Article

Template Matching and Simplification Method for Building Features Based on Shape Cognition

Xiongfeng Yan [ID], **Tinghua Ai** * [ID] **and Xiang Zhang**

School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, China;
xiongfeng.yan@whu.edu.cn (X.Y.); xiang.zhang@whu.edu.cn (X.Z.)
* Correspondence: tinghuaai@whu.edu.cn; Tel.: +86-027-6877-8375

Abstract: This study proposes a template matching simplification method from the perspective of shape cognition based on the typical template characteristics of building distributions and representations. The method first formulates a series of templates to abstract the building shape by generalizing their polygons and analyzing their symbolic meanings, then conducts the simplification by searching and matching the most similar template that can be used later to replace the original building. On the premise of satisfying the individual geometric accuracy on a smaller scale, the proposed method can enhance the impression of well-known landmarks and reflect the pattern in mapping areas by the symbolic template. The turning function that describes shape by measuring the changes of the tangent-angle as a function of the arc-length is employed to obtain the similar distance between buildings and template polygons, and the least squares model is used to control the geometry matching of the candidate template. Experiments on real datasets are carried out to assess the usefulness of this method and compare it with two existing methods. The experiments suggest that our method can preserve the main structure of building shapes and geometric accuracy.

Keywords: map generalization; building simplification; template matching; shape cognition; turning function

1. Introduction

In topographic maps, building features play a significant role in multi-scale representations, and they have been widely used in various fields, such as urban planning, navigation, and landscape analyses. To improve legibility (i.e., readability) within the map, the outlines of building polygons should be represented more succinctly and concisely, and this problem needs to be addressed by simplification.

Building simplification aims to represent objects more briefly by generalizing their polygonal shapes while capturing their geometry, community structures, and distribution patterns. This process should fully consider the individual characteristics and the overall structure of a settlement. Perceived individually, the shape structures of buildings are controllable and artificial, similar to other man-made features, which are in accordance with the design and will of the architects [1]. Therefore, there is a substantial difference between natural features (e.g., soil parcels [2]) and their simplifications, such as enhancing the orthogonality of the outline [3]. Some of these structures are meant more to convey the symbolic meanings and culture characteristics behind map objects, than to convey an exact positioning [4]. Perceived as a group, the overall structure is closely related to the geographical context, is affected by the terrain, and is usually distributed in a dense, clustered pattern [5–9]. There are relatively stable and typical templates or pattern characteristics that can be effectively applied to building generalization [10,11]. For example, in a certain residential region, the buildings may generally look similar in shape with a strong pattern, including collinear and curvilinear patterns [12], H-patterns, Z-patterns, E-patterns, grid patterns, and stair-patterns [13].

Thus, from the perspective of spatial cognition, a set of shape templates can be built that can then be regarded as simplified versions of the original buildings if the two shapes match to some extent. The template is a high-level concept of shape structure description. On the premise of satisfying the individual geometric accuracy on a smaller scale, a symbolic template can enhance the impression of well-known landmarks and reflect the patterns in the mapping area. Based on this idea, this paper proposes to simplify building features first by shape cognition and then by template matching.

The remainder of this paper is structured as follows: Section 2 presents related work, surveys, the state-of-the-art with respect to building simplification and shape recognition. Section 3 describes the details of the simplification method using template matching and attempts to answer three questions: (1) what are the characteristics of templates and why can they simplify buildings? (2) how does one find the best templates for a building that needs to be simplified? and (3) how does one match it? Section 4 presents the experiments we carried out to test and assess the method and compare it with existing methods. Finally, Section 5 summarizes the papers and notes directions for future research.

2. Related Work

2.1. General Building Simplification

The objective of building simplification is to abstract the geometry with a simpler form and preserve, or even enhance, the characteristic shape, in which a set of rules and constraints should be considered [14]. For example, the removal of edges shorter than a specified length, or the closing off or widening of isolated small intrusions or extrusions while maintaining topological consistency. Considering these requirements, many algorithms have been developed over the decades. These studies can be roughly divided into three categories: geometric, morphological, and optimization algorithms.

The main idea of a geometric algorithm is to detect short edges or small bend structures, and then handle them by the operations of removing, enlarging, squaring, and so on. Buchin [15] presented an edge-move method for building simplification or schematization that operates the local structure iteratively. In terms of local structure recognition, Chen et al. [16] used the triangle of constrained Delaunay triangulation as a primitive shape to identify the concave structures by matching the typical triangle sequences. Xu et al. [17] took the continuous adjacent four points as a basic unit to define the bend structures and remove the short edges; this method is an extension and improvement of early line simplification local algorithms.

Another category of building simplification embodies morphological methods. Damen et al. [18] introduced the morphological operators (e.g., dilation and erosion) to aggregate and simplify buildings. Some drawbacks of this approach were improved by Meijers [19] using offset curves that were derived from the straight skeleton, a result through which orthogonal characteristics can be obtained. These morphological methods handle the object as a whole. In contrast, space decomposition provides an alternative strategy to local processing. Guo and Ai [20] proposed a hierarchical rectangle differential model according to the inner component structures of a building. The model separates building polygons as a set of rectangles and builds the relation between them. The simplification result with a special scale can be derived from the rectangle composite state. Kada [21] represented building polygons as 2D line segments based on half-space modeling and cell decomposition, through which simplification can be translated into a selection of cells problem. Raster-based methods have also been studied by some researchers [22].

Optimization techniques have been widely used in automatic cartographic generalization and data abstraction [23]. For the simplification of buildings, Sester [24] presented an optimization solution that removes short edges using a set of rules first, and then adjusts the simplified building optimally based on least squares adjustment theory. A similar idea was also carried out by Bayer [25] and Liu [26]. Haunert [27] presented an approach for this task by selecting a subsequence of original edges, for which an optimization strategy was used and shape regularities were considered. In the field of

artificial intelligence, Cheng et al. [28] proposed a back-propagation neural network (BPNN) model for learning cartographers' knowledge to realize simplification. Optimization methods attempt to transfer the cartographic knowledge into a machine-understandable form [23], but it still requires a geometry-oriented processing with difficult-to-understand high-level concepts, such as shapes and patterns.

2.2. Shape Recognition in Building Simplification

Shape describes the geometric form of individual spatial objects, and significantly contributes to understanding, evaluating, and predicting geographical phenomena [29,30]. The core of shape analysis is the quantitative description of geometric characteristics and automatic recognition of structural patterns. The former transforms the shape to a descriptive metric that meets some special criteria, e.g., the represented shape is unique. It usually includes region-based methods and boundary-based methods. Region-based descriptors (e.g., position, size, skeleton, curvature, and moment) have a smaller calculation amount, but they are easily affected by noise, and the shape described by a single descriptor is indistinguishable. Boundary-based descriptors fit shape by approximately using a string or function, such as shape context [31], Fourier transform [32], and the turning function [33], that can better analyze the structural features of a shape.

In manual building generalization, cartographers could form the preliminary geospatial understanding in their brain by reading whole maps, and then they would consciously maintain the regularity of similar polygons during the process. Shape analysis recognition is well-attuned to this task. Essentially, shape analysis and recognition aim to search for some specific regularities of individual objects and community groups, often called templates or patterns, which, in turn, are taken as a sign of the workings of a regular process. When the scale is modified, the template or pattern that is represented by a uniform and simpler symbol can be used to replace the generalized buildings. Based on this idea, Rainsford et al. [34] attempted to use a series of simple alphabetic templates to match rural buildings. Sabo et al. [35,36] associated a geometric pattern with cartographic objects and introduced the concept of patterns to on-the-fly map generalization. Liu et al. [37] presented a prototype template method for simplification and employed the shape-number to control the matching accuracy. Li et al. [38] represented the buildings as node degree arrays based on a skeleton and matched them through array operations. Yan et al. [39] used the template matching method to simplify the buildings on medium-small scale topographic maps by analyzing the characteristics of the regional environment. These methods based on shape analysis and template matching are an attempt at the use of pattern in map generalization [40]. This kind of template or pattern can ignore the details and maintain the main characteristics of objects and groups, which is consistent with the simplification process. However, they are mainly concerned with the similarity of the shape and, so far, there is not a good solution for the matching problem. In the field of quality evaluation, Haunert [41] also considered the shape element as a criterion to keep building footprints as legible as they were in the original figure.

3. Methodology

3.1. The Creation of Templates

A shape template can reflect the general shape of similar spatial objects, and it can also be used to abstract the objects out of their minor details and highlight the consistency and regularity of the group. As a result of simplification, templates should be summarized as simplified building shapes intended to meet the requirements of cartographic generalization, and they also should be easily readable and clearly decodable with pleasing aesthetics [42].

At present, people mainly create a template library through manual or empirical means. Reference [34] used a series of English letters (e.g., IFPGELUOT) to build a template library for the rural houses in Denmark, Reference [37] obtained and refined some specific templates for buildings

in certain regions due to their similar characteristics, and Reference [35] created and derived geometric patterns by arranging primitives. These kinds of templates usually have the basic properties in shape representation, including abstraction, symbolization, and indetermination [1]. They can be broadly divided into four categories from the analysis of shape structure:

- Simple Shape: Simple shape uses some basic geometrics to represent buildings abstractly. Typically, buildings have a rectangular geometric form and could even be represented by the minimum bounding rectangle after generalization. Thus, we could employ a rectangle as a template where orthogonality and parallelism could also be imposed.
- Symbolic Shape: A symbolic shape template tries to depict the recognized objects through some imaginable and familiar things to make them easier to understand and be communicated. It is an effective way to enrich the template library by introducing some alphabetic letters, e.g., "E" "T" "Z" "U", or a familiar animal.
- Composite Shape: Some complex structures can hardly be represented by a simple or symbolic template and, thus, the composite shape template should be employed. Such a template has a higher level of abstraction and symbolism and can infuse some more complicated characteristics by arranging different components [43], for example, contextual and humanistic characteristics.
- Automatic Extracted Shape: Manually creating templates may be tedious and time-consuming; thus, some automated method can be introduced to template generation, such as information extraction and shape classification. Similar objects can be extracted from the training samples as templates and used to match buildings within regions with similar environmental patterns.

Figure 1 gives an overview of the above four types of templates. They have their own characteristics, and can be combined flexibly. However, in applying them in the practice of building simplification, their pertinence and effectiveness must be emphasized, and the typical characteristics of the regional environment should also be taken into consideration [39].
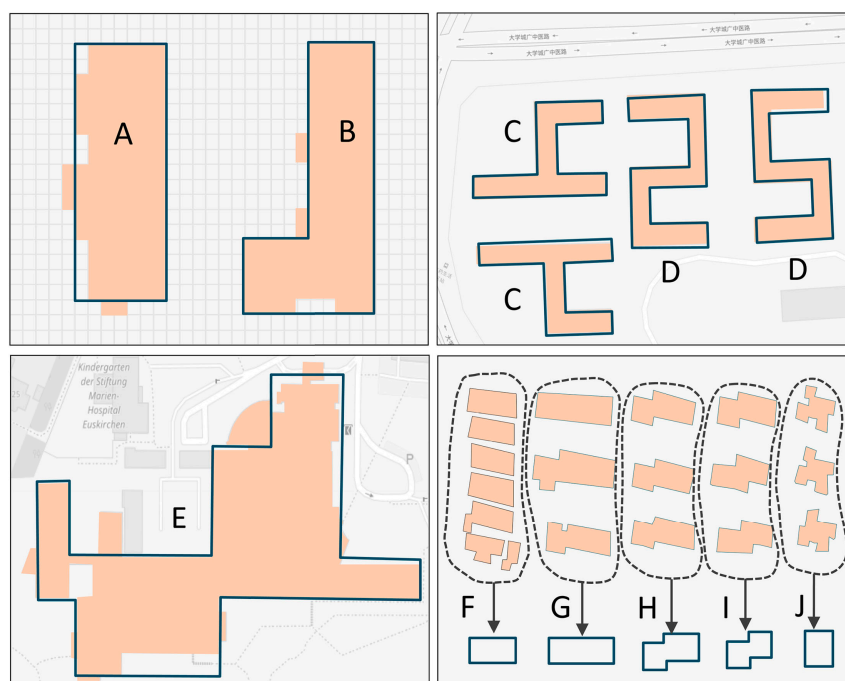

Figure 1. Illustration of four types of templates: templates (**A**,**B**) are simple shapes, (**C**,**D**) are symbolic shapes, (**E**) is a composite shape, and (**F**–**J**) are automatic extracted shapes.

3.2. Building Simplification Using the Template Matching Method

This paper tries to create a series of templates as simplified results for buildings to abstract and symbolize shapes. As shown in Figure 2, for these buildings in a certain region, a series of simple and symbolic shape templates could be created manually, and the simplification procedure for each building in this region has two main steps: shape measurement and template matching. After creating a template library, it is crucial to find a reasonable and effective measure for obtaining similarities between buildings and templates. By comparing the similarities, the "most similar" template can be used as a candidate simplified result for the original building through a geometric transformation.
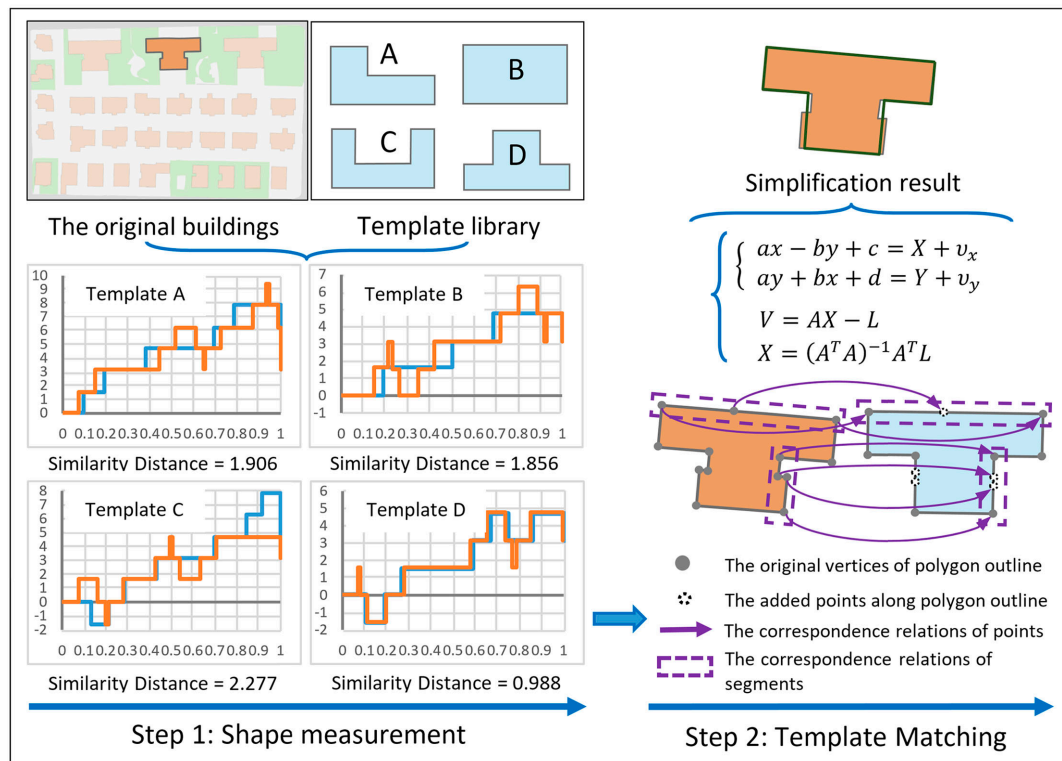
Figure 2. The procedure of building simplification using the template matching method.

3.2.1. Shape Measurement Based on the Turning Function

The turning function, also known as the tangent function, is a well-known method for shape measurement [33], and is broadly applied in spatial query [1] and quality assessment [44]. The turning function converts the polygon shape into a function, where the structure can be analyzed and the similarity can be easily computed in the function space. The details of the turning function are described below [33].

For a simple polygon A, the counterclockwise (or clockwise) tangent angle that point O on the boundary of A makes with the reference orientation (e.g., x-axis) associated with the polygon is denoted by v, and the turning function $f(s)$ is defined as the relationship between the tangent-angle v and the arc-length s. Without loss of generality, the arc-length can be homogenized, and the turning function $f(s)$ of polygon A can be illustrated, as Figure 3.
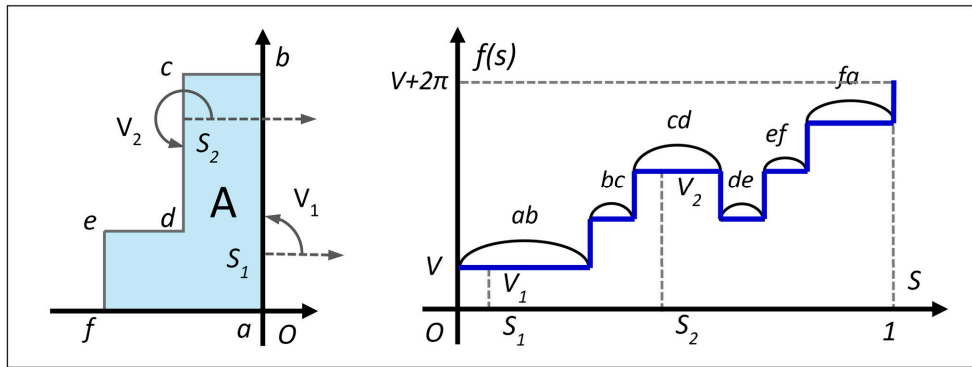
Figure 3. Defining the turning function $f(s)$ of polygon A.

The turning function $f(s)$ keeps track of the turning that occurs and converts a 2D polygon into a piecewise function, making it especially feasible for similarity measurements. Considering two polygons A and B, the similarity $S(A, B)$ can be measured by taking the distance between $f_A(s)$ and $f_B(s)$ in function space, such as the L_2 metrics. However, the result is very sensitive to the choice of the starting point on the boundary, usually through rotating the polygon and shifting the reference point to a minimum. It can be defined as follows:

$$SimilarityDistance(A, B) = \|f_A(s) - f_B(s)\|_2 = \left(\min_{\substack{\theta \in \mathcal{R} \ t \in [-1,1]}} \int_0^1 (f_A(s + t) - f_B(s) + \theta)^2 ds \right)^{\frac{1}{2}}, \quad (1)$$

where t denotes the distance of the starting point shifted along the boundary of A and θ denotes the rotation angle.

It is inferred from Equation (1) that $f(s)$ has several properties that make it especially suitable for our purpose. First, the similarity distance is invariant when translating or scaling the polygons. Second, it is also available for two mirror polygons through controlling the motion direction (in other words, moving the tangent point O counterclockwise or clockwise). In addition, the similarity distance reflects the relative differences between the template and the building in shape, thus, the rotation angle θ does not affect the match result.

3.2.2. Template Matching Based on Least-Squares Adjustment

To ensure the accuracy of the location, and the direction and size of the simplification, the "most similar" template needs to be translated, rotated, and zoomed to match the origin building. This kind of problem can be described by the 2D Helmert Transformation Model [45], which may be expressed as follows:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = k \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \quad (2)$$

where $[x\ y]^T$ and $[X\ Y]^T$ denote the outline point of the building and template, respectively. In this transformation, there is the scale component k, the rotation component θ, and the translation components Δx and Δy.

Let $k \cos \theta = a$, $k \sin \theta = b$, $\Delta x = c$, $\Delta y = d$, and add residuals $[v_x\ v_y]^T$ to this model to develop the observation equation, which can be expressed as follows:

$$\begin{cases} ax - by + c = X + v_x \\ ay + bx + d = Y + v_y \end{cases}, \quad (3)$$

For n pair of points $[(x_1, y_1), \cdots, (x_i, y_i), \cdots, (x_n, y_n)]$ and $[(X_1, Y_1), \cdots, (X_i, Y_i), \cdots, (X_n, Y_n)]$, the total observation equation can be expressed as follows:

$$V = AX - L, \tag{4}$$

where:

$$A = \begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ & \cdots & & \\ x_i & -y_i & 1 & 0 \\ y_i & x_i & 0 & 1 \\ & \cdots & & \\ x_n & -y_n & 1 & 0 \\ y_n & x_n & 0 & 1 \end{bmatrix}_{2n \times 4}, X = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}_{4 \times 1}, L = \begin{bmatrix} X_1 \\ Y_1 \\ \cdots \\ X_i \\ Y_i \\ \cdots \\ X_n \\ Y_n \end{bmatrix}_{2n \times 1},$$

Based on the least-squares adjustment, which ensures that the $V^T V$ is minimized, an estimate solution for transformation parameters is derived by the following equation:

$$X = \left(A^T A\right)^{-1} A^T L, \tag{5}$$

In the above model, at least three control points must be given. In our case, we apply a dynamic programming algorithm that was proposed by Nöllenburg et al. [46] to find the correspondence between the outline of the building and the candidate template, but other methods can also be used to solve this problem as well. The segments of two outlines can build a relationship of 1-0, 1-1, or 1-m, and the control points can be derived from the corresponding segments. To ensure that the control points are in one-to-one correspondence, some necessary points will be added along the boundaries, as shown in Figure 4.
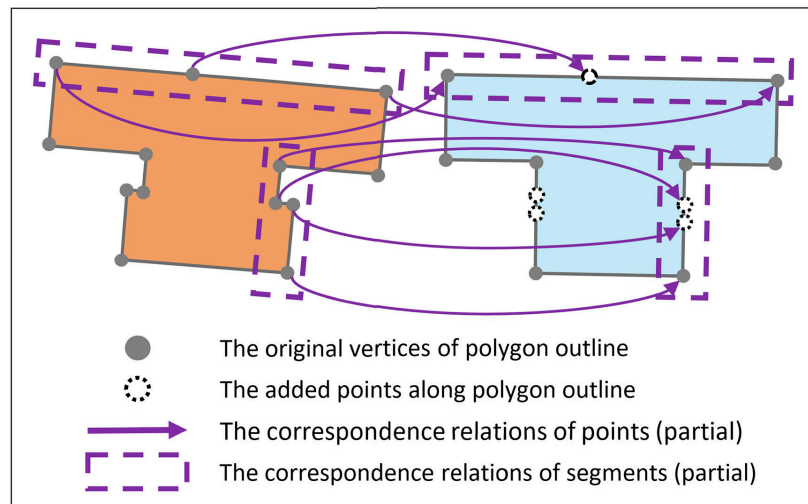

Figure 4. Illustration of the correspondence relation of segments and control points.

4. Experiments and Discussion

Experiments were carried out in this section to assess the reasonability and feasibility of the proposed approach. The contents included two aspects: (1) qualitative and statistical evaluations of the simplification, and (2) a further discussion about the scaling of templates. A comparison with existing methods is also presented.

4.1. Quality Assessment of Simplification

4.1.1. Experimental Datasets and Templates

The experimental data chosen was a large-scale 1:2000 dataset from a topographic map that was extracted from the city of Guangzhou, China (Figure 5). The study area is approximately 1.2 km × 1.3 km and contains 625 building polygons. The study area exists as several typical residential communities, which are common structures among the urban buildings. Within each community, the shape of the buildings is similar and stable. Some basic geographic features, such as roads and vegetation, are also included in this study area. They covered the buildings in the representation but did not cut them.
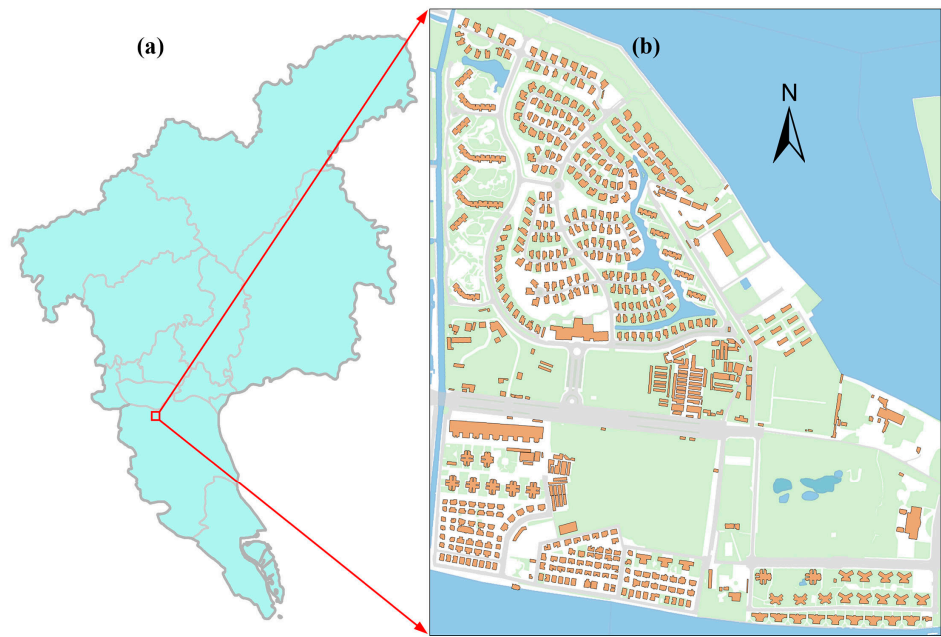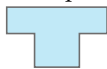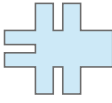

Figure 5. Experimental region and dataset: (**a**) Guangzhou region, and (**b**) the experimental dataset.

With respect to the simplification of this experimental dataset, we manually created a diversified template library. It contains 54 templates, mainly including some standard geometries (e.g., rectangle and circle), symbolic symbols (e.g., T-shaped, U-shaped, L-shaped), and some special templates derived from the typical building shapes in the dataset. Twelve examples are shown in Table 1.

Table 1. Example of a template library for an experimental dataset.

Categories	Template			
Simple template	Square-shaped	Rectangle-shaped	General-shaped	General-shaped
Symbolic template	U-shaped	Y-shaped	T-shaped	L-shaped
Composite shape				

4.1.2. Qualitative and Statistical Evaluation

We used these templates to simplify the experimental datasets with the target scale of 10,000, and the result is shown in Figure 6a (simplified buildings are superimposed on the initial buildings). There is no selection or typification in this experiment. It can be seen from the results that a majority of the buildings in the datasets were simplified by the predefined template and their shapes were well-preserved, but there are still some buildings in the dataset that were far from satisfactory.
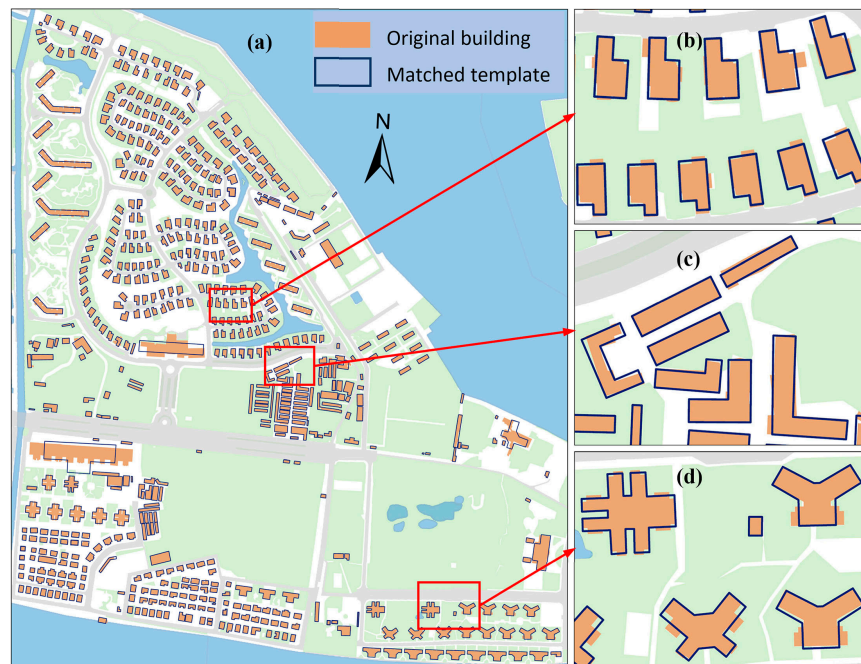

Figure 6. The result of building simplification: (**a**) the comparison of buildings before and after simplification using template matching; (**b**–**d**) showing the local details.

To be more specific, some building polygons that have similar characteristics were simplified successfully by the same template, as shown in Figure 6b. After the template matching simplification, their shapes were more regular and uniform. This situation often appears in large-scale city maps like the one employed in the experiment.

Another situation is that these buildings have an auxiliary cognizance, for example, the shapes are similar to letters (Figure 6c). As we have seen, they can also be simplified by the symbolic templates and their main tendencies are maintained after simplification. In addition to the large-scale map, this situation also often appears in a medium-small scale map, e.g., the scattered houses in a rural region. Specific to such characteristics of buildings, we did some extra tests by using a 1:50,000 scale topographic map from the central and western region in Sichuan, China. This result is shown in Figure 7, and we found that most of the scattered houses were simplified by the simple or symbolic templates, with some area-up areas failing to be simplified, which was consistent with our experiment. Regarding the performance of the template matching method on a medium-small scale map, there are more discussions in the work of Rainsford et al. [34] and Yan et al. [39].
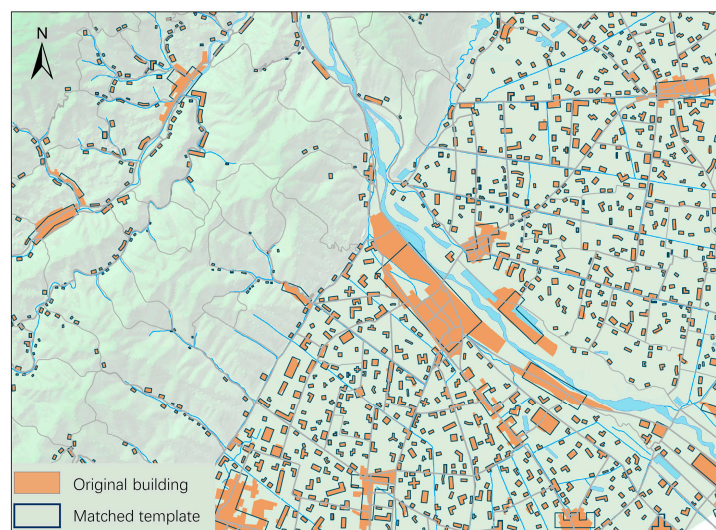
Figure 7. An extra experiment to test the scattered houses in a medium-small scale map.

Additionally, in Figure 6d there are some special buildings with a particular shape, which the simple and symbolic templates may not able to handle. We consciously created several corresponding templates to simplify them, and their orthogonality and symmetry could be imposed, thus the results achieved the desired effect. With the decrease of the target scale, these special buildings become simpler, so it is also possible that they are simplified by the simple or symbolic templates at a smaller scale, e.g., for the building located in the top-right of Figure 6d, a cross-shaped or even a square template can be used.

For the traditional methods, every building is analyzed and processed independently, thus, it is difficult to achieve a regular and uniform shape as its simplified version. A simple and the most common example is shown in Figure 8; although the three buildings in a residential community are very similar, their shapes are still different by simplifying with a traditional method on the ESRI's (Environmental Systems Research Institute, Redlands, CA, USA) ArcGIS platform (the green outlines). A better strategy is to replace them with the same shape by the template matching method (the dark outlines). This may be slightly inferior in geometric accuracy, but it has a relatively large advantage in terms of regularity and uniformity, which are helpful to enhance shape cognition and improve map readability. Additional, the traditional method does not maintain orthogonality and parallelism consciously, which needs some special algorithms for further processing [3].
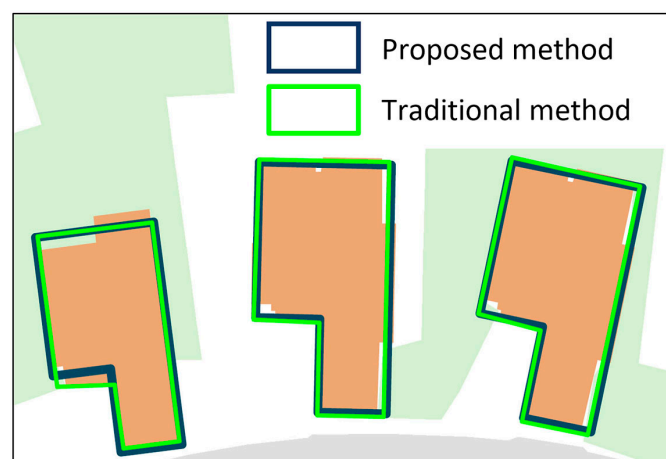

Figure 8. A contrastive analysis of the proposed method and a traditional method.

Further, this paper statistically assessed the evaluation of the quality of the simplification quantitatively. For that purpose, the surface distance was calculated. Considering two polygons A and B, the metric can be defined by comparing the areas of the intersection and union between them [47]:

$$SurfaceDistance = 1 - \frac{Area(A \cap B)}{Area(A \cap B)} \qquad (6)$$

In this experiment, the surface distances are concentrated mainly in the range of 0.0 to 0.25 and the proportion is 94.2%, meaning that the simplification by the proposed method is satisfied. The distribution diagram and numerical evaluation of surface distances for all simplified buildings are combined with the results of comparison tests, which will be presented in Section 4.3.

To analyze the cause for the local failure of simplification, we chose these 36 buildings whose surface distances are greater than 0.25 and found that there are three main aspects that may cause errors: (1) the shape of the original building is too complicated to find a suitable template in the template library; (2) there is a suitable template in the library but the wrong template was still found when calculating the similarity distance; and (3) a destructive offset of position and orientation occurs when the transformation is performed using the least squares theory. All unsatisfied buildings of the experimental dataset were counted, and these three types of errors occur for 12, 15, and nine, respectively. An example for each type of error is shown in Figure 9.
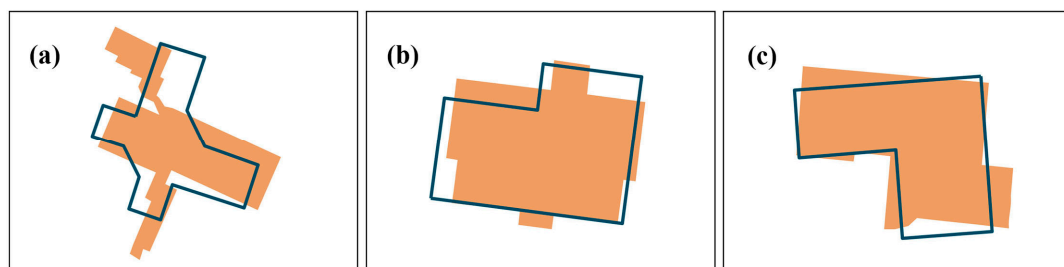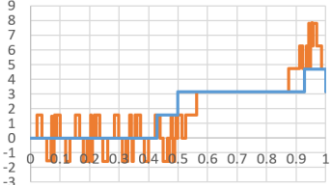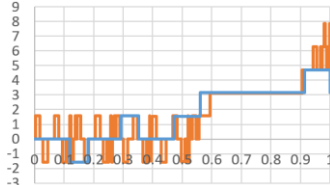

Figure 9. Example for each type of error: (**a**) the first type of error, (**b**) the second type of error, and (**c**) the third type of error.

The first type of error is what might have been expected because some building polygons are usually quite large and have complicated shapes, such as built-up areas, and, as a result, many of them cannot find a suitable template. For example, we did not create a special template for the complicated building in Figure 9a, thus, it failed to be simplified. For the second and third types of error, through in-depth analysis we found that although the turning function and least squares methods performed well, there still exists something unsatisfactory in the simplification, such as that for the one located in the south-west of Figure 6a which is falsely matched as a T-shaped template. To gather more detailed information, we chose two templates (a T-shaped template that matched in the experiment and an I-shaped that may be better matched) to test the southwest one, and the results are shown in Table 2. As we can see, the turning function of the building has a large amount of oscillation because of its comb-type shape, and both the T-shaped template and I-shaped template failed to match it correctly. In fact, due to the complexity of the shapes, there is no shape representation method that can achieve 100% correctness for similarity measurements, and the orientation is also difficult to precisely define [48]. On the other hand, the accuracy of building simplification is also closely related to the scaling of the template, which will be discussed in the next section.

Table 2. Comparative analysis of a building matched by I-shaped and T-shaped templates.

	I-Shaped Template	T-Shaped Template
Building to be simplified	Similarity Dis = 2.167 Surface Dis = Null	Similarity Dis = 2.083 Surface Dis = 0.500

4.2. Discussion of the Scaling of Template Matching

From the perspective of shape cognition, every template shape can abstract a shape structure or imply a symbolic meaning, but there may be several templates with slightly different shapes that are pertinent to the same meaning. This property is similar to calligraphy. The handwriting of everybody is different; some people are incredibly neat, while others are messy. This writing, however, is still regarded as the same letters or words. This reflects the hierarchy of the template shape, which has a great influence on the accuracy of simplification. For example, templates A and B in Figure 10 have the same symbolic meaning in shape cognition, but the results of the shape distance and surface distance are very different.
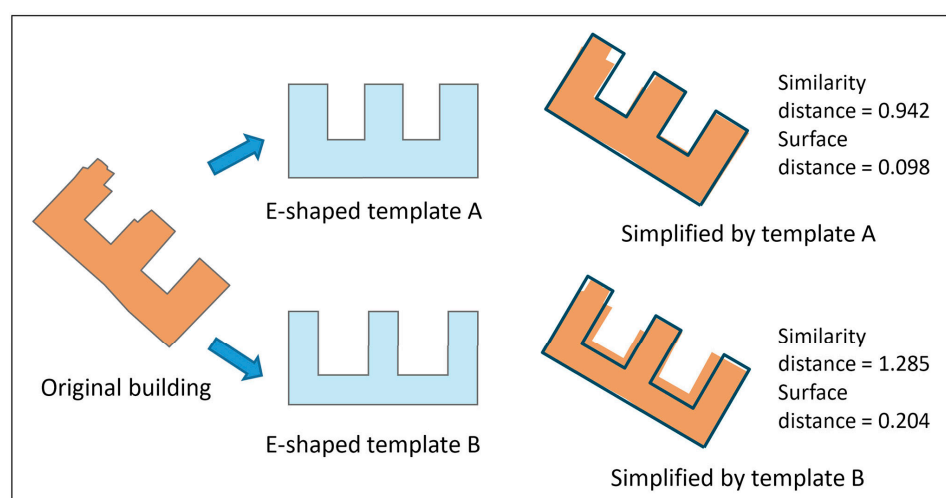
Figure 10. Illustration of the simplification results by different templates that have the same symbolic meaning.

Since the templates are created manually in the experiment, their shapes should be predefined for each simple shape or symbolic meaning by considering this property before the simplification. For example, we defined three shapes with the 3:2, 2:1, and 3:1 length-to-width ratios as rectangular templates. The more detailed the templates, the better the simplification, but also the greater the calculation; thus, one should balance the computing power and generalization requirement based on the characteristics of the dataset when creating the library.

To ensure the quality of the generalization, the scaling of templates should also be considered after the matching simplification. First, every adjusted template through the least squares method should satisfy

the cartographic constraints, such as the shortest edge greater than 0.3 mm [6,14,42]. Second, the surface distance should be less than a given threshold, and the threshold is set according to the characteristics of different types of buildings. Reference [39] suggests that an adventurous parameter value can be appropriate for buildings that are smaller than 25 mm^2 in an origin scale of 1:50,000, but a conservative value should be used for buildings with larger areas and more complicated shapes. Referring to the settings, we used 0.25 and 0.2 for this parameter, respectively. Considering the generalization constraints, the correctness of the matching simplification in the above experiment is 88.4%.

4.3. Comparison with Existing Methods

The proposed simplification algorithm chooses a candidate template to match and simplify the building. The shape measurement step has been widely covered in the earlier works. Ai et al. [1] compared the Fourier transform and the turning function on the performance of similarity measures on an ordinal scale, and concluded that the Fourier transform is suitable for natural and sinuous objects, but for artificial objects, such as the buildings in our case, the turning function is more suitable.

In this paper, we mainly carried out an experiment to compare the performance of our method and existing methods in the template matching step. As a comparison task, two existing methods were also implemented: first, the geometric matching method based on some simple operations (rotation, positioning, and scaling) was proposed by Sabo et al. [35,36], followed by the work of Yan et al. [39] who tested these operations by using the direction, gravity, and area of original building as the parameters, and then the MBR-based method that forces the minimum bounding rectangle (MBR) of the template to coincide with the building [1,35]. The comparison of the distribution of the surface distances for all buildings is presented in Figure 11, and the mean, median, maximum, minimum, and standard deviation values of the surface distance in this experiment are presented in Table 3.
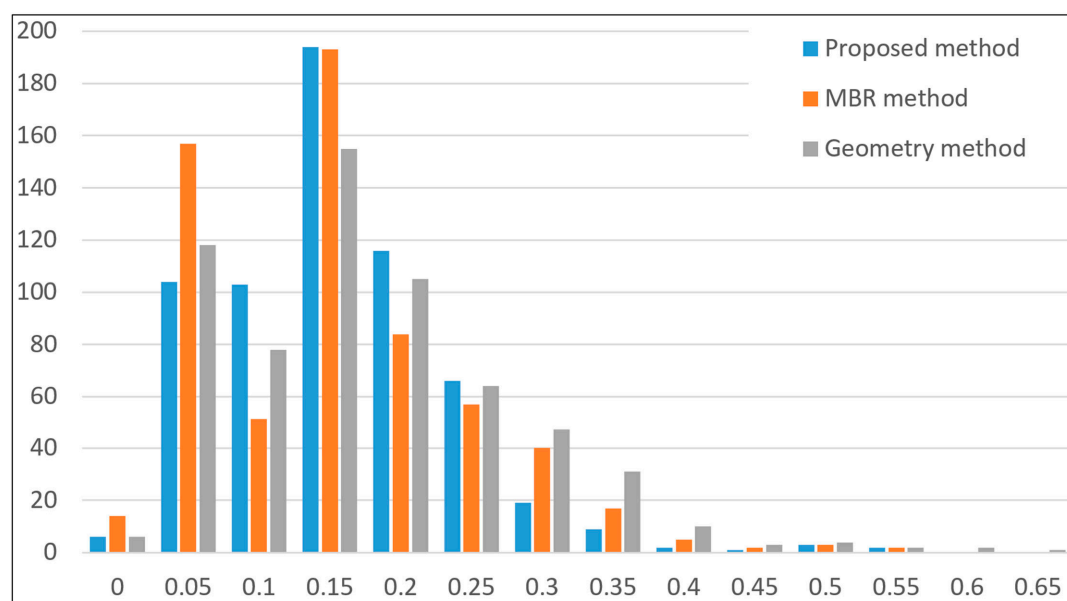
Figure 11. The distribution of surface distances of all buildings. The horizontal axis denotes the surface distances, and the vertical axis denotes the number of buildings.

Table 3. Numerical evaluation of the compared experiments.

	Mean	Median	Standard Deviation	Maximum	Minimum
Proposed method	0.126	0.126	0.075	0.520	0.0
MBR method	0.123	0.122	0.093	0.522	0.0
Geometry method	0.144	0.131	0.097	0.610	0.0

From these results, it can be seen that most of the surface distances are between 0.05 and 0.25, which means that the three methods can match templates with the original buildings well. The numerical evaluation shows that the proposed method and the MBR-based method are better than the geometric method. In detail, the mean value of the MBR-based method has a slight advantage over our proposed method, but the standard deviation comparisons show that the proposed method is more stable than the other two methods.

This may be explained by the fact that, as a kind of basic and simple method, the geometric method can be improved through the MBR-based method and our proposed method. The MBR-based method works by stretching the template to cover the original building. It is especially useful for those buildings whose shape is close to rectangular; this situation will lead the MBR-based method to have a better surface distance, as shown in Figure 12a. There are some such cases in the experimental dataset, e.g., the small buildings in the bottom-left, for which we can see that there is a peak near zero from the distribution of surface distances. However, it is unstable because the minimum boundary rectangle of the polygon is quite different from the shape in many cases. An example is shown in Figure 12b, in which the direction is difficult to be controlled by the MBR. In addition, the stretching will seriously cause the deformation of the shape, which cannot fit the original intention of template simplification, for example, if there is a slender enation on the outline (Figure 12c).
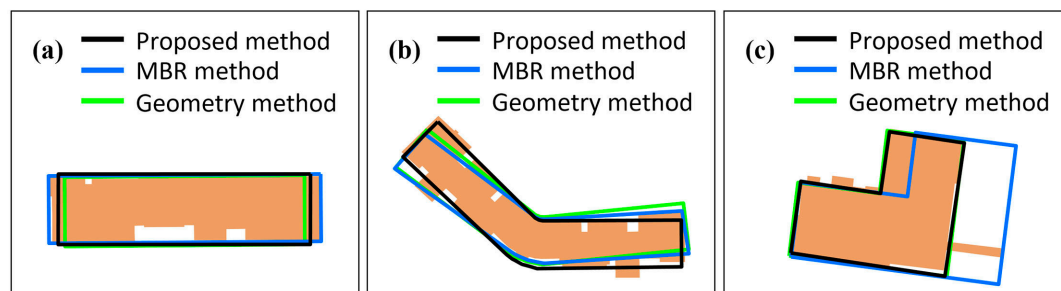

Figure 12. A visual analysis of the compared experiments: (**a**) a case where the minimum boundary rectangle of a polygon is close to the shape, (**b**) a case where the minimum boundary rectangle of a polygon is quite different from the shape, and (**c**) a case with a slender enation on the outline.

Our proposed method works by adjusting the control points of the polygon outline based on the least squares model, which is a dynamic process according to the whole structure of the shape and thus can overcome this shortcoming. However, because the shape is predefined and fixed, there are some templates that cannot match the original building with very high precision. The problem can also be improved by enriching the template library if the calculation conditions permit. Thus, it can be concluded that the MBR-based method may be useful for those buildings whose shape is close to a rectangle, but for the others, the proposed method is the best for meeting the requirements of building simplification, and it can also be consistent with predefined shapes.

5. Conclusions and Outlook

Building simplification is a classic problem in map generalization. Conventionally, both these approaches based on community structure or by eliminating small bends cannot achieve the unity of regional and individual characteristics. This paper proposes a simplification method using template matching. The basic idea is to consider the requirements of cartographic generalization and create a series of abstracted and symbolic templates for simplifying the buildings by generalizing building polygons and analyzing the typical structure characteristics and humanistic elements. The approach identifies the shape structure as a whole, and uses templates to represent the initial buildings, rather than dealing with geometric details locally. The template matching approach takes the overall distribution into consideration. For example, the buildings in a certain area can be represented

by the same template to maintain their spatial context characteristics and enhance their symbolic readability, making them more significant. In addition, the templates can preserve orthogonality.

The proposed method includes two steps for shape measurement and template matching. In the first step, we employ the turning function as the shape descriptor to measure the similarity to find the best candidate template for simplified buildings. In the second step, we carry out a transformation to adjust the geometry position of the candidate template based on the least squares model. By testing against real datasets, this research suggests that our approach can maintain the main structure of the building shapes and geometry accurately. The proposed method has also proven to be useful by the comparison with two existing methods. Furthermore, the scaling of the template and the completeness of the template library have an impact on the quality of the simplification. Consequently, the templates should be further enriched and extended in the future, for instance, by considering the situation where buildings are complicated with more than one hole and using some other sources of information (e.g., street network) to avoid some potential conflicts.

Author Contributions: Xiongfeng Yan and Tinghua Ai together conceived the template matching simplification idea; Xiongfeng Yan collected the data, designed the experiments, and wrote the manuscript; Tinghua Ai and Xiang Zhang performed the analysis and modified the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ai, T.; Cheng, X.; Liu, P.; Yang, M. A shape analysis and template matching of building features by the Fourier transform method. *Comput. Environ. Urban Syst.* **2013**, *41*, 219–233. [CrossRef]
2. Ai, T.; Ke, S.; Yang, M.; Li, J. Envelop generation and simplification of polylines using Delaunay triangulation. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 297–319. [CrossRef]
3. Lokhat, I.; Touya, G. Enhancing building footprints with squaring operations. *J. Spat. Inf. Sci.* **2016**, *13*, 1–28.
4. Mackaness, W.; Burghardt, D.; Duchêne, C. Map generalisation: Fundamental to the modelling and understanding of geographic space. In *Abstracting Geographic Information in a Data Rich World*; Burghardt, D., Duchêne, C., Machaness, W., Eds.; Springer: Cham, Switzerland, 2014; pp. 1–15.
5. Ai, T.; Guo, R. Polygon cluster pattern mining based on Gestalt principles. *Acta Geod. Cartogr. Sin.* **2007**, *36*, 302–308.
6. Steiniger, S.; Taillandier, P.; Weibel, R. Utilising urban context recognition and machine learning to improve the generalisation of buildings. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 253–282. [CrossRef]
7. Du, S.; Shu, M.; Feng, C. Representation and discovery of building patterns: A three-level relational approach. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 1161–1186. [CrossRef]
8. Ai, T.; Zhang, X.; Zhou, Q.; Yang, M. A vector field model to handle the displacement of multiple conflicts in building generalization. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 1310–1331. [CrossRef]
9. Yu, W.; Ai, T.; Liu, P.; Cheng, X. The analysis and measurement of building patterns using texton co-occurrence matrices. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1–22. [CrossRef]
10. Gaffuri, J.; Trévisan, J. Role of urban patterns for building generalisation: An application of AGENT. In Proceedings of the 8th ICA Workshop on Generalisation and Multiple representation, Leicester, UK, 20–21 August 2004.
11. Gong, X.; Wu, F. A typification method for linear pattern in urban building generalization. *Geocarto Int.* **2016**, 1–19. [CrossRef]
12. Zhang, X.; Ai, T.; Stoter, J.; Kraal, M.; Molenaar, M. Building pattern recognition in topographic data: Examples on collinear and curvilinear alignments. *Geoinformatica* **2013**, *17*, 1–33. [CrossRef]
13. Yang, W. Identify building patterns. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing, Beijing, China, 3–11 July 2008.
14. Li, Z.; Yan, H.; Ai, T.; Chen, J. Automated building generalization based on urban morphology and Gestalt theory. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 513–534. [CrossRef]

15. Buchin, K.; Meulemans, W.; Renssen, A.; Speckmann, B. Area-preserving simplification and schematization of polygonal subdivisions. *ACM Trans. Spat. Algorithms Syst.* **2016**, *2*, 1–35. [CrossRef]

16. Chen, W.; Long, Y.; Shen, J.; Li, W. Structure recognition and progressive simplification of the concaves of building polygon based on constrained D-TIN. *Geomat. Inf. Sci. Wuhan Univ.* **2011**, *36*, 584–587.

17. Xu, W.; Long, Y.; Zhou, T.; Chen, L. Simplification of building polygon based on adjacent four-point method. *Acta Geod. Cartogr. Sin.* **2013**, *42*, 929–936.

18. Damen, J.; van Kreveld, M.; Spaan, B. High quality building generalization by extending the morphological operators. In Proceedings of the 12th ICA Workshop on Generalisation and Multiple Representation, Montpellier, France, 20–21 June 2008.

19. Meijers, M. Building simplification using offset curves obtained from the straight skeleton. In Proceedings of the 19th ICA workshop on Generalisation and Multiple Representation, Helsinki, Finland, 14 June 2016.

20. Guo, R.; Ai, T. Simplification and aggregation of building polygon in automatic map generalization. *J. Wuhan Technol. Univ. Surv. Map.* **1999**, *25*, 25–30.

21. Kada, M. Aggregation of 3D buildings using a hybrid data approach. *Cartogr. Geogr. Inf. Sci.* **2011**, *38*, 153–160. [CrossRef]

22. Wang, H.; Wu, F.; Zhang, L.; Deng, H. The application of mathematical morphology and pattern recognition to building polygon simplification. *Acta Geod. Cartogr. Sin.* **2005**, *34*, 269–276.

23. Sester, M. Optimization approaches for generalization and data abstraction. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 871–897. [CrossRef]

24. Sester, M. Generalization based on least squares adjustment. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing, Amsterdam, The Netherlands, 16–23 July 2000.

25. Bayer, T. Automated building simplification using a recursive approach. In *Cartography in Central and Eastern Europe*; Gartner, G., Ortag, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 121–146.

26. Liu, P.; Ai, T.; Deng, J. Simplification and rectangularity of building-polygon based on least squares adjustment. *J. China Univ. Min. Technol.* **2008**, *37*, 699–704.

27. Haunert, J.H.; Wolff, A. Optimal and topologically safe simplification of building footprints. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, San Jose, CA, USA, 3–5 November 2010.

28. Cheng, B.; Liu, Q.; Li, X.; Wang, Y. Building simplification using backpropagation neural networks: A combination of cartographers' expertise and raster-based local perception. *GISci. Remote Sens.* **2013**, *50*, 527–542.

29. MacEachren, A. Compactness of geographic shape: Comparison and evaluation of measures. *Geogr. Ann. Ser. B-Hum. Geogr.* **1985**, *67*, 53–67. [CrossRef]

30. Wentz, E.A. Shape analysis in GIS. In Proceedings of the Auto-Carto, Seattle, WA, USA, 7–10 April 1997.

31. Belongie, S.; Malik, J.; Puzicha, J. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 509–522. [CrossRef]

32. Persoon, E.; Fu, K.S. Shape discrimination using Fourier descriptors. *IEEE Trans. Syst. Man Cybern.* **1977**, *7*, 170–179. [CrossRef]

33. Arkin, E.M.; Chew, L.P.; Huttenlocher, D.P.; Kedem, K.; Mitchell, J.S. An efficiently computable metric for comparing polygonal shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 209–216. [CrossRef]

34. Rainsford, D.; Mackaness, W. Template matching in support of generalisation of rural buildings. In *Advances in Spatial Data Handling, Proceedings of the 10th International Symposium on Spatial Data Handling, Ottawa, ON, Canada, 9–12 July 2002*; Richardson, D.E., van Oosterom, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 137–151.

35. Sabo, M.N.; Bédard, Y.; Bernier, E. Methodology for developing a database of geometric patterns to better support on-the-fly map generalization. In Proceedings of the 22th International Cartographic Conference, A Coruna, Spain, 9–16 July 2005.

36. Sabo, M.N.; Bédard, Y.; Moulin, B.; Bernier, E. Toward self-generalizing objects and on-the-fly map generalization. *Cartogr. Int. J. Geogr. Inf. Geovis.* **2008**, *43*, 155–173. [CrossRef]

37. Liu, P.; Ai, T.; Hu, J.; Cheng, X. Building-polygon simplification based on shape matching of prototype template. *Geomat. Inf. Sci. Wuhan Univ.* **2010**, *35*, 1369–1372.

38. Li, H.; Wang, Y.; Yu, Z. Template match of buildings and results analysis. In Proceedings of the 18th IEEE International Conference on Geoinformatics, Beijing, China, 18–20 June 2010.

39. Yan, X.; Ai, T.; Yang, M. A simplification of residential feature by the shape cognition and template matching method. *Acta Geod. Cartogr. Sin.* **2016**, *45*, 874–882.

40. Mackaness, W.; Edwards, G. The importance of modeling pattern and structure in automated map generalization. In Proceedings of the Joint ISPRS/ICA Workshop on Multi-Scale Representations of Spatial Data, Ottawa, ON, Canada, 7–8 July 2002.

41. Haunert, J.H. A symmetry detector for map generalization and urban-space analysis. *ISPRS J. Photogramm. Remote Sens.* **2012**, *74*, 66–77. [CrossRef]

42. Weibel, R.; Dutton, G. Constraint-based automated map generalization. In Proceedings of the 8th International Symposium on Spatial Data Handling, Vancouver, BC, Canada, 12–15 July 1998.

43. Biederman, I. Recognition-by-components: A theory of human image understanding. *Psychol. Rev.* **1987**, *94*, 115. [CrossRef] [PubMed]

44. Fan, H.; Zipf, A.; Fu, Q.; Neis, P. Quality assessment for building footprints data on OpenStreetMap. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 700–719. [CrossRef]

45. Hofmann-Wellenhof, B.; Lichtenegger, H.; Collins, J. *Global Positioning System Theory and Practice*, 5th ed.; Springer: Vienna, Austria, 2001; pp. 229–251.

46. Nöllenburg, M.; Merrick, D.; Wolff, A.; Benkert, M. Morphing polylines: A step towards continuous generalization. *Comput. Environ. Urban Syst.* **2008**, *32*, 248–260. [CrossRef]

47. Vauglin, F.; Bel Hadj Ali, A. Geometric matching of polygonal surfaces in GISs. In Proceedings of the ASPRS Annual Meeting, Tampa, FL, USA, 30 March–3 April 1998.

48. Duchêne, C.; Bard, S.; Barillot, X.; Ruas, A.; Trévisan, J.; Holzapfel, F. Quantitative and qualitative description of building orientation. In Proceedings of the 7th ICA Workshop on Progress in Automated Map Generalisation, Paris, France, 28–30 April 2003.