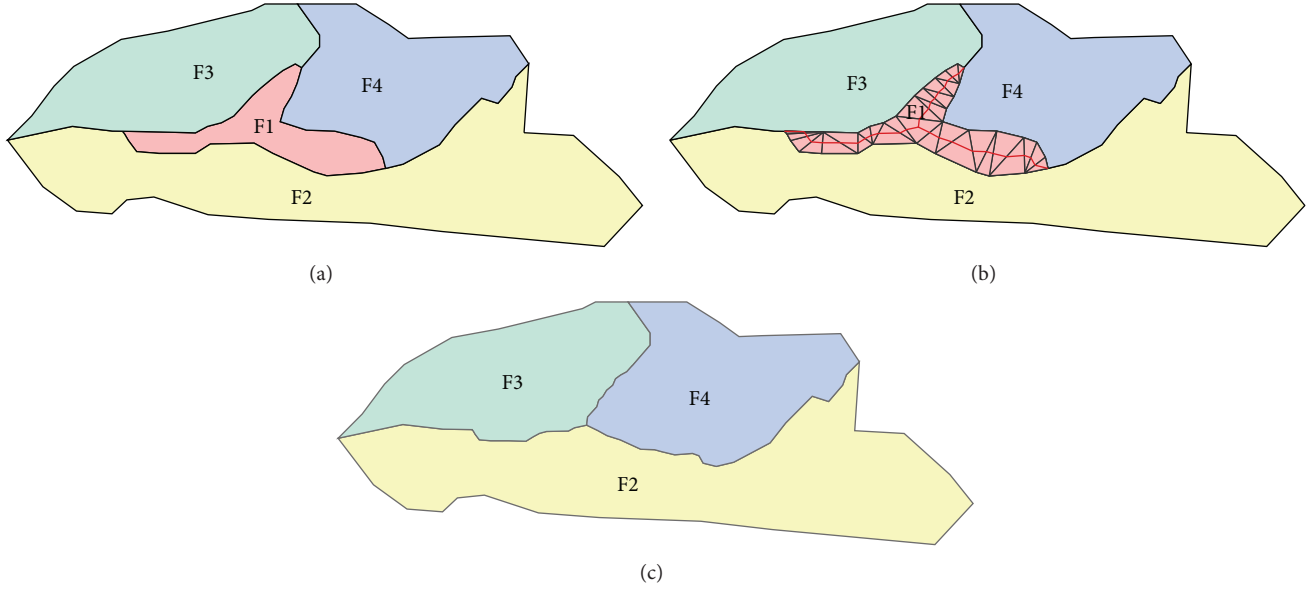
(a)

(b)

(c)

FIGURE 2: Patch amalgamation.

In the cartographic generalization, for the patches whose rasterized distance is close but parted, the single dilation erosion operation cannot achieve the effect of patch aggregation. Generally, several dilation operations need to be conducted first, before the erosion operations of corresponding times can be conducted, and this process is known as the closing operation in mathematical morphology, while the operation process of the opening operation is the exact opposite to that of the closing operation. Their formulas are as follows:

$$\text{Closing operation: } A \cdot B = ((A \oplus B) \ominus B),$$
$$\text{Opening operation: } A \bigcirc B = ((A \ominus B) \oplus B). \tag{2}$$

In practical application, the number of times of erosion and dilation needs to be determined according to the distance of the aggregation land type. Taking the ratio operation as an example, if the distance of the two patches is $n$, it needs to first conduct the dilation operation for $(n/(2^*\text{cellsize}))$ times and then conduct the erosion operation of the corresponding times in order to achieve the aggregation operation.

*3.2. Cellular Automata.* Although the algorithm of the mathematical morphology has the unique advantages of the natural parallel implementation structures, the feature generalization of a single attribute can only be conducted in this algorithm. Obviously it is difficult to cope with the multisemantic land use data generalization by using mathematical morphology alone. Therefore, it needs to introduce the cellular automata with "mode filtering" as the rule of transformation to achieve the patch amalgamation.

The cellular automata is a kind of grid dynamics model in which the cellular units with discrete and limited state interact in the local space and evolve in the discrete time dimension [20]. It includes four basic elements, namely, cellular, state, neighborhood, and transformation rule. The

raster pixel in the research is the cellular, the attribute category of the pixel is the cellular state, all the cellular whose distance to a cellular is within the scope of the determined radius $r$ are the neighborhood of the cellular, and the function used for controlling the change of the cellular state (mode filtering) is the transformation rule. In the generalization process, according to the characteristics of the mode filtering algorithm, the $5 \times 5$ neighborhood is used as model to traverse the full figure, with all the raster values within the scope of the center cellular neighborhood being extracted, and the raster values which appear the most number of times are maximum in the neighborhood and are used as the center raster values of the next moment. The land use data generalization based on cellular automata is a gradual process. Therefore, many iterative operations need to be conducted so as to achieve the ideal state of generalization.

## 4. Semantic Evaluation Theory

In the map generalization process, the operations such as merging, exaggeration, and deletion can not only change the geometrical and topological relations between the elements, but also change the semantic relation fundamentally. In the research, we did not excessively consider the complex semantic relations between the land use data. In order to contrast the generalization results of the vector and raster data, we selected the two evaluation elements (namely, semantic consistency and semantic completeness) from the semantic perspective and conducted the comparative analysis and evaluation for the two generalization results, respectively, from the levels of land type and maps.

*4.1. Semantic Consistency.* The degree of consistency with the semantic constraint is known as semantic consistency. In the level of land type, taking the land type $C_k$ as an example, $S_1$
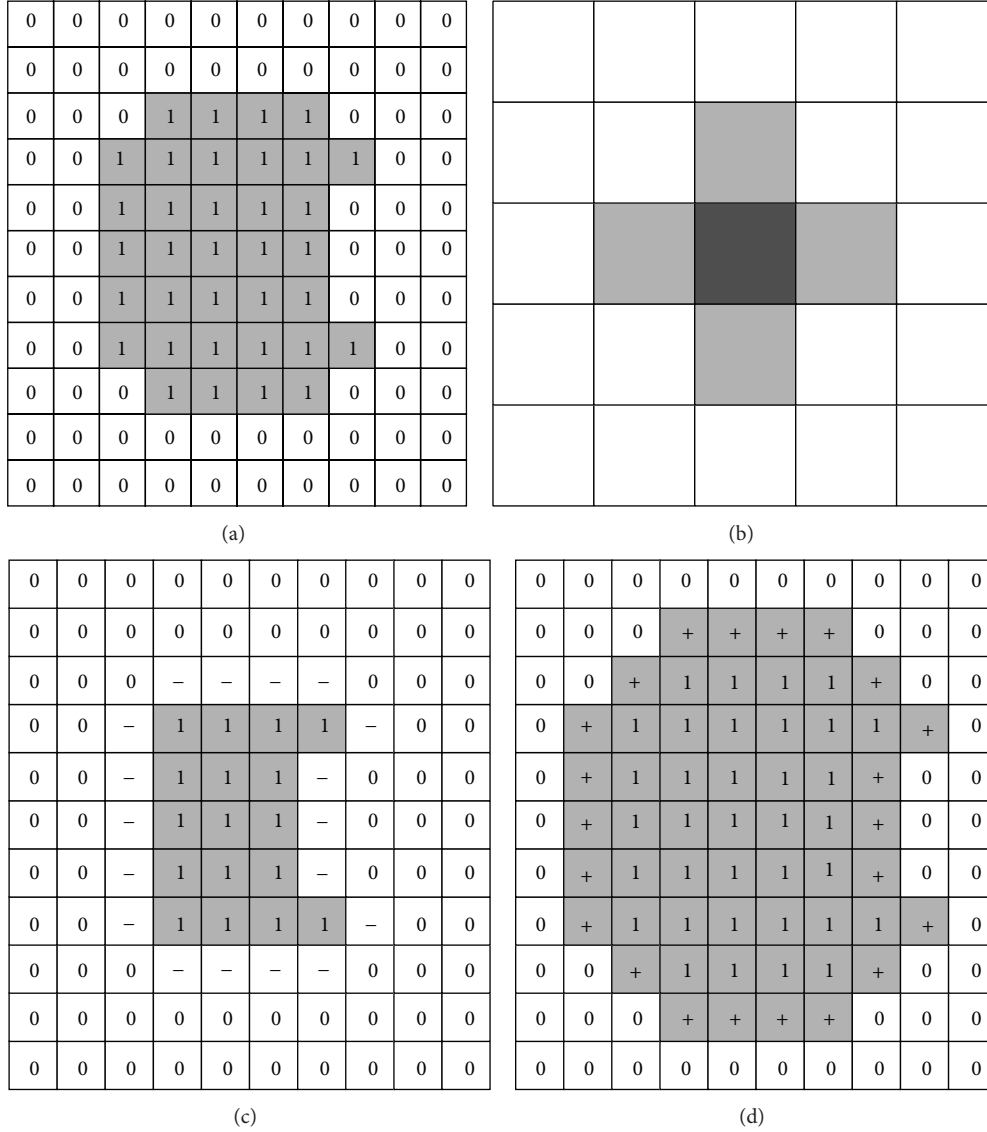
**(a)**

| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**(b)** (structuring element, cross shape)

**(c)**

| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | − | − | − | − | 0 | 0 | 0 |
| 0 | 0 | − | 1 | 1 | 1 | 1 | − | 0 | 0 |
| 0 | 0 | − | 1 | 1 | 1 | − | 0 | 0 | 0 |
| 0 | 0 | − | 1 | 1 | 1 | − | 0 | 0 | 0 |
| 0 | 0 | − | 1 | 1 | 1 | − | 0 | 0 | 0 |
| 0 | 0 | − | 1 | 1 | 1 | 1 | − | 0 | 0 |
| 0 | 0 | 0 | − | − | − | − | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**(d)**

| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | + | + | + | + | 0 | 0 | 0 |
| 0 | 0 | + | 1 | 1 | 1 | 1 | + | 0 | 0 |
| 0 | + | 1 | 1 | 1 | 1 | 1 | 1 | + | 0 |
| 0 | + | 1 | 1 | 1 | 1 | 1 | + | 0 | 0 |
| 0 | + | 1 | 1 | 1 | 1 | 1 | + | 0 | 0 |
| 0 | + | 1 | 1 | 1 | 1 | 1 | + | 0 | 0 |
| 0 | + | 1 | 1 | 1 | 1 | 1 | 1 | + | 0 |
| 0 | 0 | + | 1 | 1 | 1 | 1 | + | 0 | 0 |
| 0 | 0 | 0 | + | + | + | + | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

FIGURE 3: The operation of erosion and dilation.

and $S_2$ are used to, respectively, represent the total area of the land type $C_k$ before and after generalization, and then the calculation formula of the semantic consistency $\rho_C$ of the land type $C_k$ can be expressed as follows:

$$\rho_c = 1 - \frac{|S_2 - S_1|}{S_1}. \tag{3}$$

Further expansion is conducted for Formula (3), and the semantic consistency of the whole map on the map level can be expressed as follows:

$$\rho = 1 - \frac{\Delta S}{S_0}, \tag{4}$$

where $\Delta S$ represents the sum of the absolute value of the area change of each land type before and after the generalization, and $S_0$ represents the sum of the total area of various patches

before the generalization. Whether it is the land type level or the map level, the range of value of the semantic consistency is $[0, 1]$. The higher the value is, the better the semantic consistency is, and the better the generalization effect is.

### 4.2. Semantic Completeness.

The semantic redundancy and degree of omission are known as semantic completeness. Taking land type $C_k$ as an example, according to the deletion of the target before and after the generalization, the semantic completeness in the level of land type is defined as follows:

$$\tau_c = 1 - \frac{\Delta N_c}{N_c}, \tag{5}$$

where $\Delta N_c$ is the deletion number of the land type $C_k$ after the generalization and $N_c$ is the number of belonging to the land type $C_k$ before the generalization. Further expansion is
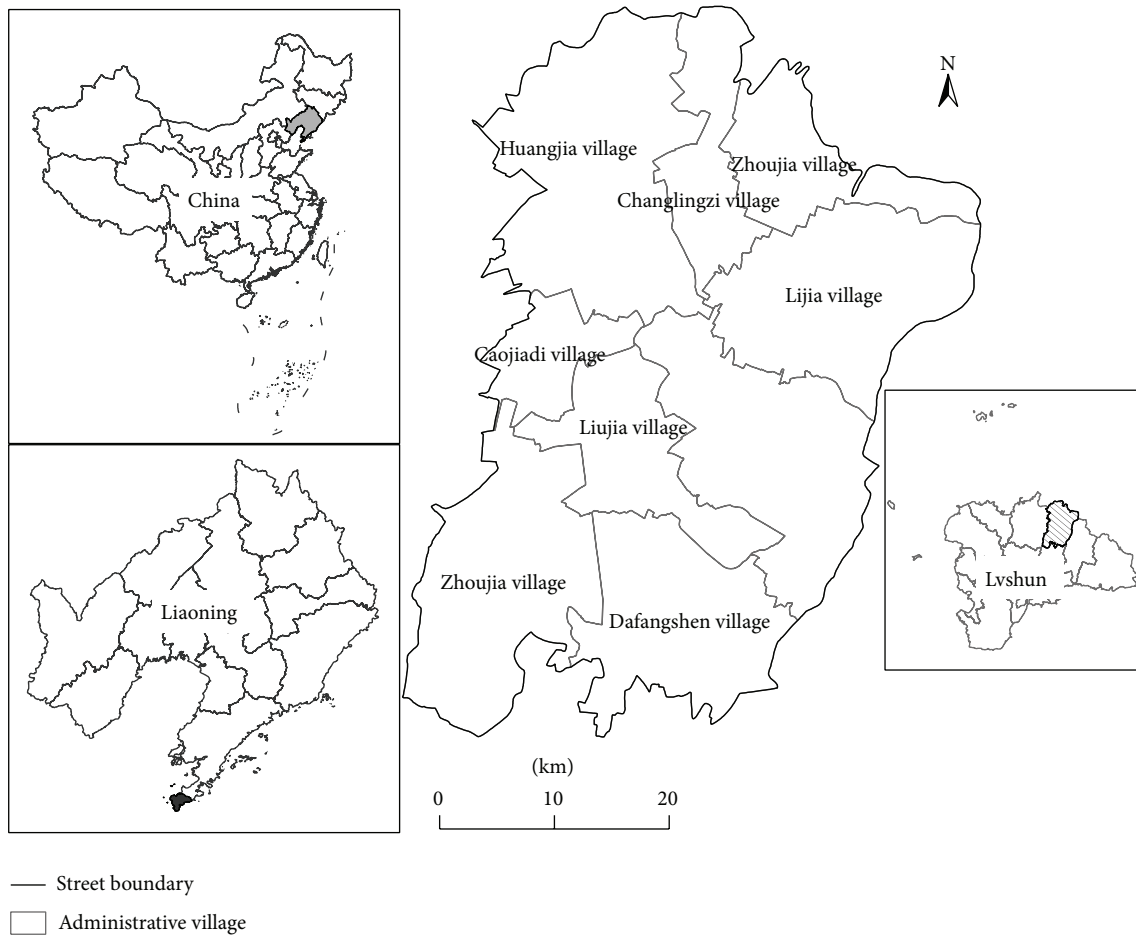
FIGURE 4: Location of study area.

conducted for Formula (5), and the semantic completeness of the map level can be expressed as follows:

$$\tau = 1 - \frac{\Delta N}{N_0}. \qquad (6)$$

In the formula, $\Delta N$ is the total number of the deletion targets of each land type after generalization and $N_0$ is the total number of the targets of each land type before generalization. The value range of the semantic completeness is $[0, 1]$. The closer to 1 the value is, the higher the semantic completeness is; on the contrary, category deletion is serious.

## 5. Case Study and Analysis of Results

In the research, we selected the patch data with a scale of 1 : 10000 of Changcheng Street, Lvshunkou District, Dalian City, as the experimental data of the research. As shown in Figure 4, Changcheng Street is located in the northeastern part of the Lvshunkou District; nine administrative villages, namely, Huangjia Village, Changlingzi Village, Zhaojia Village, Caojiadi Village, Lijia Village, Liujia Village, Zhongjia Village, Zhoujia Village, and Dafangshen Village, are under the jurisdiction of Changcheng Street; the total area is 3,266 hm$^2$. The data for land use status of Changcheng Street

is dominated by the farmland, and the unused land is dominated by the natural reserve. The distribution of the types of land use is relatively scattered, basically being able to reflect the patch generalization method proposed in the research.

According to the planning requirements of the land use, the data scale is integrated from 1 : 10000 to 1 : 50000 in the generalization process. Before the generalization, the smallest area on the map of the original vector data is 400 m$^2$. According to the cartographic generalization requirements of the 1 : 50000 scale land use data of the overall planning of the land use in the period from 2006 to 2020, we selected 10,000 m$^2$ as the smallest epigraph area and 30 m as the maximum distance of the patch aggregation. The iteration was conducted for the raster data with the resolution of 5 × 5, and the images tended to stabilize after the iteration was conducted eighty times.

The secondary component development of ArcGIS Engine under the VS 2010 programming environment was used to achieve the generation of the patch skeleton lines in the vector data, and the patch generalization of the vector data was completed in combination with ArcGIS; the scientific computing package (numpy) of the Python script program was used to achieve the integrated mode filtering program of combining mathematical morphology and cellular automata,

Cropland
Garden land
Forest land
Other agricultural lands
Urban land
Rural residential area

Mining land
Traffic and water conservancy land
Other construction lands
Water
Natural reserve land

(a) Original land use data

Cropland
Garden land
Forest land
Other agricultural lands
Urban land
Rural residential area

Mining land
Traffic and water conservancy land
Other construction lands
Water
Natural reserve land

(b) Result of vector data generalization

Cropland
Garden land
Forest land
Other agricultural lands
Urban land
Rural residential area

Mining land
Traffic and water conservancy land
Other construction lands
Water
Natural reserve land

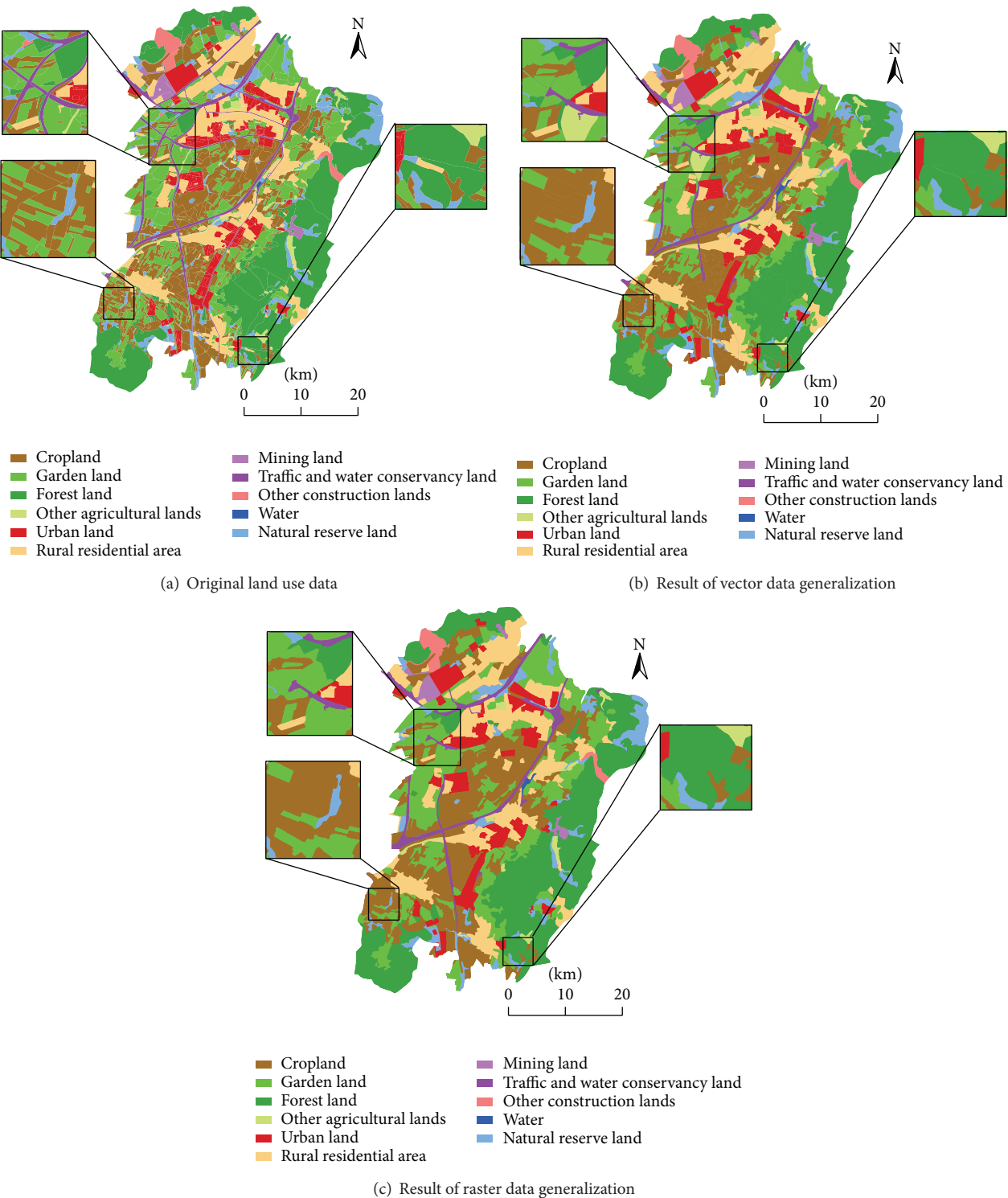(c) Result of raster data generalization

FIGURE 5: Contrast between result of generalization and original data of land use.

so as to complete the patch generalization of the raster data. The generalization results are shown in Figure 5.

The statistical results of the quantity and area of the patches before and after the generalization are shown in Table 1. Compared with the total patch number, the result of the raster data generalization was better, and the patch number was reduced from the original 1,805 to 256, while, comparing with the total area change of the patch, the effect of the vector data generalization was better, and the total areas before and after the generalization were completely consistent, mainly because the types of land use are considered in the processes of the patch amalgamation and aggregation

TABLE 1: Patch number, area change and semantic computation before and after generalization.

| | Patch number | | | Patch area | | | Semantic consistency | | Semantic completeness | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Original data | Vector data | Raster data | Original data (hm²) | Vector data (hm²) | Raster data (hm²) | Vector data | Raster data | Vector data | Raster data |
| Cropland | 545 | 258 | 44 | 703.797 | 743.035 | 718.999 | 0.944 | 0.978 | 0.473 | 0.081 |
| Garden land | 434 | 176 | 62 | 629.161 | 629.257 | 644.396 | 1 | 0.976 | 0.406 | 0.143 |
| Forest land | 154 | 56 | 21 | 824.501 | 823.27 | 826.414 | 0.999 | 0.998 | 0.364 | 0.136 |
| Other agricultural lands | 41 | 11 | 8 | 41.711 | 46.554 | 28.382 | 0.884 | 0.68 | 0.268 | 0.195 |
| Urban land | 242 | 94 | 28 | 234.213 | 241.244 | 235.752 | 0.97 | 0.993 | 0.388 | 0.116 |
| Rural residential land | 171 | 86 | 41 | 430.688 | 444.607 | 462.983 | 0.968 | 0.925 | 0.503 | 0.24 |
| Mining land | 9 | 6 | 5 | 31.316 | 31.335 | 32.728 | 0.999 | 0.955 | 0.667 | 0.556 |
| Traffic and water conservancy land | 37 | 17 | 8 | 134.213 | 91.865 | 107.572 | 0.684 | 0.801 | 0.459 | 0.216 |
| Other construction lands | 6 | 3 | 3 | 40.743 | 40.415 | 40.181 | 0.992 | 0.986 | 0.5 | 0.5 |
| Water | 1 | 1 | 1 | 1.752 | 3.33 | 3.386 | 0.099 | 0.067 | 1 | 1 |
| Natural reserve land | 165 | 60 | 35 | 195.697 | 172.88 | 165.305 | 0.883 | 0.845 | 0.364 | 0.212 |
| Total | 1805 | 768 | 256 | 3267.792 | 3267.792 | 3266.096 | 1 | 0.999 | 0.425 | 0.142 |

Note: all vector data and raster data are the statistical result of generalization of land use patch.

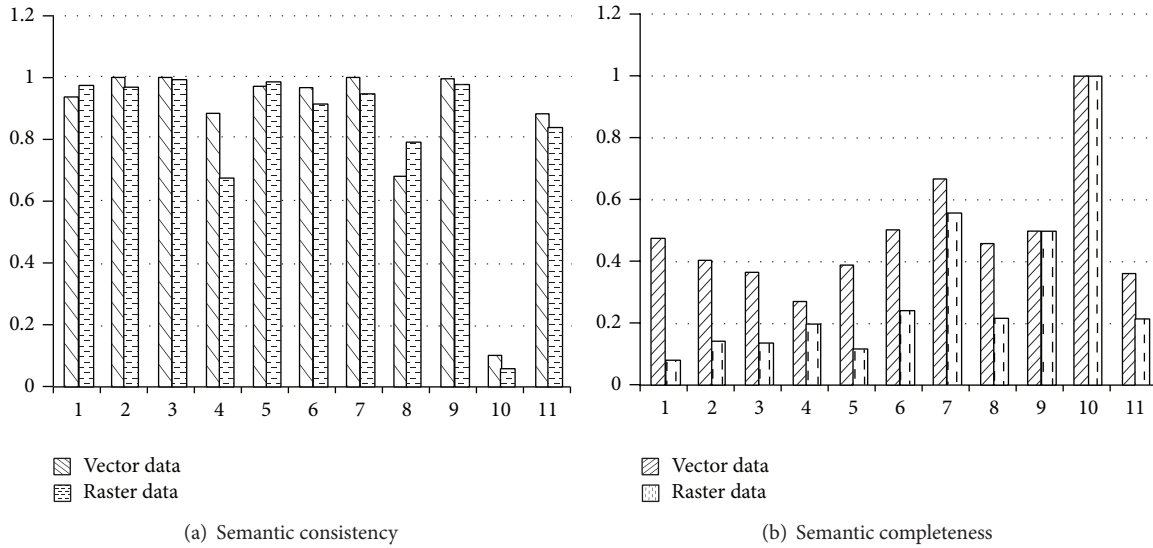(a) Semantic consistency

(b) Semantic completeness

FIGURE 6: Histogram of semantic consistency and completeness of generalization of land use patch. Note: 1: cropland; 2: garden land; 3: forest land; 4: other agricultural land; 5: urban land; 6: rural residential land; 7: mining land; 8: traffic and water conservancy land; 9: other construction lands; 10: water; 11: natural reserve land.

of the vector data, making the area of each land type maintain balance to the maximum limit before and after the generalization. Compared with the area change of each type of land use, in the result of the vector data generalization, the type of land use with maximum area reduction is the traffic and water conservancy land, and, compared with the original data, the area is totally reduced by 42.35 hm$^2$, mainly because such kinds of land contain long and narrow regions whose width is less than the epigraph distance such as highway and railway, and the subdivision processing is conducted for these long and narrow regions to merge them into other features; in the result of the raster data generalization, the area reduction of the natural reserve is maximum, up to 30.39 hm$^2$, mainly because the shapes of the land type patches are relatively complex, while the mode filtering algorithm merely has the action of boundary smoothing, and finally the smoothing processing is conducted for the projecting parts of the patch shape, reducing its area.

As can be seen from Table 1, on the level of the whole map, compared with the original data, the generalization results of the vector and raster data basically adhere to similar semantic consistency, while because the vector data generalization is relatively more balanced on the holding, its semantic consistency is higher, up to the value of 1. Compared with semantic completeness, the semantic completeness of the generalization results of the vector and raster data are not high, and especially of note is the fact that the semantic completeness of the raster data is a mere 0.142, mainly because the combination of the closing operation of the mathematical morphology and mode filtering algorithm changes the pixel values and achieves patch aggregation, significantly reducing the patch number, with a serious deletion of comprehensive categories.

Combining Table 1 and Figure 6, as can be seen from Figure 6, in the generalization results of the vector and raster

data, except for the waters, the semantic consistency of the generalization results of other land types is above 0.6, mainly because there is only one water patch whose area is small but meets the epigraph requirements in the research area; there are more secondary features around it, and many small patches are absorbed in the generalization process, resulting in too large area change rates of the water patch and reducing its semantic consistency. As can be significantly seen from Figure 6(b), because there is only one water patch before and after generalization, the semantic completeness of the two generalization results reach the maximum value 1. Comparing the two generalization results, for the land types with the exception of water, the semantic completeness value of the vector data is higher than that of the raster data, fundamentally explaining the reason why the land type deletion of the raster data generalization result is relatively serious with respect to the vector data on the whole map scale, namely, the semantic completeness is lower.

## 6. Conclusions and Discussions

In the research, we used the buffer overlap, Delaunay triangle subdivision, mathematical morphology, cellular automata, and other theoretical methods to achieve the patch generalization of the vector data and raster data and compared the advantages and disadvantages of different spatial data in the patch generalization process, so as to provide the basis for the selection of generalization data for the land use patch.

Analyzing from the perspective of the data structure, the object oriented vector data is more suitable for quantitative and qualitative researches, while the highly structured raster data have more advantages in aspects such as rapid modeling and extracting neighborhood. The research shows that, after the generalization of the raster data, the degree of reduction for the total number of the patches is greater, and the

aggregation effect of the small patches is obvious, ensuring that the spatial layout of the patches is more standardized. However, in the aspect of maintaining the total area balance of the patch, the generalization of the vector data has more advantages.

Analyzing from the scale of land type, the semantic consistency of each type of land use of the two kinds of data before and after the generalization differs a little, and the semantic consistency of most land types can reach above 0.6, indicating that the patch area of each land type maintains balance to a certain extent before and after the generalization. Comparing the semantic completeness of each land type, the raster data generalization result is relatively low with respect to the vector data, and the degree of semantic omission is large. The number of the water patches does not change before and after the generalization, and the semantic completeness of the water patches reaches the value of 1 in the generalization results of the two kinds of data.

Comparing on the map level, the semantic consistency is high for the generalization results of the two kinds of data, and especially the degree of the semantic consistency of the vector data generalization results with the original data reaching the value of 1, while the semantic completeness of the generalization results of the two kinds of data is relatively low, and the semantic completeness of the raster data generalization result is only 0.142, indicating that the two kinds of data differ little on the total area balance before and after generalization, while on the deletion of land type, the raster data is more serious.

In summary, in the process of cartographic generalization, it is better to select the vector data to conduct the generalization for the items whose area requirements and semantic requirements before and after generalization are strict; on the contrary, the raster data can be selected for generalization that more emphasizes the overall layout specification of the map, while effective smoothing, aggregation, and other operations are achieved for the irregular small patches.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

## References

[1] Q. Wenqing, "The Theory and approaches of geographic—feature oriented generalization in GIS environment," *Acta Geographica Sinica*, vol. 53, no. 4, pp. 303–313, 1998.

[2] L. Qiao, L. Chun, X. Changyuan et al., "Cartographic generalization of digital land use current situation map," *Bulletin of Surveying and Mapping*, vol. 3, pp. 30–32, 2000.

[3] Z. Lijun, *Study on the Methods of Land Use Master Planning Map*, East China Normal University, Shanghai, China, 2011.

[4] Z. Wei, *Study on the Methods of Land Use Map Genefalization Based on ArcGIS*, Zhejiang University, Zhejiang, China, 2006.

[5] L. Yang, *An Approach for Quality Evaluation in Polygon Map Generalization*, Central South University, Hunan, China, 2009.

[6] Y. Zhihua, *The Study on Cartographic Generalization Based on Topological Relations*, China University of Geosciences, Beijing, China, 2010.

[7] Z. Guang, *The Study of Map Generalization Based on Multi-Agent Systerm*, Zhejiang University, Zhejiang, China, 2011.

[8] S. Zhikuan and T. Guoan, "Quality assessment of land use data generalization," *Geographical Research*, vol. 30, no. 2, pp. 233–242, 2011.

[9] N. Zhanguo, *The Mapping Generalization of the Digital Land Distribution Product Map of GIS*, Anhui University of Science and Technology, Anhui, China, 2007.

[10] M. S. Monmonier, "Raster-mode area generalization for land use and land cover maps," *Cartographica*, vol. 20, no. 4, pp. 65–91, 1983.

[11] B. Su and L. Zhilin, "An algebraic basis for digital generalization of area-patches based on morphological techniques," *The Cartographic Journal*, vol. 32, no. 2, pp. 148–153, 1995.

[12] B. Su, Z. Li, and G. Lodwick, "Morphological transformation for the elimination of area features in digital map generalization," *Cartography*, vol. 26, no. 2, pp. 23–30, 1997.

[13] W. Huilian, W. Fang, W. Bao-Shan et al., "The algorithm for aggregation of polygon applied to automatic map generalization," *Engineering of Surveying and Mapping*, vol. 14, no. 3, pp. 15–18, 2005.

[14] Z. Su, *Research on the method of land-use cartographic generalization based on the combination of mathematical morphology and cellular automata [M.S. thesis]*, Northwestern University, Shanxi, China, 2011.

[15] W. Jie and G. Qingsheng, "An improved algorithm for combination of land-use data," *Geomatics and Information Science of Wuhan University*, vol. 37, no. 9, pp. 1116–1119, 2012.

[16] Y. Jun, K. Fanqiang, X. Jianchao, Q. Ge, X. Li, and P. Xie, "Land use patch generalization based on semantic priority," *Abstract and Applied Analysis*, vol. 2013, Article ID 151520, 8 pages, 2013.

[17] A. Tinghua and G. Renzhong, "Extracting center-lines and building street network based on constrained Delaunay Triangulation," *Acta Geodaetica et Cartographica Sinica*, vol. 29, no. 4, pp. 348–354, 2000.

[18] A. Tinghua, G. Renzhong, and C. Xiaoneidong, "Simplification and aggregation of polygon object supported by Delaunay Triangulation structure," *Journal of Image and Graphics*, vol. 6, no. 7, pp. 703–709, 2001.

[19] A. Tinghua and L. Yaolin, "Aggregation and amalgamation in land-use data generalization," *Geomatics and Information Science of Wuhan University*, vol. 27, no. 5, pp. 486–492, 2002.

[20] K. Xinli and B. Fuling, "Review on researches of geographical cellular automata," *Journal of Xianning University*, vol. 29, no. 3, pp. 103–106, 2009.