Pattern Recognition of Complex Distributed Ditches

Chengyi Liu, Fang Wu, Xianyong Gong * ©, Ruixing Xing and Jiawei Du

Institute of Geospatial Information, Information Engineering University, Zhengzhou 450001, China; liuchengyi@zju.edu.cn (C.L.); wufang_6306126.com (F.W.); xingrxgis@whu.edu.cn (R.X.); whxdjw@126.com (J.D.)
* Correspondence: gongxygis@whu.edu.cn

Abstract: The ditch pattern reflects the distribution characteristics of an agricultural drainage system and needs to be detected to enrich the data source before map generalization. Due to several breaks, the connectivity of the ditches is destroyed between ditches and rivers in map representation, making its structure complex. Previous studies have primarily focused on extracting parallel ditches, paying less attention to complex ditches with breaks. The pattern recognition of ditches does not merely involve the extraction of parallel relation. It involves the extraction of different level relations, which is a great challenge. Therefore, this study proposes a novel model to describe the complex structure of ditches. Our work consisted of the following three main contributions: (1) the collinear relation was defined to detect the groups of ditch segments separated by breaks, (2) the detection method of parallel relation was improved throughout the multi-parameter combined constraints, and (3) the main-tributary relation was proposed to build a connection between parallel groups and rivers. The experimental results showed that the proposed method was solved effectively in complex ditch pattern recognition.

Keywords: data enrichment; spatial pattern recognition; complex structure; map generalization

1. Introduction

Map generalization is the selection and simplified representation of details appropriate to the target map scale while preserving the main structures [1]. The cartographers use the detailed knowledge to analyze the context of a map for map generalization. In the past three decades, researchers have attempted to formalize this knowledge and built automated systems to simulate the process of man-made map generalization [2–5]. The knowledge of map generalization was divided into the following three types: (1) geometric knowledge (geometric characteristics); (2) structural knowledge (spatial characteristics); and (3) procedural knowledge (various operations and strategies) [6]. Spatial pattern—a typical structural knowledge—is a kind of high-level knowledge implicit in the spatial relations of geo-features. The detection of spatial pattern is meant to enrich the original data and serve for multi-scale representation [6,7].

Ditches are typical man-made drainages and have a great influence on transporting material and draining water in agricultural districts, reflecting the layout of farmland and affecting the landscape of a farmland ecosystem [8]. They are commonly found in groups in the rural environment and are widely used to describe such a landscape in maps, especially thematic maps of agriculture [9]. Hence, the generalization of ditches should emphasize their structure or spatial pattern, which would be enriched previously. However, the following two aspects make the data enrichment of ditch structure more difficult.

Firstly, the mixed distribution with both discrete and continuous characteristics makes the pattern recognition of complex ditches more challenging. Spatial patterns widely exist in geo-features, such as buildings, roads, rivers, and ditches. These spatial patterns are constructed by adjacency relation. However, their detection strategies differ from each other. Geo-features can be divided into discrete (e.g., buildings), continuous (e.g.,
roads and rivers), and mix-distributed geo-features (e.g., ditches) based on various distribution. Discrete geo-features extensively use constrained Delaunay triangulation to detect proximity relation [10,11]. Some researchers divided continuous geo-features into segments and constructed connection relations of each segment [12–14]. However, mix-distributed geo-features have both discrete and continuous characteristics, unlike the other two geo-features mentioned above. The current detection strategies for adjacency relation are not applicable for mix-distributed geo-features. Therefore, it is a challenge to detect various relations and their complex pattern. Focusing on the last, this study is particularly concerned with different relations detection of complex mix-distributed ditches based on human visual perception and the Gestalt theory [15].

The existing methods mainly aim to detect parallel ditches and typification [7,9]. Moreover, few studies have been developed focusing on complex distributed ditches. Connectivity is an important factor for pattern recognition of ditches considering good draining water. However, there are many breaks—underground pipes—between ditches and rivers in map data for the passage of agricultural machines (Figure 1). These breaks have a negative impact on the continuity of ditches, making it challenging to preserve their spatial patterns in map generalization. The breaks make the adjacency relations of ditches do not remain at an equal level similar to that of buildings, roads, and rivers. Hence, we propose a novel multi-relation structure model to describe these complex distributed ditches.

![Figure 1. Breaks between ditches: (a) a break in the ditch of real image data and (b) map data, (c) a break at the junction of ditches of real image data, and (d) map data.](image)

The following are the main contributions of this study. First, collinear relation was identified to repair the continuity of ditch segments. After that, several factors, including orientation, length, and arrangement direction, were considered to improve the quality of detecting parallel relation. Lastly, the main-tributary relations were detected to preserve the connection between parallel groups and river segments. As a consequence, a bottom–up model with three relations described the complex structure of ditches effectively.

The remainder of this paper is organized as follows: The following section (Section 2) refers to related work. Section 3 presents the definition of a complex structure of ditches based on the Gestalt theory, including collinear, parallel, and main-tributary relations. Their differences are analyzed in detail. After that, the relevant identified parameters are presented. The method of detecting complex pattern is proposed in Section 4. Section 5 discusses the experimental results. Finally, this paper ends with conclusions (Section 6).
2. Related Work

The spatial pattern is an important issue of human visual perception and spatial relation reasoning [16], which is beneficial in enriching the map data and formalizing the structural knowledge of the complex distribution of ditches. It involves two main procedures for spatial pattern recognition: spatial pattern definition and pattern detection.

First, various spatial patterns were defined by the function and geometric topology of geo-features.

- Buildings were classified into industrial, inner city, urban, suburban, and rural areas [17]. Similarly, the topological connectivity of ditches must be considered for good drainage function, which is an important task in ditch pattern recognition. On the other hand, function had a close correlation with geometric topology. Buildings with single-family, multifamily, and nonresidential attributes have different geometric and distribution characteristics [18]. Regnauld [19] proposed that buildings have linear, circular, star, and other irregular patterns. Zhang [10] expanded building patterns to include collinear, curvilinear, align-along-road, grid-like, and unstructured patterns. These pattern definitions of discrete geo-features are combinations of homogeneous adjacency relations at a fixed map scale.

- Relatively, the patterns of continuous geo-features are defined by homogeneous adjacency relations at different map scales. Road networks interweave to form patterns such as stroke, grid, star, and ring [13,20,21]. Similarly, main-tributary relations are built by Horton, Strahler, or Shreve hierarchical models in river networks [12,22,23]. Furthermore, river patterns, such as dendritic, parallel, pinnate, rectangular, and trellis, are defined by different morphological structures because the concept of river patterns has changed from geomorphology to cartography [24–26].

- In their work, they undifferentiated treated adjacency relations and considered adjacency relations as the basic unit of spatial pattern in the same level. However, several heterogeneous geo-features were gradually considered as a complete unit based on integrity in the Gestalt principle [15] and contour integration by human visual perception [27]. Liu [28] defined major, minor, and parallel relations to combine broken ponds as a whole and established the multi-hierarchical structure. A combined collinear pattern was proposed to describe complex collinear buildings [29]. Therefore, combined relations based on visual integrity and multi-hierarchical structure make it possible to define complex spatial patterns.

Second, the effective methods for detecting spatial patterns were also the curial procedure. As mentioned above, adjacency relation is the basis for extracting spatial patterns.

- For discrete geo-features, a graph composed by geo-feature nodes and adjacency relations was usually employed to model adjacency relations. After that, several geometric indices were inferred to describe the similarity of adjacency geo-features. Proximity, size, shape, area, orientation, and arrangement direction were widely used for recognizing building patterns [30–32].

- On the other hand, for continuous or mix-distributed geo-features, connection angle, length, orientation, and other geometric and topological indices were used to detect spatial patterns of roads and rivers [25,26]. Zhang [33] proposed an index called parallel factor for recognizing a two-lane road. The orientation and centroid distance of ditches were considered for a parallel pattern [9]. Furthermore, Tian [34] adopted orientation, length, and perceptual distance to detect parallel ditches. They only focused on parallel patterns based on homogeneous adjacency relations. However, the various relations between ditches and rivers should be a concern.

To summarize, the reviewed work focused on characteristics of homogeneous adjacency relations in spatial patterns, and multiple constraints were designed for pattern recognition. However, their method of detecting patterns based on homogeneous adjacency relations hardly solves ditch pattern recognition of heterogeneous adjacency relations. Ignoring differences in multi-level relations would lead to an incomplete extraction of ditch
pattern. Therefore, we propose a definition of multi-level relations based on the Gestalt principle and human visual perception. Then, the algorithm to detecting various relations was designed based on their different definitions. Furthermore, the connection of these relations was considered and makes the method orderly to detect the complex pattern of ditches.

3. Complex Structure of Ditches

To clearly describe the complex structure of ditches, relations between ditches and rivers would be defined based on human visual perception [15,27]. Ditches have a good sense of integrity, regularity, and balance as man-made drainages. Parallel relations are the most common spatial patterns of ditches. However, breaks divide one complete ditch into several ditch segments, transforming single parallel relations (1-1 problem) into complex relations (m-n problem). Hence, several ditch segments should be combined based on the Gestalt principles, such as proximity, similarity, common region, continuity, and common orientation [10,20,35,36]. Moreover, these breaks also destroyed the connection between ditches and rivers. It is required to define another new relation to describe the relation of this connection. Therefore, we propose a novel model of complex ditch structure to enrich the ditch data before map generalization (Figure 2).

The model incorporated three relations. (1) Collinear relation: it was defined to extract a complete ditch from ditch segments, converting the m-n problem into the 1-1 problem. (2) Parallel relation: parallel distribution is a typical pattern of ditches and easily perceived by human visual perception. (3) Main-tributary relation. Ditches have a close affiliation with rivers to drain water considering the function of ditches. Additionally, unstructured ditches have a main-tributary relation with rivers. The concept of an unstructured ditch has not been properly explained. Hence, this paper focuses on the main-tributary relation between parallel ditches and rivers.

Figure 2. The model of complex ditch structure.
3.1. Collinear Relation

Collinear relation ditches were composed of several ditch segments, which were approximately aligned straight in their common orientation. These ditch segments were arranged closely in the same direction with their orientation. Figure 3a illustrates an example of the collinear relation and highlights three red line segments: \( d_2, d_3, \) and \( d_4 \). These segments are straight and arranged closely along the same orientation and direction. They can be easily perceived as a whole and be formed as a collinear relation due to the Gestalt principles such as proximity, similarity, and common orientation and direction. However, \( d_5 \) and \( d_6 \) can hardly be considered as a whole because \( d_6 \) is not a straight line. Moreover, the distance between \( d_7 \) and \( d_8 \) is large. Therefore, they are difficult to be perceived as a whole.

![Figure 3. Three relations of ditch group: (a) collinear relation group, (b) parallel relation group, and (c) main-tributary relation group.](image)

3.2. Parallel Relation

Parallel relation ditches were composed of several equal-length ditches (or collinear group), which were approximately aligned straight with almost equal intervals in the common orientation [34]. Furthermore, they were arranged in the vertical direction of ditch orientation. Figure 3b shows an example of parallel relation. We considered \( \{d_2, d_3, d_4\} \) and \( \{d_7, d_8\} \) as a whole based on the perception of collinear relation in Section 3.1. Thereafter, \( d_1, \{d_2, d_3, d_4\}, d_6, \{d_7, d_8\}, \) and \( d_9 \) were assembled into a set of parallel relations. Specifically, \( d_5 \) was ruled out due to its short length. Moreover, \( d_{10} \) was removed owing to its different orientation and arrangement directions.

3.3. Main-Tributary Relation

Main-tributary relations are composed of parallel relation ditches and river segment. Specifically, the starting point or end point of a ditch is close to the river segment, while the orientation of the ditch is subvertical to the river segment. Figure 3c shows an example of the main-tributary relation. Ditches are situated within the mesh surrounded by rivers and main roads. Furthermore, \( d_1, \{d_2, d_3, d_4\}, \{d_5, d_6\}, \) and \( d_7 \) represent a set of parallel relations. We considered \( D_2 \) to be the main river segment due to its close arrangement with end points of parallel ditches and their subvertical orientation of ditch, and ruled out \( D_1 \). Furthermore, we did not consider \( d_8 \) and \( d_9 \) because they were not straight lines and far from the river segment.

4. Methodology

4.1. Basic Idea of Complex Pattern Recognition

Two steps were implanted to realize the complex pattern recognition of ditches, as shown in Figure 4.
The first part is called ditch grouping. Roads and rivers have a strong separation effect on ditches. The region was divided into several meshes by the river network and main road network. Ditches were grouped by mesh based on containing relation. Therefore, the meshes surrounded by main roads and rivers were regarded as grouping units of ditches; ditches in the same mesh were treated as a group. Figure 5a shows that the mesh is constructed by main roads \((R_1, R_2)\) and river segments \((D_1, D_2)\). Moreover, ditches \((d_1-d_{15})\) were formed into a group.
Figure 5. Process of complex pattern recognition: (a) original ditches, (b) collinear relation detection, (c) parallel relation detection, and (d) main tributary relation detection.

The second part involves identifying three relations and building the model to describe complex patterns based on three relations, which are highlighted as below.

- Firstly, collinear relation will be recognized in Section 4.2. $d_6$-$d_8$ were considered as collinear ditches (red lines in Figure 5b) by the collinear relation detection rules discussed in Section 4.2. A new line ($d_{16}$ in Figure 5c), from the starting point of the first line to end point of the last line in the collinear group arrangement, is used to represent one collinear group. Similarly, $d_{17}$ represents the collinear group $\{d_9, d_{10}\}$.

- Furthermore, ditches (or collinear groups) are treated as basic units to recognize parallel relation in Section 4.3. $\{d_1$-$d_4\}$, $\{d_5, d_{16}, d_{17}, d_{11}\}$, and $\{d_{12}$-$d_{14}\}$ represent three parallel ditches groups based on the parallel relation rules discussed in Section 4.3. However, we eliminated parallel ditches with shorter average lengths due to the alternate arrangements of other parallel groups. Therefore, $\{d_{12}$-$d_{14}\}$ was not recognized as a parallel group (Figure 5c).

- Finally, following the rules in Section 4.4, the main-tributary relation was identified between parallel ditches and river segments. Figure 5d shows that parallel group $\{d_1$-$d_4\}$ follows river segment $D_1$, and parallel group $\{d_5, d_{16}, d_{17}, d_{11}\}$ follows river segment $D_2$.

Through the sequential detection of collinear, parallel, and main-tributary relation, the complex pattern of ditches was finally identified.

### 4.2. Recognizing Collinear Relation

Collinear relation ditches were composed of several ditch segments, which were separate straight lines. These segments had the same orientations. Furthermore, their direction of arrangement was the same as their orientation. Moreover, they were arranged closely in the form of a straight line without breaks.

Therefore, the following five parameters were proposed to identify the collinear relation based on orientation, sinuosity, arrangement direction, projection length, and distance.
1. The angle of ditch segment, $\theta$: It is the orientation of the line between the start point and end point of the ditch. The value of $\theta$ was the angle between the $x$-axis and the ditch $d_1$ in Figure 6a. It ranges from $0^\circ$ to $180^\circ$.

![Figure 6. Parameters of detecting collinear relation: (a) angle between the $x$-axis (dotted line) and the ditch, (b) Sinuosity, the red line connecting the start point and end point of the ditch, (c) Arrangement difference, (d) Projection length between ditches, the red line represents the projection of $d_1$ to $d_2$, (e) Ditch distance, the red dotted line is the nearest line between $d_1$ and $d_2$.](image)

2. The sinuosity of ditch segment, $\alpha$: It is the sum of the weighted angle values of each part in the ditch segment. The weight was the ratio of the length of each part to the total length of the ditch segment. It was calculated as follows:

$$\alpha = \left| \frac{\sum d_i \theta(i)}{\sum d_i} - \theta \right|$$

where $d_i$ is the length of the $i$th part of the ditch segment, $\theta(i)$ represents the orientation of the $i$th part of the ditch segment, and $\theta$ denotes the angle of the line connected by the start point and end point of the ditch (see the red line in Figure 6b). The greater the $\alpha$ is, the more tortuous the ditch. Otherwise, the ditch was closer to a line.

3. The difference of arrangement, $\beta$: It is described by the angles between a candidate ditch and a line connected by nodes of two candidate ditches (see Figure 6c). It was calculated as follows:

$$\beta = \min(\beta_1 - 90^\circ, \beta_2 - 90^\circ, \beta_3 - 90^\circ) \in [0^\circ, 180^\circ].$$

The arrangement of the ditches was closer to collinear relation when $\beta$ was closer to $90$.

4. Projection length, $l_{prj}$: It represents the projection length of one ditch to another. Figure 6c shows that the length of red line segment is the projection length of ditch from $d_2$ to $d_1$. It was calculated as follows:

$$l_{prj}(d_1, d_2) = \begin{cases} 0, & l_1 \leq 0 \land l_2 \leq 0 \\ l_1 - l_2, & 0 \leq l_2 \leq l_1 \land l_1 - l_2 \leq d_1 - l_1 \\ l_1, & l_1 - l_2 \leq 0 \land l_1 \leq d_1 \\ l_1 - l_2, & 0 \leq l_2 \leq l_1 \land l_1 - l_2 \leq d_1 \leq l_1 \\ l_1, & l_1 - l_2 \leq 0 \land l_1 \leq d_1 \end{cases}$$

$$l_1 = v_1 \cos \beta_1$$

$$l_2 = v_2 \cos \beta_2$$
where $\beta_1$ represents the angle between $v_1$ and $d_1$, and $\beta_2$ represents the included angle between $v_2$ and $d_1$. When the value of $l_{prj}(d_1, d_2)$ is smaller, the more the arrangement of the two ditches deviates from the vertical orientation of $d_1$.

5. Distance between ditches, $\lambda$: It is the minimal distance between adjacent ditches in the candidate group, i.e., the length of the red dotted line between $d_1$ and $d_2$ in Figure 6e. The range of $\lambda$ was $(0, +\infty)$. The value of $\lambda$ close to 0 implies that these ditch segments are closely aligned, similar to a line.

Each parameter threshold was determined to identify the collinear relation. We set the identification rules as follows: (1) $\min(180 - |\theta(i) - \theta(j)|, 180 - |\theta(i) - \theta(j)|) \leq \theta_0$, (2) $a \leq \alpha_0$, (3) $|\beta-90| \leq \beta_0$, (4) $l_{prj} \leq l_0$ and (5) $\lambda \leq \lambda_0$. The ditches were identified as collinear relations only when all these constraints were met.

The threshold was set separately for each parameter to control the difference in ditch segments. For example, $\theta_0$ controlled the orientation difference of ditch segments to ensure that they extended in the same direction. It was the upper limit of the orientation difference. The smaller the $\theta_0$, the more similar the orientation of ditches. However, it was meaningless to set a very small value, e.g., zero, to extract collinear relation. Furthermore, $\alpha_0$ ensured that the candidate ditch segment remained a straight line. We did not consider the tortuous ditches. Additionally, $\beta_0$ and $l_0$ ensured that the direction of arrangement remained the same as the orientation of ditch segments. Hence, the tolerable direction difference was greater if the value of $\beta_0$ was bigger. We set $l_0$ very small (strictly zero; however, if there is a computational error, it should be set close to zero) to ensure that they were arranged in the direction of orientation. Finally, $\lambda_0$ was used to ensure that they were closely arranged. It prevented the ditches with loose arrangements from being identified as a collinear relation.

4.3. Recognizing Parallel Relation

Parallel relation ditches had similar lengths and orientations. They were arranged in the direction of vertical orientation. Particularly, the collinear group was considered a whole to identify parallel relations. We proposed three parameters to extract parallel relation based on the similarity of the length, orientation, and arrangement direction:

1. The angle of ditch, $\theta$: It denotes the orientation of the ditch. The smaller the difference of $\theta$ between the ditches, the higher the orientational similarity is and the more parallelly arranged the ditches.

2. Length similarity coefficient, $\delta$: This parameter describes the degree of deviation between ditch length and the average length of the candidate parallel ditch group. It was proposed by Zahn [37], as follows:

\[
\delta = \frac{\sqrt{n} \cdot |d_i - \bar{d}|}{\sqrt{\sum_{i=0}^{n} (d_i - \bar{d})^2}}.
\]  

(6)

where $d_i$ denotes the length of ditch and $\bar{d}$ represents the average length of the parallel candidate ditch group. When $\delta_i$ is smaller, the difference of ditch length in the group is smaller.

1. Directional projection ratio, $\mu$: It is the ratio of the projection length of $d_i$ to $d_j$ to the length of the ditch $d_j$. It was calculated as follows:

\[
\mu(d_i, d_j) = \frac{l_{prj}(d_i, d_j)}{d_i}.
\]  

(7)

Its value ranges from 0 to 1. Furthermore, $l_{prj}(d_i, d_j)$ was calculated as shown in Equation (3). The higher the $\mu$ is, the more perpendicular the ditch arrangement direction is to the vertical orientation, and the more consistent the parallel relation arrangement.
The following rules were set to identify parallel relation of ditches: (1) \( \min|180-\theta(i)-\theta(j)|,|\theta(i)-\theta(j)| \leq \theta_0 \), (2) \( \delta_i \leq \delta_0 \), and (3) \( \mu(d_i, d_j) \geq \mu_0 \). The ditches were identified as parallel relations when all these constraints were met.

In the rules mentioned above, \( \theta_0 \) denotes the upper limit of orientation difference of ditches. Furthermore, \( \delta_0 \) controlled the length difference of candidate parallel ditches. More ditches of various lengths were classified into parallel groups with increasing \( \delta_0 \). We found that it was effective to set it close to 3.0 to extract parallel groups. Moreover, \( \mu_0 \) ensured that candidate parallel ditches were arranged in the same direction. The smaller difference in the arrangement direction was allowed with increasing \( \mu_0 \). Therefore, \( \mu_0 \) was set greater than 0.5.

4.4. Recognizing Main-Tributary Relation

Two parameters were used to identify the main-tributary relation between a river and ditches based on the proximity and vertical arrangement between the river segment and its parallel ditch group.

1. The nearest distance, \( d(R_i, d_j) \): It shows the minimum distance between the river segment \( R_i \) and the ditch \( d_j \).
2. Angle difference between ditch and river, \( A(R_i, d_j) \): It represents the absolute value of angle difference between the river segment \( R_i \) and the ditch \( d_j \).

We set the following recognition rules of the main-tributary relation: (1) \( d(R_i, d_j) \leq d_0 \) and (2) \( |A(R_i, d_j) - 90| \leq A_0 \).

\( d_0 \) was estimated by visual perception; it was usually set less than 100 m. \( A_0 \) was set less than 30°, considering that the angle of parallel groups and river segments is a large acute angle.

5. Experiment and Discussion

5.1. Data

The Netherlands has a large and complex man-made drainage system; its ditches form a typical agricultural landscape. The pattern of ditches should be identified first to enrich the ditch data and furtherly serves for automatic generalization. Therefore, we used the topographic datasets at the scale of 1:10k from Noordoostpolder in the Netherlands. There are ditches, rivers, and roads (see Figure 7). Different ditches were also represented by center points in gray. We observed that many ditches were surrounded by rivers and roads. Moreover, they were discontinuous due to many breaks.

5.2. Results

The ditches were grouped by a river and main road meshes after data preprocessing. A total of 111 ditch groups were achieved by ditch grouping. After that, the structure recognition was performed for each ditch group, and corresponding parameters were determined after constant debugging. The first structure was collinear relation. We set \( \theta_0 = 10^\circ, \alpha_0 = 15^\circ, \phi_0 = 5^\circ, l_0 = 0.01, \) and \( \lambda_0 = 100 \) m. Collinear ditches were identified and represented in red lines (Figure 8a). Thereafter, we set \( \delta_0 = 3.0 \) and \( \mu_0 = 0.7 \); all parallel relations were identified, which were connected by center points of parallel ditches (Figure 8b). Furtherly, parallel ditches with shorter average lengths were cut to eliminate the alternation between different parallel groups (Figure 8c). Finally, we set \( d_0 = 50 \) m, \( A_0 = 30^\circ \). Main-tributary relations were identified between parallel ditch groups and river segments, which were represented by an arrow from the parallel ditch region to their main river segment (Figure 8d).
Figure 7. The experimental data.

5.2. Results

The ditches were grouped by a river and main road meshes after data preprocessing. A total of 111 ditch groups were achieved by ditch grouping. After that, the structure recognition was performed for each ditch group, and corresponding parameters were determined after constant debugging. The first structure was collinear relation. We set $\theta_0 = 10^\circ$, $\alpha_0 = 15^\circ$, $\beta_0 = 5^\circ$, $l_0 = 0.01$, and $\lambda_0 = 100$ m. Collinear ditches were identified and represented in red lines (Figure 8a). Thereafter, we set $\delta_0 = 3.0$ and $\mu_0 = 0.7$; all parallel relations were identified, which were connected by center points of parallel ditches (Figure 8b).

Furtherly, parallel ditches with shorter average lengths were cut to eliminate the alternation between different parallel groups (Figure 8c). Finally, we set $d_0 = 50$ m, $A_0 = 30^\circ$. Main-tributary relations were identified between parallel ditch groups and river segments, which were represented by an arrow from the parallel ditch region to their main river segment (Figure 8d).

We implemented two other experiments to better demonstrate the algorithms: extracting parallel patterns by parallel factors (Test 1) [33] and extracting parallel ditches by perceptual distance (Test 2) [34]. In Test 1, we set the parallel factor to be less than 30. The tolerance of angles between ditches was set to be $30^\circ$. In Test 2, the tolerance of perceptual distance was set to 400 m. Additionally, we set $\delta_0 = 3.0$ and adopted an independent projection constraint of $\mu_0 = 0.7$, which was the same as the setting of $\mu_0$ in this paper. The results of detecting parallel ditches are shown in Figures 9 and 10.

The results had much circuitous places extracted by the parallel factor (Figure 9a) and many breaks extracted by perceptual distance (Figure 9b) compared to parallel results extracted by our method. It is because their methods did not consider the m-n problem when detecting parallel relation. Therefore, collinear relation and parallel relation were mixed together in Test 1. Moreover, short ditches were considered into the parallel group due to lack of length-similarity constraint (see Figure 10b,e). The length-similarity constraint and projected ratio were considered in Test 2. However, short ditch segments with a low projected ratio made it difficult to detect parallel groups, leading to several breaks in Figure 10c,f. Moreover, the m-n problem became the big barrier while applying these methods. Nevertheless, the proposed method first detected collinear relation to combine “many” ditch segments into “one” complete ditch. The parallel results extracted by our method were better than their results because of converting the m-n problem into the 1-1 problem.
We detected 325 collinear ditch groups in Figure 8a. There were 132 parallel ditch groups in Figure 8b and 54 parallel groups after removing parallel groups in an alternate area in Figure 8c. Finally, we identified 61 groups of main-tributary relations between these parallel groups and river segments in Figure 8d.

We implemented two other experiments to better demonstrate the algorithms: extracting parallel patterns by parallel factors (Test 1) [33] and extracting parallel ditches by perceptual distance (Test 2) [34]. In Test 1, we set the parallel factor to be less than 30. The tolerance of angles between ditches was set to be 30°. In Test 2, the tolerance of perceptual distance was set to 400 m. Additionally, we set $\delta_0 = 3.0$ and adopted an independent projection constraint of $\mu_0 = 0.7$, which was the same as the setting of $\mu_0$ in this paper. The results of detecting parallel ditches are shown in Figures 9 and 10.

Furthermore, considering the effect of the breaks, we furtherly identified parallel relation based on the result of extracting the collinear relation in Test 1 and 2. The parallel extraction results in Figure 11 show a more coherent parallel relation (red lines) than the results shown in Figure 9. This indicates that the extraction of collinear relation effectively improves the quality of extracting a parallel relation. However, the difference in ditch length reduced the quality of parallel relation extraction (Test 1), as shown in Figure 12c. Length difference was considered as one of the important factors in Test 2. Therefore, perceptual distance divided different length ditches into different groups, as shown in Figure 12b,d. However, the perceptual distance was a comprehensive parameter combined with angle, length, distance, and projection length. The difference of one factor enhances the overall difference of ditches, which is highly sensitive to length difference and distance difference, as shown in Figure 12b,d, respectively. Finally, these parallel groups should be removed due to the limited importance of parallel ditches with short lengths.
one factor enhances the parallel groups, leading to several breaks.

Figure 9. (a) Extracted parallel ditches in Test 1, (b) Extracted parallel ditches in Test 2.

Figure 10. Detailed comparisons of extracted parallel results: (a,d) the results of our method, (b,e) Test 1 results, (c,f) Test 2 results.

Figure 11. Extracted parallel ditches after extracting collinear relation: (a) Test 1 results, (b) Test 2 results.
5.3. Evaluation

Two cartography experts and six professional cartographers were invited to visually detect the relations (collinear, parallel, and main-tributary) for evaluating the quality of the extracted results. Each relation was identified by at least three cartographers. Firstly, the cartographers identified the targeted relation. Thereafter, their identifying results would be checked by one of the experts. The expert would determine the final result of each relation based on the results of these cartographers. After detecting three different relations, the manual results of identifying complex patterns were completed. These manual recognition results are shown in Figure 13.

There were 338 collinear groups, 60 parallel groups, and 65 main-tributary groups (see Figure 11). We took the results of manual recognition as the standard and chose precision, recall, and F1-score to evaluate the effect of the methods (Table 1). The precision is defined as the ratio of the number of correctly extracted relations to the total number of extracted relations. Furthermore, the recall is defined as the ratio of the number of correctly extracted relations to the number of visually identified relations. Moreover, the F1-score is defined as the ratio of twice the product of precision and recall to the sum of precision and recall. All these metrics mentioned above ranged from 0 to 1.0. Meanwhile, the closer the value...
of metrics is to 1.0, the higher the similarity between the experimental group and the standard group.

Table 1. The performance evaluation metrics for pattern extraction.

| Relation Type        | Precision | Recall | F1-Score |
|----------------------|-----------|--------|----------|
| Collinear            | 0.968     | 0.947  | 0.956    |
| Parallel             | 0.907     | 0.996  | 0.949    |
| Parallel (Test 1)    | 0.680     | 0.024  | 0.046    |
| Parallel (Test 2)    | 0.924     | 0.558  | 0.696    |
| Main-tributary       | 0.950     | 0.905  | 0.927    |

Each evaluated metric of extracting different relations by our method was more than 0.90, implying that our method can recognize the complex structure of ditches satisfactorily. Moreover, we compared parallel recognition results of Tests 1 and 2 with manual results based on the extracted result of collinear relation. Three metrics of Test 1 results are all below 0.7, indicating poor performance in extracting parallel relations. Although the precision of Test 2 results is better than the results of Test 1 and our method, the recall and F1-score of the Test 2 result are both below 0.7, implying unrobust performance of extracting parallel extraction results. Nevertheless, all metrics of our method are greater than 0.9, implying effective and robust parallel extraction results. Furthermore, some differences between manual and automatically extracted results should be concerned in Figure 14.

Detailed collinear results show that ditches are divided into a collinear group extracted by our method due to meeting distance constraint (see green lines in Figure 14a). However, humans divided them into different groups. The threshold of distance is difficult to measure accurately. Furthermore, humans recognized geo-feature alignment in a regular way. Ditch (id 2456 in black dotted box) in Figure 14a and ditch (id 871 red line) in Figure 14b are not straight lines; they only have straight parts. However, these differences were not easily noticeable for human visual perception. Finally, these lines were identified as the collinear relation (Figure 14c) and parallel relation (Figure 14d).

5.4. Discussion

This study aims to describe the structure of complex distributed ditches by pattern recognition to enrich the data source and facilitate automatic generalization. Therefore, we propose a novel model to describe the complex pattern of ditches. The advantages of the proposed method in detecting complex pattern are as follows:

(1) The m-n problem can be solved by detecting collinear relations. It is still a big challenge in map generalization, which was studied for pattern recognition [29] and typification [38]. The model combined several discrete ditch segments as a whole if they meet the rules of collinear relation. Hence, the m-n problem was converted into the 1-1 problem, significantly improving the quality of subsequent pattern recognition.

(2) The quality of detecting parallel relation by multi-characteristic constraints was improved. Angle difference, length similarity coefficient, and directional projection ratio were used to extract parallel relation based on proximity and similarity in the Gestalt principle. The results were consistent with human visual perception.

(3) The method established an affiliated connection between ditch and river, which could manage the linkage effect in river generalization. The main topology and distribution characteristics of the ditch were identified (this problem was mentioned but not dealt with [39]), which could be easily preserved in map generalization for different map scales.

Additionally, some issues were further studied or improved. First, the key parameter thresholds would have a great influence on the results of detecting patterns. We conducted more experiments using different thresholds to better demonstrate the proposed method. Moreover, some examples of main-tributary relations were presented, which should be
furtherly studied because of their complexity. Finally, the process of generalizing complex ditches was discussed based on the proposed model.

Figure 14. Detailed comparisons of extracted results by human and our method: (a) collinear relations extracted by our method, (c) collinear relations extracted by human, (b) parallel relations extracted by our method, (d) parallel relations extracted by human visual perception.

5.4.1. Impact of Parameter Threshold

In collinear relation recognition, the parameter $\lambda_0$ controlled the looseness of collinear ditch arrangement, which had a great influence on collinear extracted results (Figure 15). We set $\lambda_0 = \{10 \text{ m}, 70 \text{ m}, 100 \text{ m}, 120 \text{ m}\}$ to detect collinear relation. The extracted results showed that the arrangement of collinear ditches is closer when $\lambda_0$ is smaller. However, it does not mean the smaller value provides a better result. Figure 15a shows that the number of collinear relations is less than manual results (Figure 14c). It is the same with manual results when $\lambda = 70 \text{ m}$ (Figure 15b). The best working range of $\lambda_0$ for our method was 20–100 m.
In parallel relation recognition, angle difference $\theta_0$ and directional projection ratio $\mu_0$ can have a serious impact on extracted results. We set $\theta_0 = \{10^\circ, 15^\circ, 20^\circ\}$ and $\mu_0 = \{0.5, 0.7, 0.9\}$ to get nine different results (Figure 16). We contrasted the original parallel results to reflect the differences of parameter changes on parallel extraction clearly. The left ditch (id 295) was included in the parallel group by changing $\theta_0$ from $10^\circ$ to $15^\circ$. After that, the extracted results did not change even though the angle difference was changed to $20^\circ$.

Parallel groups composed of short ditches changed when $\theta_0$ remained fixed and $\mu_0$ changed. Similar extracted results can be seen in Figure 17. The number of ditches in a parallel group reduced from 10 to 7 when $\mu_0$ changed from 0.5 to 0.9. More parallel groups would be eliminated with decreasing $\theta_0$ and increasing $\mu_0$. We recommend the range of $\theta_0$ from $10^\circ$ to $20^\circ$ and $\mu_0$ from 0.5 to 0.7.
5.4.2. Some Special Examples of Main-Tributary Relation

The main-tributary relation was detected by the angle and distance between parallel ditches and river segments. However, there are special examples of main-tributary relation that should be considered. The main road crossed the ditches, resulting in ditch segments being in a different group (Figure 18a). Therefore, the proposed method could not determine their collinear relation and main-tributary relation. Furthermore, the upper drainage of the ditch was not necessarily a river but possibly ditches (Figure 18b). The ditch (id 1003) may be the upper drainage of the left ditches (red lines). Therefore, the topology and distribution characteristics between ditches and rivers should be further analyzed.

5.4.3. Further Study Based on the Model of Complex Pattern

This paper proposes a novel model to describe the complex pattern of ditches. Moreover, it is important to use this model for the generalization of ditches. The current studies of typification for parallel ditches can give us some ideas. Sandro [9] built an envelope called a “canvas” of parallel ditches. Furthermore, Tian [7] calculated the number and the position of parallel ditches after typification based on the K-means algorithm. We can design the method to make these algorithms more suitable for the complex pattern of ditches (see Figure 19). After identifying the complex pattern of ditches, collinear ditches should be merged into complete lines, as shown in Figure 19b. After that, the envelopes of parallel ditches can be drawn (see green area in Figure 19c). The number and position of the center points of the new ditches can be calculated by the K-means algorithm. The new ditches would be the extension lines from center points to the boundary of the canvas at the average direction of parallel ditches (see dotted lines in Figure 19e). Finally, the main-
tributary relation between the parallel group and rivers should be preserved, and new ditches should be extended to the river lines, as shown in the red dotted box of Figure 19f. Moreover, ditch density differences in different regions and the degree of typification [38] should be studied further.

Figure 19. Generalization for the complex pattern of ditches: (a) the original data, (b) the result of merging collinear ditches after identifying a complex pattern, (c) canvas of parallel ditches (green area), (d) center points of parallel ditches, and (e) the output of the K-means algorithm. The dotted lines show the extension lines from center points to the boundary of the canvas in the average direction of parallel ditches. (f) The result after preserving the main-tributary relation between parallel ditches and river, which is shown in the red dotted box.

6. Conclusions

Pattern recognition is an important issue for data enrichment and map generalization. This paper demonstrated a model to describe the complex pattern of mix-distributed ditches. Three relations—collinear, parallel, and main-tributary—were defined to describe the complex structure of ditches. Collinear relation combined ditch segments via the principle of integrity, maintaining the continuity of ditch. Parallel relation—the main spatial pattern in farmland—was used to combine discrete similar parallel ditches into a group. Main-tributary relation repaired the connection between parallel ditches and rivers. The three relations describe complex ditches at different levels. Moreover, a complex pattern of ditches was detected from the primary level to high level, which is in the order of collinear, parallel, and main-tributary. The experimental results show that the proposed method can detect the main topology and distribution characteristics of ditches and describe their structure in an orderly manner.

Further research will focus on improving the main-tributary relation detection method. Furthermore, different strategies for ditch generalization based on the proposed complex structure will be studied.

Author Contributions: Conceptualization, Chengyi Liu and Fang Wu; methodology, Chengyi Liu; software, Chengyi Liu; validation, Chengyi Liu, Fang Wu and Ruixing Xing; formal analysis, Chengyi Liu, Jiawei Du and Ruixing Xing; investigation, Chengyi Liu and Ruixing Xing; resources, Xianyong Gong; data curation Chengyi Liu and Ruixing Xing; writing—original draft preparation, Chengyi Liu; writing—review and editing, Fang Wu, Xianyong Gong and Jiawei Du; visualization, Chengyi Liu; supervision, Jiawei Du; project administration, Fang Wu and Xianyong Gong; funding acquisition, Fang Wu and Xianyong Gong. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41801396, 41471386.
References

1. ICA (International Cartographic Association). Multilingual Dictionary of Technical Terms in Cartography; Steiner: Wiesbaden, Germany, 1973.

2. Powitz, B. Computer-assisted generalization—An important software tool for gis. Int. Arch. Photogramm. Remote Sens. 1993, 29, 664–673.

3. Patrick, T.; Julien, G. Improving map generalization with new pruning heuristics. Int. J. Geogr. Inf. Sci. 2012, 26, 1309–1323. [CrossRef]

4. Jantien, S.; Marc, P.; Vincent, V. Fully automated generalization of a 1:50k map from 1:10k data. Cartogr. Geogr. Inf. Sci. 2014, 41, 1–13. [CrossRef]

5. Touya, G.; Duchêne, C.; Taillandier, P. Multi-Agents Systems for Cartographic Generalization: Feedback from Past and On-Going Research. Technical Report; IGN (Institut National de l’Information Géographique et Forestière); LaSTIG, Équipe COGIT. hal-01682131. 2018. Available online: https://hal.archives-ouvertes.fr/hal-01682131/document (accessed on 6 October 2020).

6. Steiniger, S.; Weibel, R. Relations among Map Objects in Cartographic Generalization. Cartogr. Geogr. Inf. Sci. 2007, 34, 175–197. [CrossRef]

7. Tian, J.; Yang, W.; Chen, L. Combining Two Clustering Ideas for Typification of Ditches. In Geoinformation for Informed Decisions, Lecture Notes in Geoinformation and Cartography; Abdul Rahman, A., Boguslawski, P., Anton, F., Nor Said, M., Mohd Omar, K., Eds.; Springer: Cham, Switzerland, 2014; pp. 173–178. [CrossRef]

8. Zhang, X.; Ai, T.; Stoter, J. Characterization and detection of building patterns in cartographic data: Two algorithms. Int. Arch. Photogrammet Remote Sens. Spat. Inf. Sci. 2010, 38, 261–266. [CrossRef]

9. Zhu, L.; Jiang, C.; Xie, X.; Xie, M. Ecological and Landscape Effects of Ditches on Farmland Ecosystem. In Proceedings of the 2009 3rd International Conference on Bioinformatics and Biomedical Engineering, Beijing, China, 11–13 June 2009; pp. 1–4. [CrossRef]

10. Sandro, S.; Massimo, R.; Matteo, Z. Pattern recognition and typification of ditches. In Advances in Cartography and Giscience, Lecture Notes in Geoinformation and Cartography; Ruas, A., Ed.; Springer: Heidelberg, Germany, 2011; pp. 425–437. [CrossRef]

11. He, X.; Deng, M.; Luo, G. Recognizing Linear Building Patterns in Topographic Data by Using Two New Indices Based on Delaunay Triangulation. ISPRS Int. J. Geo-Inf. 2020, 9, 231. [CrossRef]

12. Horton, R. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Geol. Soc. Am. Bull. 1945, 56, 275–370. [CrossRef]

13. Marshall, S. Streets and Patterns; Spon Press: New York, NY, USA, 2014; p. 318.

14. Heinzle, F.; Anders, K. Characterising space via pattern recognition techniques: Identifying patterns in road networks. In The Generalisation of Geographic Information: Models and Applications; Mackaness, W., Ruas, A., Sarjakoski, T., Eds.; Elsevier: Amsterdam, The Netherlands, 2007; pp. 233–254. [CrossRef]

15. Wertheimer, M. Laws of Organization in Perceptual Forms. In A Source Book of GESTALT Psychology; Ellis, W.D., Ed.; Harcourt, Brace: New York, NY, USA, 1938; pp. 71–88, Original Work Published in 1923; Available online: https://insights.ovid.com/psycritiques/psyc/196808/000/source-book-gestalt-psychology/25/01258377 (accessed on 12 January 2021).

16. Freundschuh, S.M. Is there a relationship between spatial cognition and environmental patterns? In Proceedings of the Theories & Methods of Spatio-Temporal Reasoning in Geographic Space, International Conference GIS-from Space to Territory: Theories & Methods of Spatio-Temporal Reasoning, Pisa, Italy, 21–23 September 1992. [CrossRef]

17. Steiniger, S.; Lange, T.; Burghardt, D. An approach for the classification of urban building structures based on discriminant analysis techniques. Trans. GIS 2008, 12, 31–59. [CrossRef]

18. Meinel, G.; Hecht, R.; Herold, H. Analyzing building stock using topographic maps and GIS. Build. Res. Inf. 2009, 37, 468–482. [CrossRef]

19. Regnauld, N. Contextual building typification in automated map generalization. Algorithmica 2001, 30, 312–333. [CrossRef]

20. Heinzle, F.; Anders, K.H.; Sester, M. Pattern recognition in road networks on the example of circular road detection. In Geographic Information Science; Raubal, M., Miller, H.J., Frank, A.U., Goodchild, M.F., Eds.; Springer: Berlin, Germany, 2006; Volume 4197, pp. 153–167. [CrossRef]

21. Touya, G. A Road Network Selection Process Based on Data Enrichment and Structure Detection. Trans. GIS 2010, 14, 595–614. [CrossRef]

22. Strahler, A. Quantitative analysis of watershed geomorphology. Trans. Am. Geophys. Union 1957, 38, 913–920. [CrossRef]

23. Shreve, R.L. Statistical law of stream numbers. J. Geol. 1966, 74, 17–37. [CrossRef]
24. Simon, D.; Gerald, C. The morphology and formation of floodplain-surface channels, Cooper Creek, Australia. *Geomorphology* **2004**, *60*, 107–126. [CrossRef]

25. Zhang, L.; Guilbert, E. Automatic drainage pattern recognition in river networks. *Int. J. Geogr. Inf. Sci.* **2013**, *27*, 2319–2342. [CrossRef]

26. Jung, K.; Shin, J.; Park, D. A new approach for river network classification based on the beta distribution of tributary junction angles. *J. Hydrol.* **2019**, *572*, 66–74. [CrossRef]

27. Field, D.J.; Hayes, A.; Hess, R.F. Contour integration by the human visual system: Evidence for a local “association field”. *Vis. Res.* **1993**, *33*, 173–193. [CrossRef]

28. Liu, C.; Wu, F.; Gong, X.; Xing, R.; Luo, D. Complex linear pattern recognition for regular ponds groups. *Acta Geod. Cartogr.* **2020**, *49*, 256–266. [CrossRef]

29. Xing, R.; Wu, F.; Gong, X.; Du, J.; Liu, C. An axis-matching approach to combined collinear pattern recognition for urban building groups. *Geocarto Int.* **2021**. [CrossRef]

30. Ai, T.; Guo, R. Polygon cluster pattern mining based on Gestalt principles. *Acta Geod. Cartogr. Sin.* **2007**, *36*, 302–308. [CrossRef]

31. Wei, Z.; Guo, Q.; Wang, L.; Yan, F. On the spatial distribution of buildings for map generalization. *Cartogr. Geogr. Inf. Sci.* **2018**, *45*, 539–555. [CrossRef]

32. Yang, L.; Zhang, L.; Ma, J. Interactive visualization of multi-resolution urban building models considering spatial cognition. *Int. J. Geogr. Inform. Sci.* **2011**, *25*, 5–24. [CrossRef]

33. Zhang, H.; Wu, F.; Gong, X.; Xu, J.; Zhang, J. A parallel factor-based method of arterial two-lane roads recognition. *Geomat Inf. Sci. Wuhan Univ.* **2017**, *42*, 1123–1130. [CrossRef]

34. Tian, J.; Xiong, F.; Fang, H.; Yang, W.; Chen, L. A new typification method for ditches with almost parallel distribution. *Geomat Inf. Sci. Wuhan Univ.* **2014**, *39*, 1452–1456. [CrossRef]

35. Yang, B.; Luan, X.; Li, Q. An Adaptive Method for Identifying the Spatial Patterns in Road Networks. *Comput. Environ. Urban Syst.* **2010**, *34*, 40–48. [CrossRef]

36. Gong, X.; Wu, F.; Qian, H.; Ma, K. The parameter discrimination approach to multi-connected linear pattern recognition in building groups. *Geomat Inf. Sci. Wuhan Univ.* **2014**, *39*, 335–339. [CrossRef]

37. Zahn, C. Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Trans. Comput.* **1971**, *20*, 68–86. [CrossRef]

38. Wang, X.; Burghardt, D. A mesh-based typification method for building groups with grid patterns. *Int. J. Geo-Inf.* **2019**, *8*, 168. [CrossRef]

39. Touya, G. River Network Selection based on Structure and Pattern Recognition. In Proceedings of the 23rd International Cartographic Conference, Moscow, Russia, 4–10 August 2007; pp. 4–9. [CrossRef]