Sensors and Materials, Vol. 36, No. 7 (2024) 3047–3058
MYU Tokyo

S & M 3717

A Simplification Method for Continuous Tessellation Area Data Based on Discretization Idea

Pengda Wu,[1,2] Yao Cheng,[2*] Yanyan Zeng,[3] and Wentao Cao[4]

[1]School of Information Engineering, China University of Geosciences,
No. 29, Xueyuan Road, Haidian District, Beijing 100083, China
[2]Chinese Academy of Surveying and Mapping,
No. 28, Lianhuachi West Road, Haidian District, Beijing 100036, China
[3]Beiing Institute of Surveying and Mapping, No. 15, Yangfangdian Road, Haidian District, Beijing 100038, China
[4]Wuhan Geomatics Institute, No. 209, Wansongyuan Road, Wuhan 430022, China

(Received April 25, 2024; accepted July 9, 2024)

Keywords: simplification method, tessellation area data, area discretization, area reconstruction

The simplification operation has always been a challenging research problem in the field of cartography generalization and expression. Continuous tessellation area data of land cover, whose source data come from high-resolution remote sensing satellite sensors, are an important type of remote sensing application data to express the spatial distribution information of natural and artificial objects on Earth's surface. In this paper, a novel and practical simplification method for continuous tessellation area data is proposed. The first step is area discretization, in which a general topological structure containing semantic features is established, and land patches are divided into multiple arcs on the basis of this topological structure. The second step is arc simplification, in which arcs are classified into artificial and natural types according to their semantic and morphological features, and corresponding simplification algorithms are assigned. The final step is area reconstruction, based on the semantic and topological information, and simplified arcs are reconstructed to patch polygons. The geography and national conditions census data of a city in Guizhou Province, China, are used to verify the rationality and effectiveness of the method. Experimental results showed that the method proposed in this article is suitable for the simplification processing of artificial and natural land features. The simplification results can maintain the overall shape of the patches and reduce the complexity of the arcs effectively, and the semantic information of each patch after simplification is also correct.

1. Introduction

In recent years, a series of major projects have been carried out in China, such as the National Geographic Survey and the Third Land and Resources Survey, forming a massive amount of vector land cover data to express the spatial distribution of natural and artificial objects on Earth's surface. These data are mainly the interpretation results of high-resolution remote

*Corresponding author: e-mail: chengyao@casm.ac.cn
https://doi.org/10.18494/SAM5090

ISSN 0914-4935
https://myukk.org/

sensing images, which are derived from remote sensing satellite sensors. They are also called continuous tessellation area data due to their seamless and non-overlapping spatial characteristics. Continuous tessellation area data strongly support the construction of national ecological civilization through numerical analysis at the micro level and multi-scale expression at the macro level. The core of multi-scale expression is the automatic generalization of continuous tessellation area data, that is, from large-scale data (patches) to various small-scale data to comprehensively reflect regional land cover conditions at multiple levels.[1–4] Simplification is one of the key steps in the automatic generalization of continuous tessellation area data. The basic idea is to simplify the internal details of polygons while maintaining the original area size and shape structure as much as possible.[5–9] Considering the large data volume, diverse shapes, and complex structures, patch simplification has become a difficult issue in land cover data generalization.

Scholars have designed a large number of algorithms and models to simplify polygons. Traditional methods mainly use morphological methods to simplify polygons based on grid structures. These methods are simple, efficient, and easy to implement, but rich semantic information contained in patch data cannot be used. In recent years, polygon simplification supported by vector data structures has gradually become a research hotspot. Because the polygons on the map are represented as closed areas surrounded by contour lines, the basis for their simplification is transformed to the line element simplification algorithm. Early line element simplification algorithms, such as the classic Douglas–Peucker (D–P) algorithm,[10] Visvalingam–Whyatt algorithm,[11] and L–O algorithm,[12] were mostly targeted at a single line element. However, patch data have diverse classes and complex shapes, making it difficult for a single line element simplification algorithm to meet different actual needs. In addition, existing algorithms mostly process discrete and independent features, but are not applicable to situations where land cover data are continuous and connected to each other.

To this end, we propose a new simplification method for continuous tessellation area data, which is mainly based on a discretization idea. The main processes are as follows: (1) area discretization – a general point-arc-polygon topological structure is established, and land cover patches (polygons) are divided into multiple arcs on the basis of the general topological structure; (2) arc simplification – arcs are classified into different types according to their semantic and morphological features, and appropriate simplification algorithms are assigned to different arcs; (3) area reconstruction – on the basis of patch semantic and topological information, simplified arcs are reconstructed to patch polygons.

2. Methods

2.1 Area discretization

Continuous tessellation area data are seamless expressions of the land surface in the computer. There are many shared common edges in the area (patch) data. If each patch polygon is used as a unit for simplification, the simplified results of adjacent patch shared common edges will be inconsistent, which is unacceptable for human cartographers.[13] To this end, we discretized the areas into arcs on the basis of the topological structure and used each arc as a simplification unit to perform the complete simplification of continuous tessellation area data.

The topological data structure is a way of organizing spatial data based on the principles of topological geometry. It expresses the adjacency or association relationships between geographic entities by recording information such as the nodes connected to the arcs, the left- and right-side polygon markers of the arcs, and the arcs that make up the polygons. It can effectively decompose and organize the entire space according to the spatial topological relationship between various geographical entities, and it has the stability in map projection. It is also conducive to a direct or indirect query of the composition, proximity, and other information among geographical entities.(14,15)

On the basis of the classic topological structure, we propose a general topological structure containing geometric, topological, and semantic information, in which the topological, geometric, and semantic information respectively correspond to the spatial, position, and attribute information of the patch data. The general topological structure can effectively decompose and organize the entire space according to the composition relationship of each node, arc, and polygon sub-object. This will provide basic support for discretizing and reconstructing the area data. The general topological structure is shown as follows.

(1) Node: records the relationship between nodes and arcs

TopoNode //topological node

{

Point //position coordinates of the point in two-dimensional space (geometric information)

ArcNumber //the number of arcs passing through the node (topological information)

ArcList //the list of all arcs passing through this node (topological information)

LayerID //the unique identifier of the layer where the node is located (semantic information)

FeatureID //the unique identifier of the entity to which the node belongs (semantic information)

}

(2) Arc: it is referred to the directed arcs. Each arc has only one first node and one last node. Apart from the nodes, there are no other common points between arcs. The intersection point of two arcs must be a node. In the arc data structure, the relationship between arcs, nodes, and polygons and their semantic information are recorded.

TopoArc //topological arc

{

sNode //the first node of the arc (topological information)

eNode //the end node of the arc (topological information)

LineString //an ordered set of points on an arc (geometric information)

LPolygon //the left neighbor polygon of the arc (topological information)

RPolygon //the right neighbor polygon of the arc (topological information)

LayerID //the unique identifier of the layer where the arc is located (semantic information)

FeatureID //the unique identifier of the entity to which the arc belongs (semantic information)
CommonID //shared edge identifier (semantic information)
}

(3)Polygon: it is used to represent the area in the data. Polygons are divided into two categories: network polygons (NetPolygon) and island polygons (IslandPolygon). A network polygon is an irreducible polygon surrounded by one or more arcs, which corresponds to the area one-to-one and forms the outer boundary of the area. Island polygons are separated from each other and have no common arcs between them.

NetPolygon //network polygon
{
ArcList //the arcs that make up the polygon (topological information)
ArcNumber //the number of arcs that make up the polygon (topological information)
IslandList //contained island polygons (topological information)
IslandNumber //the number of island polygons contained (topological information)
}
Island Polygon //island polygon
{
ArcList //the arcs that make up the polygon (topological information)
ArcNumber //the number of arcs that make up the polygon (topological information)
OuterPolygon //the outer polygon containing the island (topological information)
}

2.2 Arc simplification

Natural and artificial features are the main components of land cover data. Specifically, the garden land, woodland, and grassland belong to natural features, and buildings, roads, and excavations belong to artificial features The shapes of artificial features are more regular and neater, with obvious traces of artificial transformation. For example, the boundaries of general buildings are usually curved at right angles, and general roads usually extend in parallel postures. However, the distribution of artificial structures on the map is more scattered and fragmented. The shapes of natural features are smoother; for example, farmland and gardens are naturally divided by mountains and rivers, and the distribution range and area of natural features on the map are also larger and more diverse. For this reason, when simplifying the boundaries of the patches, we divide them into two categories: artificial and natural arcs. If the left and right polygons of an arc contain artificial features, the arc is marked as an artificial arc; otherwise, it is marked as a natural arc.

2.2.1 Artificial arc simplification

There are two key points in the simplification of artificial arcs. One is to fully consider the shape characteristics of the arcs and keep the rectangular characteristics without significant changes before and after simplification; the second is to evaluate the variabilities of the concave and convex arcs based on the changed scale and fill in concave parts or remove convex parts to perform the simplification process.[16] The shape structures of artificial arcs are classified and defined according to different distribution situations of the concave and convex.

(1) Classification and definition of the shape structure of processing units

Definition 1: Circular arc structure. If the angle between each node on the arc and its front and rear nodes is obtuse and the angles are approximately equal, then the arc is circular. According to the length of the arc and the size of the bending angle, it can be further subdivided into three structures: circular arch, sector, and full circle, as shown in Fig. 1.

Definition 2: Ladder-like structure. If the angle between each node on the arc and its front and rear nodes is a right angle or approximately a right angle, then the arc has a ladder-like structure, as shown in Fig. 2.

Definition 3: U-shaped structure. If the first and last points of four adjacent points on an arc are on different sides of the line segment formed by the middle two points, and the first and last edges of the three edges formed by the combination of the four points are on the different side of the middle edge, the structure enclosed by the four points is called the U-shaped structure. According to the different lengths of the edges, the U-shaped structure can be further subdivided into a flush U-shaped structure and an uneven U-shaped structure, as shown in Figs. 3(a) and 4(a), respectively.

Definition 4: Sharp-angle structure. If the angle between a node on an arc and its front and rear nodes is less than 10°, the arc is regarded as a sharp-angle structure, as shown in Fig. 5.

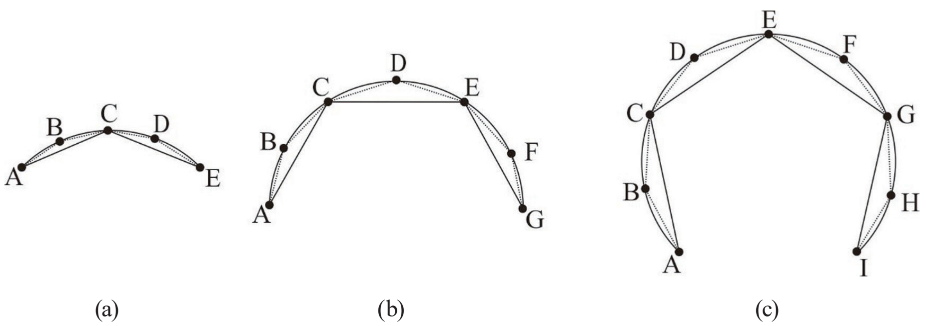


Fig. 1. Schematic of circular arc structure and its simplification result. (a) Circular arch, (b) sector, and (c) full circle.

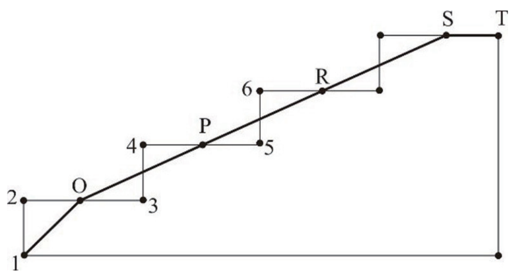


Fig. 2. Schematic of ladder-like structure and its simplification result.

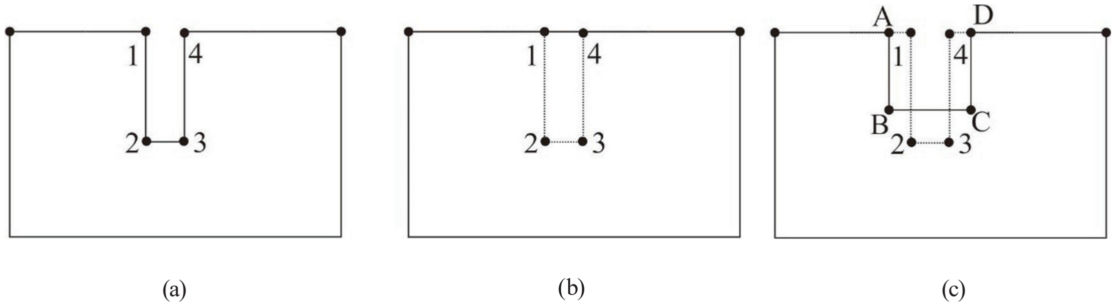


Fig. 3. Schematic of (a) flush U-shaped structure and its simplification results (b) 1 and (c) 2.

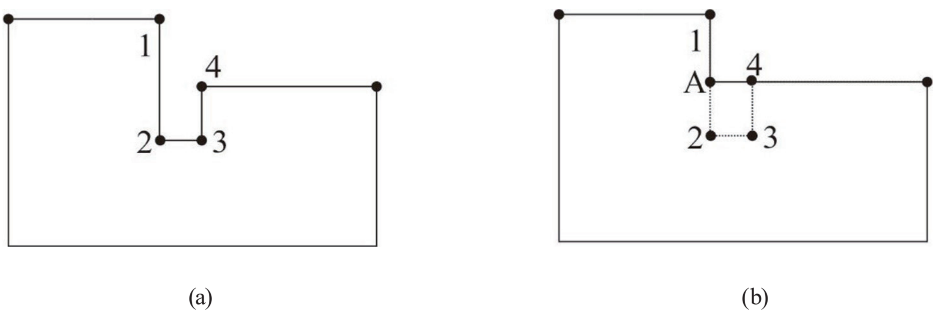


Fig. 4. Schematic diagram of (a) uneven U-shaped structure and (b) its simplification result.

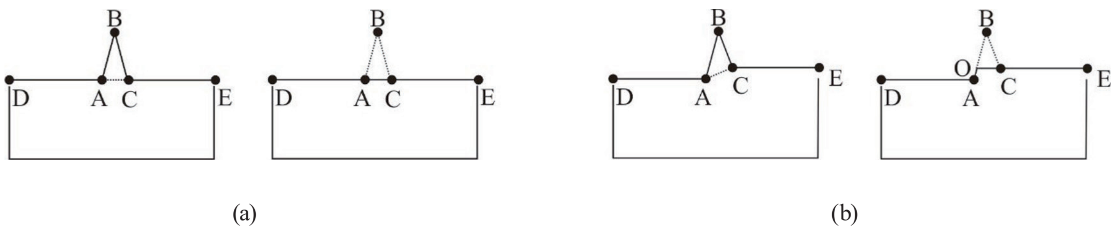


Fig. 5. Schematic of sharp-angle structure and its simplification result. Situations (a) 1 and (b) 2.

(2) Simplification methods for different shape structures

The basic principle of the simplification of artificial features is to reduce redundant details of arcs while maintaining the morphological characteristics and area size of the features so that the artificial features can be expressed more concisely and effectively on the target scale map. The simplification of artificial features should maintain the consistency of the visual perception of the graphics by the human eye. Therefore, the most basic criterion for simplification is that the structure whose length is smaller than the minimum visual length threshold (LT) of the human eye on the map (0.3 or 0.4 mm[17]) should be simplified.

- Simplification of circular arc structure.

The basic idea for the simplification of a circular arc structure is to sample the points on the curve and turn the curve into a polyline. Usually, the minimum visual length is taken as the simplification threshold and the D-P algorithm is used as the simplification method. As shown in Fig. 1, the simplification result of the circular arch AE is arc ACE [Fig. 1(a)], the simplification result of sector AG is arc ACEG [Fig. 1(b)], and the simplification result of circle AI is arc ACEEGI [Fig. 1(c)].

- Simplification of ladder-like structure.

The ladder-like structure can be regarded as the repeated accumulation of the inverted Z-shaped structure. When the scale becomes smaller, the structure can be abstracted into a jitter-free line. The simplification algorithm uses the first node of the arc as the initial node and takes three adjacent nodes as a processing unit. The midpoint of the longer one of the two edges composed of three nodes is extracted. The same processing is repeated up to the end node. All midpoints are connected and taken as the final simplification result. As shown in Fig. 2, the three points 1, 2, and 3 are the initial processing units. The length of the two sides composed of the three points is $S_{23} > S_{12}$; then, the midpoint O of the edge 23 is taken as the simplification node. In the same way, the midpoint P, R, S is extracted, and the final simplification result arc 1OPRST is obtained, as shown by the thick line in the figure.

- Simplification of U-shaped structure.

For the flush U-shaped structure: If $S_{12} \times S_{23} < LT^2$ or $S_{23} \times S_{34} < LT^2$, directly delete points 2 and 3 and connect points 1 and 4, as shown in Fig. 3(b); otherwise, exaggerate the structure and move points 1, 2, 3, and 4 to A, B, C, and D. It is required to make $S_{BC} =$ LT and $S_{AB} \times S_{BC} = S_{12} \times S_{23}$, as shown in Fig. 3(c).

For the uneven U-shaped structure, if $S_{23} \times S_{34} < LT^2$, directly delete points 2 and 3, and extend point 4 to point A, as shown in Figs. 4(a) and 4(b).

- Simplification of sharp-angle structure.

If the start and end nodes of the sharp corners A and C are approximately on a straight line and $S_{ABC} < LT^2$, then the area of the sharp corner is very small, and the two points A and C are directly fitted to one point, as shown in Fig. 5(a). If the start and end nodes of the sharp corners A and C are not on a straight line and $S_{ABC} < LT^2$, then delete point B and move point C to point O, as shown in Fig. 5(b).

2.2.2 Natural arc simplification

The simplification of natural arcs is different from the processing of artificial arcs. It emphasizes the optimization of arc details, eliminates small bends, and makes the arcs smoother.[18] During the simplification, the most likely problem is that intersection anomalies will occur where the arcs are dense or the arcs have large curvatures. For this reason, we used a line global simplification algorithm[19] that considers spatial relationships as constraints. The specific steps are as follows.

Step 1: On the basis of the general topological structure, determine and record the topological relationship of each arc before simplification. There are usually only two types of topological relationship in patch data: connected and separated.

Step 2: Set the simplified distance threshold (DT, 03 or 0.4 mm) and use the D–P algorithm to obtain the set of local extreme points whose distance is greater than DT.

Step 3: Calculate the size R of the minimum visible target in the L–O algorithm on the basis of the target and original scales.

Step 4: Using the local extreme point as the circle point, the R value as the detection step, and the first node of the arc as the initial node, apply the L-O algorithm to simplify the arc and end at the end node of the arc.

Step 5: Update the patch data topology structure and determine whether the topological relationship of the arcs observed before and after simplification has changed. If it has changed, optimization will be performed, such as shifting or deletion; if there has been no change, the arc simplification ended.

2.3 Area reconstruction

After the simplification operation, the arcs should be reorganized to area data to ensure the consistency of the data structure before and after simplification. Area reconstruction is the process of constructing a polygon through the simplified arcs based on the general topological structure. A new reconstruction area must comply with the definition of a polygon, that is, three or more line segments connect in sequence forming a closed plane geometric figure, and the semantic information of the graphic identification must be unique. In this article, we propose a reconstruction method under semantic and spatial relationship constraints. The specific process is as follows.

Step 1: On the basis of the associated nodes and polygon topology information contained in the topological structure, the discrete arcs are organized as the topological polygon NetPolygon.

Step 2: The topological polygons are traversed and the semantic and topological information for arcs is updated with inconsistent topological information before and after simplification.

Step 3: On the basis of the semantic information in the topological structure, arcs with the same semantic information are reorganized into geographical entities.

3. Experiments and Analysis

The simplification method for continuous tessellation area data is embedded to the WJ-III map workstation, which was developed by the Chinese Academy of Surveying and Mapping. The geography and national conditions census data of a city in Guizhou Province, China, are used to verify the rationality and effectiveness of the method. The original data scale is 1:10,000, with 5214 patches. The land feature types are mainly natural features such as cultivated land, woodland, and grassland, artificial features are scattered among them, and the simplification target scale is 1:100,000. Some patch data obtained in the experimental area before and after simplification are shown in Fig. 6.

The experimental process is as follows.

Step 1: The general topological structure containing semantic information for the original data is

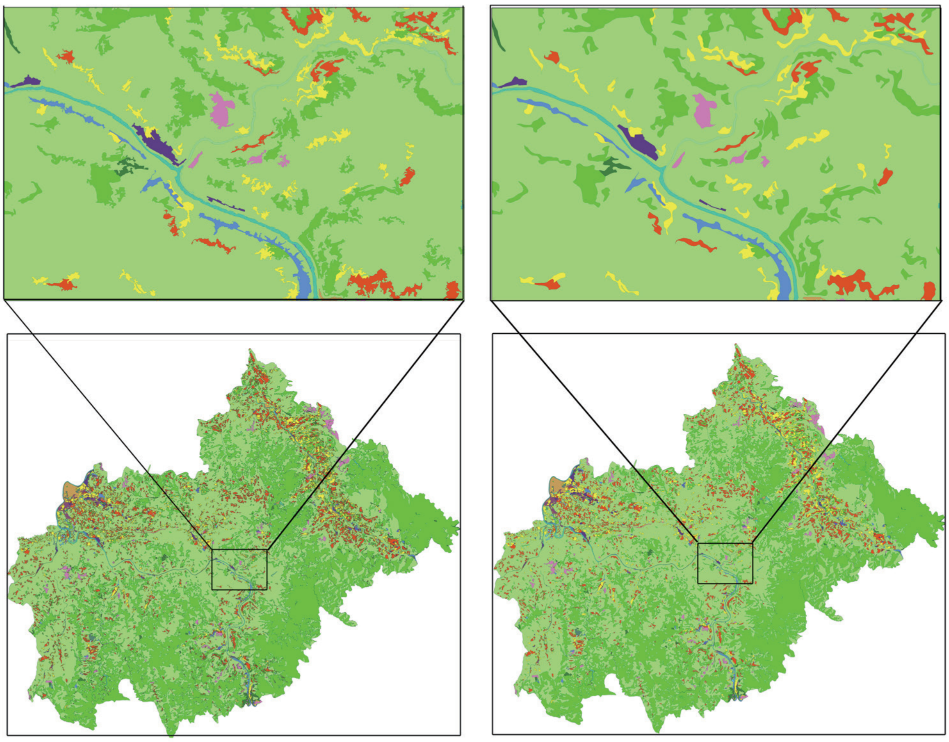

(a) (b)

Fig. 6. (Color online) Comparison of some patch data obtained in the experimental area (a) before and (b) after simplification.

established, and topological preprocessing is performed to eliminate topological errors such as redundant and false nodes in arcs.

Step 2: According to the semantic information contained in the arcs, the arcs are divided into two types of pattern: artificial and natural.

Step 3: The natural arcs are simplified using a line global simplification algorithm that considers spatial relationship constraints. The simplification distance threshold DT takes the minimum visible length threshold of 0.4 mm (distance on the map). The size R of the minimum visible target in the L-O algorithm is 36 m (actual distance).

Step 4: The artificial arcs are simplified in sequence in the order of circular arc structure, ladder-like structure, U-shaped structure, and sharp-angle structure. The simplification algorithm is performed iteratively until there is no arc that meets the simplification conditions.

Step 5: After all the arcs in the experimental area are processed, a topological polygon NetPolygon is constructed for the discrete arcs in the experimental area.

Step 6: On the basis of the topology and semantic information in NetPolygon, the geographic entities are reconstructed.

As illustrated in Fig. 6, the redundant details of the boundary of each patch data are reduced after simplification, while the morphological characteristics and area size of the patches remain.

Each land patch is expressed more concisely and clearly at the 1:100000 scale. At the same time, the topological and semantic information of the patch data before and after simplification are consistent, and there are no gaps and overlaps.

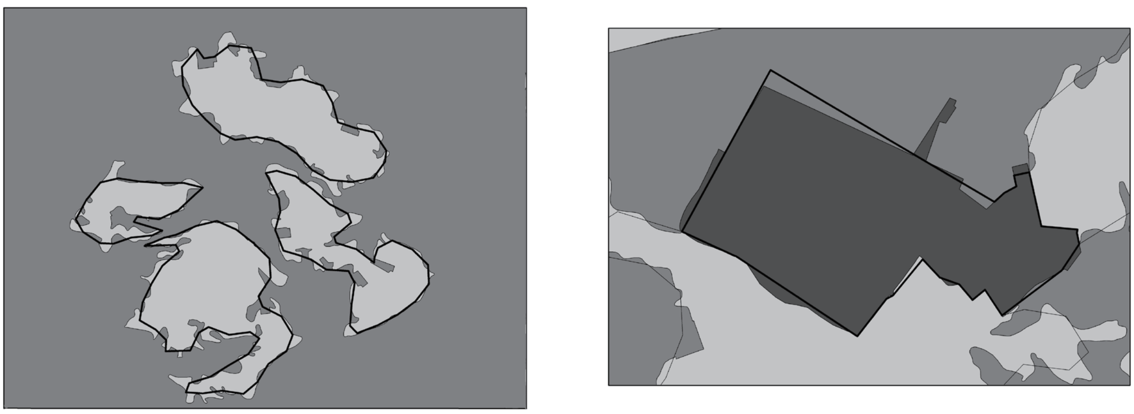
To further verify the rationality of this method, certain natural and artificial arcs obtained in the experimental area before and after simplification are zoomed-in, as shown in Fig. 7.

As illustrated in Fig. 7, for the simplification result of natural woodland arcs, the method in this article not only maintains the overall shape characteristics of the curve, but also makes the curve smoother; for the simplification result of artificial building arcs, the method in this article eliminates most of the concave and convex parts with small areas, and effectively retains and enriches the rectangular characteristics at the corners.

4. Conclusions

The simplification operation is a difficult problem for data generalization. Continuous tessellation area data of land cover, whose source data come from high-resolution remote sensing satellite sensors, are important remote sensing application data for expressing the spatial distribution information of natural and artificial objects on Earth's surface. In this paper, a novel and practical simplification method for continuous tessellation area data based on a discretization idea is proposed, which mainly includes three steps: area discretization, arc simplification, and area reconstruction. Compared with existing methods, the proposed method can effectively avoid the simplification inconsistency situation for the shared common edges of land use patches, and the simplification results can maintain the overall shape of the patches and reduce the complexity of the arcs. After verification with actual data, the following conclusions are drawn:

(1) The method proposed in this article is suitable for the simplification processing of artificial



(a) (b)

Fig. 7. Comparison of natural and artificial arcs obtained before and after simplification. The thick line represents the simplified result. (a) Natural arcs of woodland and (b) artificial arc of building.

and natural land features.

(2) For the simplification of natural features, the method in this article not only maintains the overall shape characteristics of the area, but also makes the curve smoother.

(3) For the simplification of artificial features, the method in this article can eliminate redundant details of the boundary and effectively retain the rectangular characteristics.

(4) There are no topological errors such as self-intersection, overlap, and capping in the simplification results, and the semantic information of each patch is preserved.

Further research will focus on a more refined shape feature recognition of the land cover patches and use more real data to test the appropriate parameter values required in this method. At the same time, we will further consider the requirement that the area ratio of each land cover class before and after the generalization does not change significantly to optimize the proposed simplification method.

Acknowledgments

This work was supported by the National New Fundamental Surveying and Mapping Construction Beijing Pilot Project and the National New Fundamental Surveying and Mapping Construction Wuhan Pilot Project.

References

1 J. Haunert and A. Wolff: Int. J. Geogr. Inf. Sci. **12** (2010) 1871. https://doi.org/10.1080/13658810903401008
2 T. Cheng and Z. Li: Cartogr. **2** (2006) 135. https://doi.org/10.3138/0172-6733-227U-8155
3 C. Li, Y. Yin, P. Wu, and W. Wu: Cartogr. Geogr. Inf. Sci. **3** (2021) 210. https://doi.org/10.1080/15230406.2020.1863862
4 Q. Zhu, Y. Lei, X. Sun, Q. Guan, Y. Zhong, L. Zhang, and D. Li: Remote Sens. Environ. **272** (2022) 112916. https://doi.org/10.1016/j.rse.2022.112916.
5 J. Du, F. Wu, J. Yin, C. Liu, and X. Gong: Cartogr. Geogr. Inf. Sci. **4** (2022) 313. https://doi.org/10.1080/15230406.2021.2013944
6 W. Park and K. Yu: Pattern Recognit. Lett. **9** (2011) 1267. https://doi.org/10.1016/j.patrec.2011.03.013
7 P. Duan, H. Qian, H. He, L. Xie, and D. Luo: Geomatics Inf. Sci. Whu. **5** (2020) 744. https://doi.org/10.13203/j.whugis20180434
8 B. J. Kronenfeld, L. V. Stanislawski, B. P. Buttenfield, and T. Brockmeyer: Int. J. Cartogr. **1** (2020) 22. https://doi.org/10.1080/23729333.2019.1631535
9 J. Du, F. Wu, R. Xing, X. Gong, and L. Yu: Geocarto Int. **14** (2022) 4158. https://doi.org/10.1080/10106049.2021.1878288
10 D. Douglas and T. Peucker: Can. Cartogr. **2** (1973) 112. https://doi.org/10.3138/FM57-6770-U75U-7727
11 M. Visvalingam and J. Whyatt: Cartogr. J. **1** (1993) 46. https://doi.org/10.1179/000870493786962263.
12 Z. Li and S. Openshaw: Int. J. Geogr. Inf. Sci. **5** (1992) 373. https://doi.org/10.1080/02693799208901921
13 T. Ai and H. Wu: Geomatics & Inf. Sci. of Whu. **5** (2000) 426. https://doi.org/10.13203/j.whugis2000.05.010
14 F. Luo, T. Ai, and H. Wang: Geomatics Inf. Sci. Whu. **6** (2004) 558. https://doi.org/10.13203/j.whugis2004.06.023
15 O. Van and C. Lemmen: Comput. Environ. Urban Syst. **5** (2001) 509. https://doi.org/10.1016/S0198-9715(00)00052-1.
16 R. Guo and T. Ai: Geomatics Inf. Sci. Whu. **1** (2000) 25. https://doi.org/10.13203/j.whugis2000.01.005.
17 J. Muller: Cartogr. J. **1** (1987) 27. https://doi.org/10.1179/caj.1987.24.1.27
18 F. Wu and K. Zhu: Geomatics Inf. Sci. Whu. **6** (2008) 600. https://doi.org/10.13203/j.whugis2008.06.023.
19 C. Li, P. Guo, Y. Yin, P. Wu, and T. Gu: Acta Geo. et Cartogr. Sinica. **4** (2017) 498. https://doi.org/10.11947/j.AGCS.2017.20160546.

About the Authors

Pengda Wu received his B.S. degree from Wuhan University, Wuhan, in 2013, and his M.S. degree from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences (CAS), in 2016. Since 2016, he has been an assistant researcher at the Chinese Academy of Surveying and Mapping. His research interests include spatial cognition, multiscale expression, and spatiotemporal knowledge graphs. (wupd@casm.ac.cn)

Yao Cheng received her M.S. degree from Wuhan University, China. Since 2014, she has been an assistant researcher at the Chinese Academy of Surveying and Mapping. Her research interests are in digital map generalization, spatial cognition, and multiscale expression.
(chengyao@casm.ac.cn)

Yanyan Zeng received her B.S. degree from China University of Petroleum Shandong in 2010 and her Ph.D. degree from the University of the Chinese Academy of Sciences, Beijing in 2015. Since 2015, she has been a senior engineer at the Beijing Institute of Surveying and Mapping. Her research interests are in GNSS data processing and new fundamental surveying and mapping. (zengyanyan1989@163.com)

Wentao Cao received his B.S. and M.S. degrees from Wuhan University, China, in 2011 and 2013, respectively. Since 2013, he has been an engineer at Wuhan Geomatics Institute. His research interests are in geographic information data processing, new fundamental surveying and mapping, and 3D real scene. (cao6008@126.com)