Article

Semi-Automatic Extraction of Rural Roads under the Constraint of Combined Geometric and Texture Features

Hai Tan [1,*], **Zimo Shen** [2,3] **and Jiguang Dai** [2,3]

1   Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, Beijing 100048, China
2   School of Geomatics, Liaoning Technical University, Fuxin 123000, China; 472020770@stu.lntu.edu.cn (Z.S.);
    daijiguang@lntu.edu.cn (J.D.)
3   Institute of Spatiotemporal Transportation Data, Liaoning Technical University, Fuxin 123000, China
*   Correspondence: tanh@lasac.cn

Abstract: The extraction of road information from high-resolution remotely-sensed images has important application value in many fields. Rural roads have the characteristics of relatively narrow widths and diversified pavement materials; these characteristics can easily lead to problems involving the similarity of the road texture with the texture of surrounding objects and make it difficult to improve the automation of traditional high-precision road extraction methods. Based on this background, a semi-automatic rural road extraction method constrained by a combination of geometric and texture features is proposed in this paper. First, an adaptive road width extraction model is proposed to improve the accuracy of the initial road centre point. Then, aiming at the continuous change of curvature of rural roads, a tracking direction prediction model is proposed. Finally, a matching model under geometric texture constraints is proposed, which solves the problem of similarity between road and neighbourhood texture to a certain extent. The experimental results show that by selecting different types of experimental scenes or remotely sensed image data, compared with other methods, the proposed method can not only guarantee the road extraction accuracy but also improve the degree of automation to a certain extent.

1. Introduction

Road data play an important role in many fields, such as urban planning, traffic management, map updating, disaster management, road monitoring, public health, unmanned aerial vehicle (UAV)-based visual navigation, driving assistance systems, and agricultural development [1–3]. With the rapid development of remote sensing satellites and sensors, higher-resolution remote sensing image acquisition is becoming increasingly easy [4]. As remote sensing images represent basic ground feature recognition data, an increasing number of scholars have been attracted to invest in road extraction research. Rural roads play a key role in rural planning and are an important form of infrastructure for ensuring social and economic development in rural areas. In China, for example, the rural road foundation is large. According to the statistics of the white paper "sustainable development of China's transportation", China's rural roads account for 83.8% of the country's total highway mileage. In this context, the extraction of rural roads is particularly important.

Domestic and foreign researchers have proposed a large number of road extraction methods. According to the need to provide a priori samples, this paper divides these road extraction methods into model-driven methods and data-driven methods.

Model-driven is a data processing method without samples. Its advantages lie in applying models directly to data, with low manual participation and high automation. The main representatives of this type of method are knowledge-driven methods and object-oriented methods.

The knowledge-driven method constructs the knowledge model related to the road, and then establishes hypothesis test models between the knowledge model and image-processing results, so as to achieve knowledge-based road extraction [5]. Among these methods, Hedman et al. [6] extracted roads according to their corresponding linear geometric feature knowledge. Baltsavias [7] summarized the feasibility of using existing knowledge and geographic data to improve automation efficiency. Grote et al. [8] used a digital surface model (DSM) and high-resolution colour-infrared images to extract suburban roads. Although knowledge-driven methods can use existing information to improve the road extraction efficiency, determining how to adapt the existing knowledge to various road scenes is the most difficult core problem for the knowledge-driven method.

Other typical examples of model-driven methods include object-oriented methods. In an object-oriented method, a segmentation algorithm is first used to separate the road area from other areas, and then the appropriate classification and post-processing method are selected to extract the roads. For example, Lei et al. [9] used the gray consistency of the road surface and the mutability of the road edges in the gray images to segment the studied region, then used shape features to select the road region. After multi-resolution segmentation, Kumar et al. [10] used a fuzzy membership function and image object attribute value to define different classes in order to extract roads. Bakhtiari et al. [11] first used the Canny operator to detect the contours of roads, then used a support vector machine (SVM) to classify the images after segmentation. Object-oriented methods take the object region as the processing unit, lessen the deficiencies of pixel-level analyses, and improve the spatial smoothness of the road extraction effect [12]. However, the image segmentation results obtained with these methods are greatly affected by the image quality. In cases involving texture similarity, shadows, or occlusion, it is easy to obtain segmentation results that are inconsistent with the actual features, and the road extraction accuracy following scene transformations is difficult to ensure.

Some differences exist in road image data among different sources. Rural road data, in particular, contain many road materials and curvature continuity changes. Model-driven methods are based on road features, and it is difficult to analyse these differences on the basis of limited parameter analyses. These problems make it difficult to ensure the accuracy of road extractions in model-driven methods; thus, data-driven methods have been developed.

Data-driven methods are based on the characteristics of different data. By manually selecting a priori feature samples, models are established to fit and learn the parameters of the selected features to facilitate the discrimination of roads. Deep learning and template matching are typical data-driven methods.

In the deep learning road extraction method, the model learns a prior set of road data and then identifies the roads in the test set using the discriminant function. Heermann et al. [13] proposed the back-propagation (BP) algorithm, allowing road extraction methods based on neural networks to develop rapidly. In recent years, with more applications of deep learning in the field of road extraction, globally aware road detection networks with multi-scale residual learning (GAMSNet) [14], boundary and topologically aware neural networks for road extraction (BT-RoadNet) [15], multitask road-related extraction networks (MRENet) [16], deep structured self-driving networks (DSDNet) [17], and other new networks continue to emerge. Deep learning methods provide new opportunities for the semantic expansion of remote sensing image interpretations [18]. However, the effectiveness of road extraction techniques with deep learning methods is greatly affected by the quality of the sample set, and noise and occlusion problems cause fractures in most road extraction results. It is difficult to obtain extraction accuracy and recall values over 90% simultaneously when using current deep learning methods [19,20]. In addition, no vector topology relationship exists in the extraction results, and the results thus require a large amount of postprocessing to yield product-level data.

As another data-driven method, the template matching method establishes a template by manually selecting local road samples, analyses the similarity of the parameter

information contained in the pre-selected data and template data, and selects the area with the greatest similarity to complete road extraction. Wang et al. [21] extracted initial roads through extensive contour analysis and then used the snake model to optimize the road locations. Leninisha et al. [22] improved the geometrically active deformation model and proposed an extended geometrically active deformation model with improved accuracy and efficiency. Additional, classic template matching methods include circular template [23], T-shaped template [24], sector template [25], and rectangular template [26] methods. The circular template automatically generates the initial template by morphological gradient; then it searches other road points between the starting point and the end point according to the iterative interpolation method to complete the road tracking and matching. The T-shaped template uses the angle texture feature to obtain the initial road points; then calculate the road width and road direction; finally, gray least square matching is used to locate the optimal road points. Based on the principle that the edge of the road near the ground object is the same as the road direction, the sector template proposes the multiscale line segment orientation histogram (MLSOH) descriptor [25]: the extracted information of the line segment near the road is counted, and the direction with the maximum probability is found as the dynamic tracking direction of the road; then, a sector descriptor is established to extract the centre point of the road by using the texture feature of the road. However, the applicability of these traditional template matching algorithms in complex areas, such as in regions containing obstacles or shadows, needs to be improved. Dai et al. [27] proposed a semi-automatic extraction method of rural roads with high resolution remote sensing images combined with multiple features. The MLOSH descriptor was used to calculate the road tracking direction and reduce the influence of local curvature changes on tracking; then, a multi-circle template was established; finally, the proposed panchromatic and hue, saturation and value (HSV) spatial interactive matching model was used to match and track the roads. This method solves some problems caused by the diversification of rural road materials.

On the one hand, template matching, as a data-driven method, avoids the disadvantage that it is difficult to analyse the differences of rural roads driven by the model. On the other hand, compared with the same data-driven deep learning method, template matching starts from the local part of the image and has strong human interference, strong error correction ability and high extraction accuracy (the accuracy and recall can reach more than 95%), thus meeting the requirements for practical applications. However, rural roads have the characteristics of continuous curvature changes, relatively narrow widths and diversified pavement materials, leading to a low degree of automation in template matching road extraction methods. Based on these challenges and on template matching methods, this paper proposes a semi-automatic rural road extraction method constrained by the combination of geometric features and texture features; this method not only ensures the road extraction accuracy but also improves the degree of automation. The specific contributions of this work are as follows:

(1) An adaptive road width extraction model is proposed. In the existing methods, the detection of road width is slow and the accuracy is low due to the need to set the threshold [25]. According to the good edge fitting of the extraction results of rural road line segment sequence [28], the road width is extracted by calculating the projection distance. The efficiency and accuracy of road width extraction are improved, and the quality of initial road centre point is improved.

(2) The existing direction prediction model is improved. Based on the principle of MLSOH descriptor to determine the direction, Dai et al. [28] first uses the line segment sequence with better segment fusion to replace the discrete segment. Then, the line segment sequence near the road is divided into the artificially specified angle range. Finally, the range of the maximum cumulative length of the line segment sequence is selected as the road tracking direction. However, the width of rural roads is narrow, and the road direction needs an accurate angle value. Therefore, this paper adjusts

the cumulative length of line segment sequence in [28] to the length of single line segment sequence, and obtains a more accurate and stable road direction.

(3) The proposed method solves the matching problem when the road is similar to the background. Compared with urban roads, rural roads, as low-grade roads, have the characteristics of narrower road width and diversified road materials. These easily lead to the similarity of road and background texture in the image. As a result, the traditional road extraction method has a low degree of automation on the premise of ensuring high accuracy. To solve this problem, we abandon the idea that the road matching model only relies on texture spectral features, and add geometric weights into the matching model to form a dynamic matching model incorporating geometric information. The model can solve the matching problem when the road is similar to the background by analysing the geometric and texture information of the road.

2. Materials and Methods

2.1. Experimental Data

This paper mainly uses high-resolution orthophotographic panchromatic images. The image data are selected from different remote sensing satellites, namely, Gaofen (GF)-2 and GF-7 data. Among the data, the GF-2 data represent Enshi city, Hubei Province, China and Dandong city, Liaoning Province, China, and the GF-7 data represent Zhangjiakou city, Hebei Province, China. The detailed image parameters are listed in Table 1.

Table 1. Remote sensing image information used in the experiments.

Data Catalogue	Band Type	Band Range (μm)	Resolution (m)
GF-2	Panchromatic	0.45–0.90	1
GF-7	Panchromatic	0.45–0.90	0.65

2.2. Methodology

The experimental method of this paper is divided into four parts: Section 2.2.1 introduces the pre-processing work, including line segment sequence extraction to obtain road structure features and L0 filter to improve the internal homogeneity of the road. We select manual input points after pre-processing. Section 2.2.2 introduces the adaptive road width extraction model and determines the accurate initial road centre point. Section 2.2.3 introduces the tracking direction prediction model to provide accurate tracking direction for the next matching. Section 2.2.4 introduces the rural road matching model under geometric texture constraints. If the matching is successful, the road direction prediction is carried out again to track the next road point. If it fails, the results are output. The flow-process-chart of this process is shown in Figure 1.

2.2.1. Pre-Processing

According to the unique linear characteristics of the road, we use the extraction of line segment sequence [28] to obtain the prior information of the road. As shown in Figure 2b, the line segment sequence encodes and groups the road edge segments to obtain continuous road edge information.

In addition, aiming at the problem of image noise, we use L0 filter [29]. By removing the small non-zero gradient, the unimportant details are smoothed while the significant edge of the image is enhanced (retaining the large gradient), as shown in Figure 3. This not only improves the internal homogeneity of the road, but also distinguishes the road from the surrounding features.
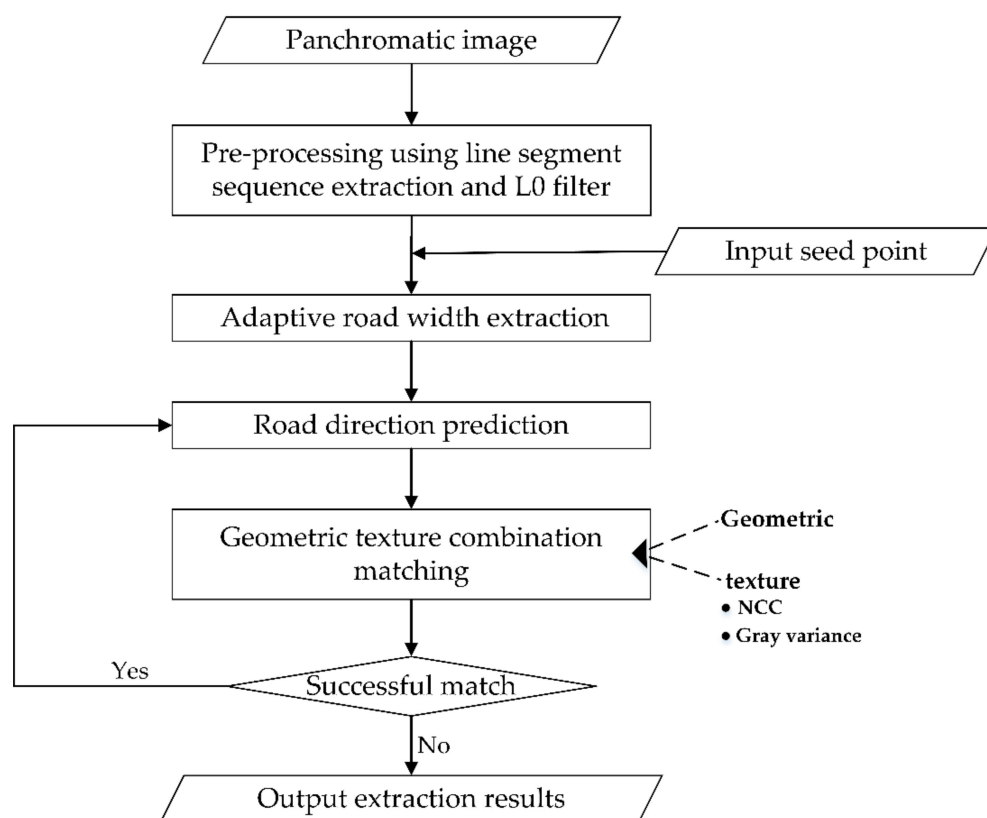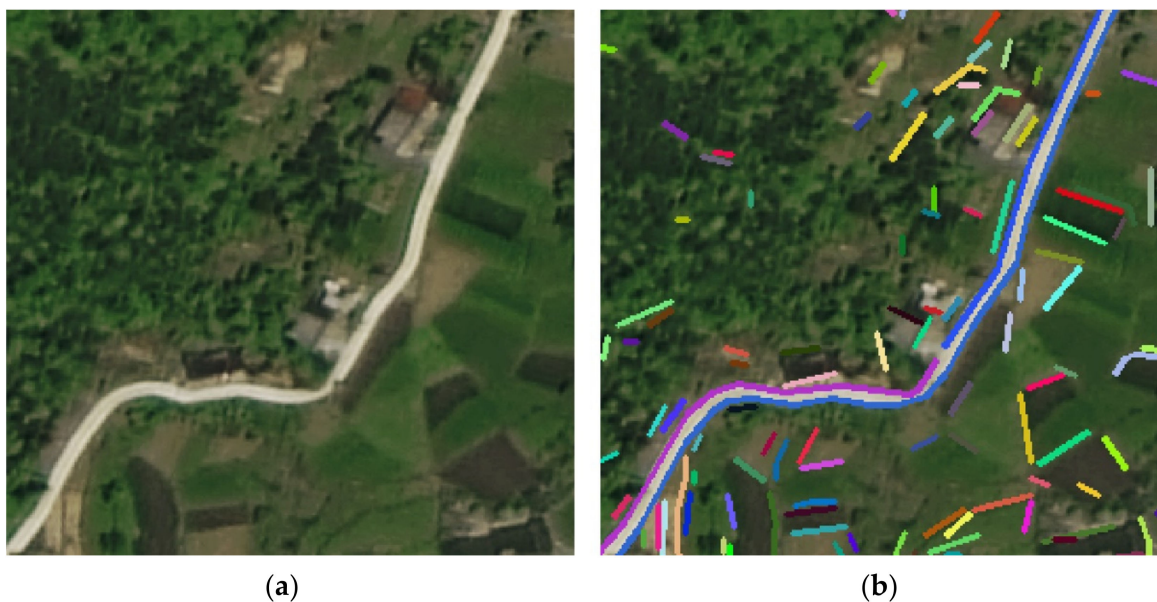
Figure 1. Flow-process chart.

Figure 2. Extraction of line segment sequence. (**a**) The original image. (**b**) Result of line segment sequence.
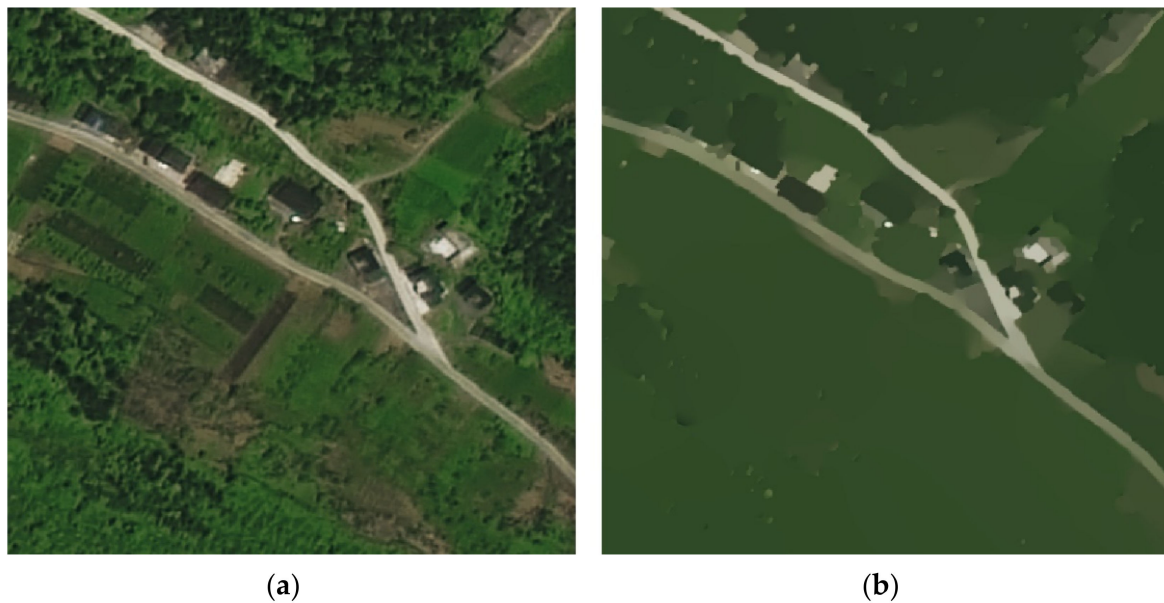
(**a**) (**b**)

Figure 3. L0 filter effect comparison. (**a**) The original image. (**b**) Image after L0 filter.

After pre-processing, an artificial point P_1 was selected at a clear position of road boundary to start road tracking and matching.

2.2.2. Adaptive Road Width Extraction Model

In this paper, an adaptive road width extraction model is designed to obtain relatively accurate initial road centre point. Figure 4 shows the overall process of the model. The specific steps are as follows:
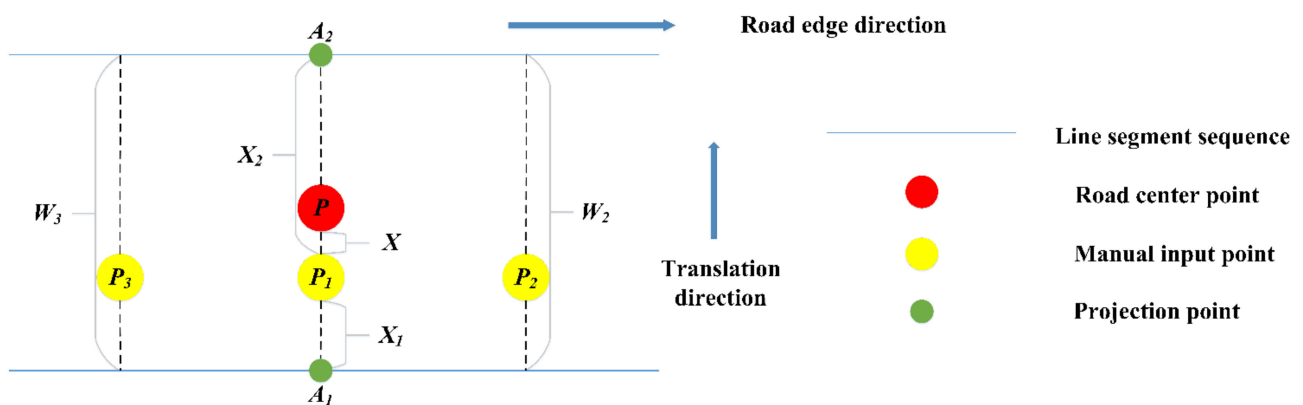
Figure 4. Adaptive road width extraction model.

(1) Based on the initial point P_1 input manually, points with 5 pixels ahead of and behind P_1 are selected along the road prediction direction to obtain a total of three points, P_1, P_2, and P_3, on the road.

(2) Starting with P_1, the projection of P_1 on the edge line segment on one side of the road is calculated to obtain the projection point A_1 and the projection distance X_1. Then, the projection of P_1 on the edge segment on the other side of the road is calculated to obtain the projection point A_2 and the projection distance X_2. The sum of X_1 and X_2 is the road width W_1. Following this method, the road widths W_2 and W_3 corresponding to P_2 and P_3, respectively, are calculated, and the average road width W among the three points is obtained.

(3) The translation direction is selected as the projection direction on the side of the road farthest away from P_1, and the translation distance X is determined, as shown in

Equation (1). The initial point P_1 is translated along the translation direction with a distance X to the road centre point P.

$$X = \frac{W}{2} - \min(X_1, X_2) \tag{1}$$

2.2.3. Tracking Direction Prediction Model

In view of the good edge fitting ability of line segment sequence, this paper adjusts the cumulative length of the line segment sequence counted in [28] to the length of single line segment sequence. The specific process is as follows:

(1)　A rectangular search box is established with the current road point as the centre and 2 times the road width as the side length.
(2)　Count the length of the line segment sequence in the rectangular box. The corresponding direction of each line segment sequence is the direction of the line segment in the rectangular box of the line segment sequence.
(3)　The direction corresponding to the longest line segment sequence is taken as the tracking direction.

Figure 5a statistics the direction distribution of line segment sequence in the search box. In Figure 5b: the yellow point is the tracked road centre point. The green point is the current road point. The blue box is the search box.
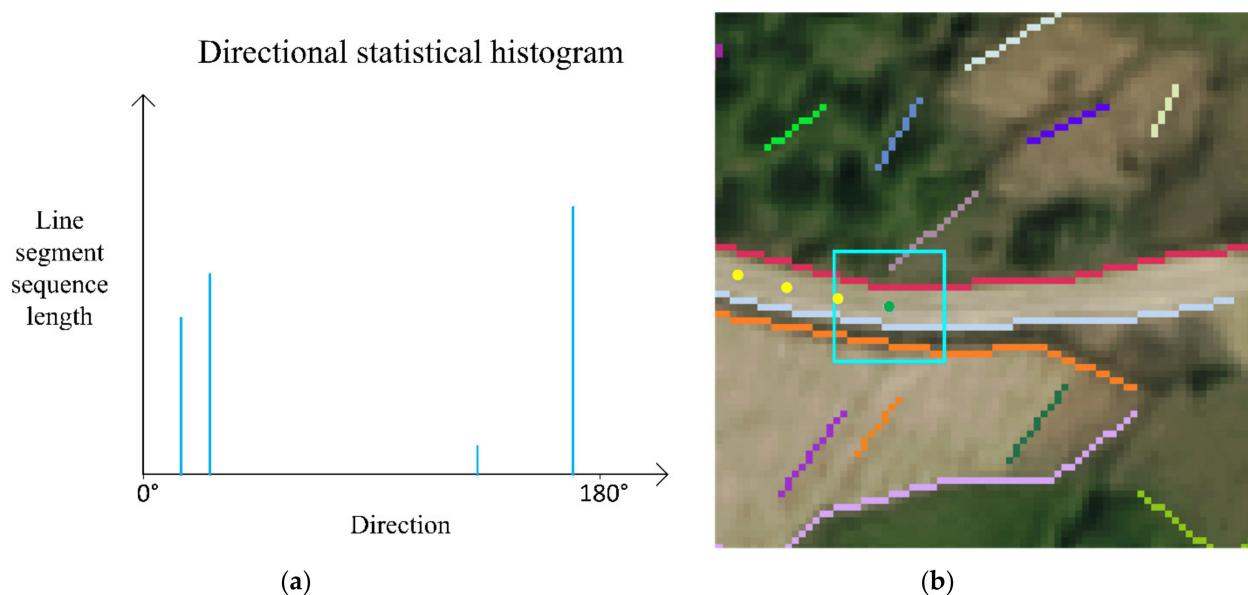
Figure 5. Tracking direction prediction model. (**a**) Direction distribution statistics of line segment sequence. (**b**) Line segment sequence detected by the current centre point.

2.2.4. Geometric Texture Combination Matching Model

In this paper, geometric weight is added to the matching model and a matching model of geometric texture combination is proposed. First, the line segment sequence detected by the reference point is divided into the corresponding matching points according to the direction. On this basis, it is fused with the texture measure value corresponding to the matching point to calculate the geometric texture combination measure value, and the maximum combination measure value is the best matching point. As shown in Figure 6, after using the method in this paper, the best matching template changes from the yellow matching template deviating from the road centre to the green matching template in the road centre. The matching model flow of this paper is as follows:
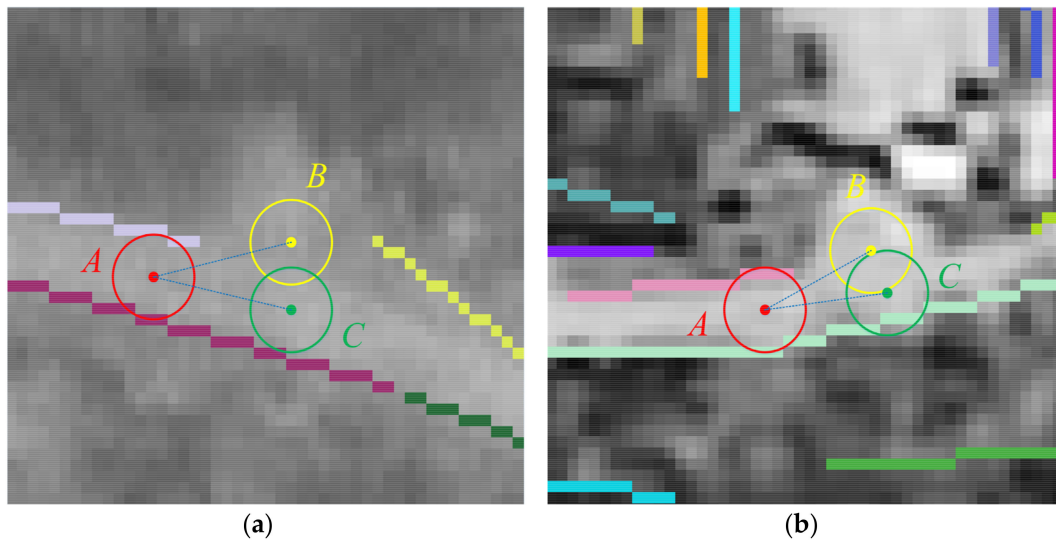
(**a**) (**b**)

Figure 6. The matching model mainly solves the following two problems: (**a**) The similarity of road and surrounding surface textures. (**b**) The partial occlusion of roads.

1. Template creation

Based on the characteristics of rural roads, multi-circle template is built based on predicting road directions, and a reference template and seven matching templates are obtained [27].

2. Geometric similarity measure

Using the road structure information provided by the line segment sequence, the line segment sequence detected by the reference point is divided according to the direction. The geometric similarity measure is calculated.

(1) The line segment sequence detected by the reference point is selected as the reference line segment sequence $\{L_1, L_2, \ldots, L_n\}$. The reference line segment sequence is divided into the closest matching points according to the indicated direction: that is, intervals of $[\theta - 30°, \theta + 30°]$.

(2) The confidence value K_i of each reference line segment sequence is calculated using Equation (2), where $Length_i$ is the length of the i-th reference line segment sequence. The length of each line segment sequence is used to compare the lengths of all detected surrounding line segments to obtain the confidence value K_i of the reference line segment sequence. The greater this confidence value is, the greater the probability that this line segment sequence is a road edge and the more accurate the indicated direction.

$$K_i = \frac{Length_i}{\sum\limits_{i=1}^{n} Length_i} \tag{2}$$

(3) The confidence values of the reference line segment sequence are accumulated in the direction indicated by the matching point, then obtain the geometric measurement values $\{G_1, G_2, G_3, G_4, G_5, G_6, \text{ and } G_7\}$ corresponding to each matching point. The larger the measurement value is, the closer the direction indicated by the matching point is to the road direction.

3. Texture similarity measure

In this paper, the gray variance and normalized cross-correlation coefficient (NCC) are fused to obtain the texture similarity measure.

(1) Calculate the gray variance in the matching template.

Gray variance measures the texture homogeneity of the image in a given region.

In Equation (3), N is the number of pixels in the matching template; $I(x_j, y_j)$ is the gray value of the j-th pixel; $Graymean1$ is the gray mean of the reference template; and v_i is the variance in the i-th matching template.

$$v_i = \frac{\sum\limits_{j=1}^{N} \left(I\left(x_j, y_j\right) - Graymean1\right)^2}{N} \tag{3}$$

Since the texture measurement value is jointly acted upon by the gray variance and NCC, it is necessary to normalize the two action values to equalize the influence of the gray variance and NCC in the determination of the texture measurement value. This paper uses the linear normalization function provided by OpenCV. In Equation (4), src_i is the i-th initial value; src_x is all initial values; max is the maximum value of the normalized range, which is 1 in this paper; min is the minimum value of the normalized range, which is 0 in this paper; and dst_i is the i-th normalized value.

$$dst_i = \frac{[src_i - min(src_x)] \times (max - min)}{max(src_x) - min(src_x)} + min \tag{4}$$

where v_i is normalized with Equation (4) to obtain the normalized variance values $\{V_1, V_2, V_3, V_4, V_5, V_6, \text{ and } V_7\}$.

(2)　Calculate the NCC between the matching template and the reference template.

In the model proposed in this paper, the reference template is used as the reference graph, the matching template is used as the real-time graph, and the correlation coefficient between the two is calculated to determine the matching effect.

$$p_i = \frac{\frac{1}{N}\sum\limits_{j=1}^{N} \left(S_j - \overline{S}\right)\left(g_j - \overline{g}\right)}{\sqrt{\frac{1}{N}\sum\limits_{j=1}^{N}\left(S_j - \overline{S}\right)}\sqrt{\frac{1}{N}\sum\limits_{j=1}^{N}\left(g_j - \overline{g}\right)}} \tag{5}$$

In Equation (5), N is the number of pixels in the reference graph and the real-time graph; S_j is the gray value of the j-th pixel in the real-time graph; $\overline{S}$ is the average gray value of the pixels in the real-time graph; g_j is the gray value of the j-th pixel in the reference graph; $\overline{g}$ is the average gray value of the pixels in the reference graph; and p_i is the correlation coefficient corresponding to the i-th real-time graph and the reference graph. The pixels of the real-time graph and the reference graph must individually correspond to the same position in their respective regions.

The correlation coefficient p_i satisfies Equation (6), and the similarity between the two images is measured in the range of $[-1, 1]$. The closer to 1 the value is, the stronger the similarity.

$$|p_i| \leq 1 \tag{6}$$

The obtained p_i values are normalized using Equation (4), and the normalized correlation coefficients $\{P_1, P_2, P_3, P_4, P_5, P_6, \text{ and } P_7\}$ are obtained.

(3)　The gray variance and NCC are fused to obtain the texture measurement.

In Equation (7), V_i is the normalized variance value of the i-th matching template and P_i is the normalized correlation coefficient of the i-th matching template. Since larger variance is correlated with smaller matching degree, the size of $(1 - V_i)$ is used to represent the matching degree. The term t_i represents the texture measurement value corresponding to the i-th matching point, and this term is normalized using Equation (4) to obtain the normalized matching point texture measurement values $\{T_1, T_2, T_3, T_4, T_5, T_6, \text{ and } T_7\}$.

$$t_i = (1 - V_i) \times P_i \tag{7}$$

4. Matching model

(1) Calculation of the combined measurement values.

$$C_i = G_i + T_i \tag{8}$$

In Equation (8), G_i is the geometric measurement value corresponding to the *i*-th matching point and T_i is the texture measurement value corresponding to the *i*-th matching point. Since the value ranges of both values are [0, 1] and their corresponding values are directly proportional to the matching effect, these values are added to obtain the final combined measurement value C_i, with a value range of [0, 2]. The maximum value among $\{C_1, C_2, C_3, C_4, C_5, C_6,$ and $C_7\}$ is selected, and the corresponding matching point indicates the best matching point.

(2) Template comparison. In the equations below, *Graymean*2 is the gray mean of the best matching template and *Gray*2 is the gray value of the corresponding matching point. $\overline{Graymean(A)}$ is the average gray value of reference template set *A*, and $\overline{Gray(A)}$ corresponds to the average gray value of the reference point. Set *A* is composed of 5 recently obtained reference templates. If fewer than 5 reference templates have been obtained, set *A* is composed of all currently obtained reference templates. In this way, our template comparison is flexible and avoids the contingency caused by a single comparison. In addition, the grayscale is divided into 16 equal levels with sizes of *g* [27].

$$\left| Graymean2 - \overline{Graymean(A)} \right| \le g \tag{9}$$

$$\left| Gray2 - \overline{Gray(A)} \right| \le g \tag{10}$$

If Equations (9) and (10) are satisfied at the same time, the texture similarity between the best matching template and the previous tracking points is confirmed to adhere to the constraints. At this time, the centre point of the best matching template, that is, the seed point, is fine-tuned to be equidistant from the road edges to obtain an optimized seed point. Then, the best-matching template can be retained as a seed point, and the tracking of the next point can be continued on this basis. If these conditions are not met, the step size is gradually increased to 5 times the road width [25] to span part of the occluded area. If the constraint is still not satisfied, the final result is printed.

3. Results

3.1. Comparison Method

In this paper, manual input points are used to semi-automatically extract roads, so four template-matching methods are selected for comparison with the proposed method: the circular template proposed by Lian et al. [23], the T-shaped template proposed by Lin et al. [24], the sector descriptor method proposed by Dai et al. [25] and the semi-automatic method of extracting rural roads from high-resolution remote sensing images based on a multi-feature combination proposed by Dai et al. [27]. Through these comparisons, the feasibility and automation of the proposed method are verified.

3.2. Evaluating Indicators

Performance indicators such as precision, recall (integrity), and quality are important parameters for evaluating road extraction methods [30]. Precision refers to the percentage of roads that are correctly extracted. Recall is the ratio between the reference road data correctly matched with the total length of the reference road map. Intersection over union

(IoU) and F1 combine precision and recall into an individual metric, which is used as the final road quality index [31]. These metrics are calculated as follows:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{13}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{14}$$

where *TP* is the length of correctly extracted roads, *FP* is the length of non-road pixels extracted as roads, and *FN* is the length of roads that was not extracted by the algorithm.

The evaluation index used in this paper also considers the number of manually input points and the road extraction time to determine the road extraction efficiency. In addition, to calculate the evaluation index, ArcMap 10.2 software was used to hand-draw the actual ground-truth data of the experimental data for comparison with the experimental results.

3.3. Experimental Analysis

3.3.1. Experiment 1

Figure 7a shows a GF-2 orthophoto with a spatial resolution of 1 m and a size of 4000 × 4000. The image displays a rural mountainous area in Enshi city, Hubei Province, China (see Table 1 for specific information).

In the experimental area shown in Figure 7b, it is obvious that the curvatures of roads in rural areas change continuously. Using the method proposed in this paper, first, the local road tracking direction is obtained from the tracking direction prediction model, and the outputs fit the road edges well. Second, the tracking step is shortened in curved road areas to avoid situations in which the tracking process extends beyond the road edge. The combination of these two steps enables continuous seed point tracking in most curved-road sections, and only one seed point must be input. The T-shaped template method [24] and the circular template method [23] do not consider the basic road direction data provided by the structural road information, so a large number of manually input points are needed for road tracking at curvature-change areas in these methods. The MLSOH descriptor is added in the sector template method [25] and the multi-feature combination method [27], so large numbers of artificial points are not required in areas with slow curvature changes; however, the direction predictions fail in road areas with large bending degrees, and small numbers of manually input points are required for processing.

In the experimental area displayed in Figure 7c, the road is mainly constructed of earth and has a certain degree of bending, and sand and stones are located on both sides of the road, resulting in the road having a texture similar to that of the surrounding features. The circular template method and T-shaped template method depend on texture matching in this area. Due to the lack of direction predictions, tracking beyond the road boundary easily occurs. Although the sector template method can eliminate the influence of some curves, the resulting direction predictions are not sufficiently restrictive in the convergent texture area, and a small number of artificial points are also needed for tracking in this area. The multi-feature combination method basically does not need to supplement points to adapt to the changes in curvature in this area. However, in the convergent texture area, the road converges in one direction and does not meet the HSV spatial interactive matching conditions. Therefore, panchromatic matching is still used, and this process invalidates the effect of the multi-feature combination method in excluding the convergent texture area. The method proposed in this paper further constrains the geometric information of roads to enlarge the geometric measurement value of the actual road direction. On this basis, the

texture measurement value is added to extract the roads in the convergent texture region, and there is no need to manually supplement points in the region shown in Figure 7c.
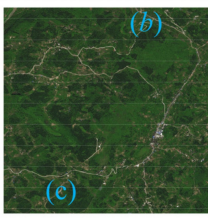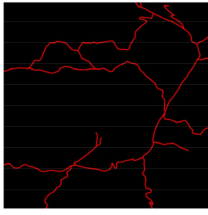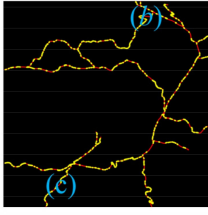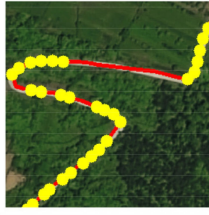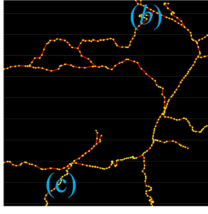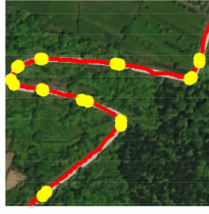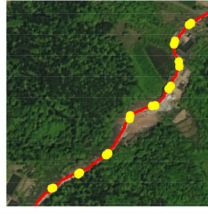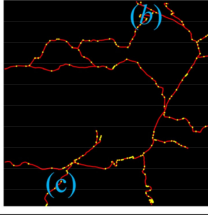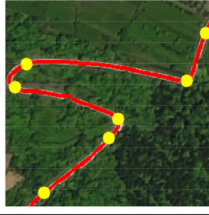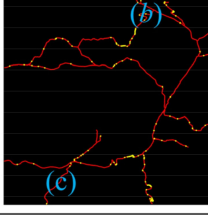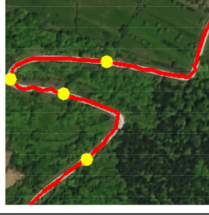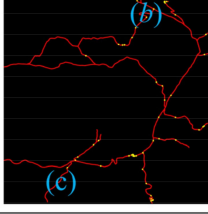

Figure 7. Road extraction results of the five methods analysed in Experiment 1. (**a**) Global graph. (**b**) Local magnification of region b. (**c**) Local magnification of region c.

3.3.2. Experiment 2

Figure 8a shows a GF-7 orthophoto with a spatial resolution of 0.65 m and a size of 4000 × 4000. The image displays a mountainous area in Zhangjiakou city, Hebei Province, China (see Table 1 for specific information).

In the experimental area shown in Figure 8b, the lower half of the road is low-grade and is therefore narrow, so the area of the template is small during the road tracking process. However, in the road extraction method based on template matching, the texture characteristics in the template are the decisive factors applied to obtain the matching effect. If the area covered by the template is small, then the selection of the best-matching template also changes when the road texture changes slightly, leading to poor tracking stability in the narrow road section. After adding the geometric measurement information, the geometric measurement value in the main direction is significantly higher than those in the other surrounding directions. Therefore, seed point tracking can be constrained in the direction of the road to overcome the interference of uneven road textures in the mountainous area. Our method requires the fewest manually input points on narrow roads, ensures accuracy, and is confirmed to perform well when applied to narrow roads.

The road shown in Figure 8c is a dirt road at the edge of a mountainous area. Due to the large traffic flow, the texture contrast between the road and the surrounding ground is poor. The T-shaped template method [24] and circular template method [23] do not consider any road edge information, so large numbers of seed points need to be manually input in the convergent texture area to continuously track the road when using these methods. The sector template method [25] predicts the road direction according to the MLSOH descriptor; this method can ensure that the road tracking is always oriented in the road direction, but it also requires four points to be input to eliminate the influence of texture similarity. The multi-feature combination method [27] uses the HSV spatial matching model to solve areas in which the texture of the road is similar to that of the surrounding ground objects. However, in areas where the contrast in the HSV colour space is not sufficiently strong, the required conditions of the HSV spatial matching model are not met [27]. Therefore, two seed points also need to be input in the region shown in Figure 8c. Our method calculates the sum of the geometric measurement value and the texture measurement value of the matching template, and this method does not consider the texture measurement value as the only factor when determining the best-matching template; thus, the road tracking method proposed in this study ensures that the correct tracking direction is followed under texture constraints. The road section shown in Figure 8c can be tracked well without requiring the manual addition of points.

3.3.3. Experiment 3

Figure 9a shows a GF-2 orthophoto with a spatial resolution of 1 m and a size of 4000 × 4000. The image displays a rural area in Dandong city, Liaoning Province, China (see Table 1 for specific information).

The road shown in the experimental area in Figure 9b is not an urban trunk road; rather, it is a low-grade earthen road and is similar to the surrounding ground material. In this case, if the initial point selection is not accurate, the initial information is not sufficient to provide an optimal reference template. The T-shaped template method [24] requires a manual visual determination of the initial road point, so artificial deviation arises when using this method. In the circular template method [23], sector template method [25], and multi-feature combination method [27], the road width is calculated using an adaptive correction model [25] to obtain the centre point of the road. However, due to the great limitations involved in setting the gradient threshold, the adaptability of this type of method is poor in road areas where the road texture is not clearly distinguished from that of the surrounding features. The adaptive road width extraction model proposed in this paper is based on obtaining a strong fit of the road edge contained in the line segment sequence, and this method is not limited to the determination of artificial vision or by a gradient threshold; thus, the proposed method can accurately calculate the initial road

centre point. As shown in Figure 9b, the method proposed in this paper requires only one seed point to be input to achieve road extraction. Although the multi-feature method also needs to input only one seed point, the overall road extraction deviates from the road centre due to the poor accuracy of road width calculation using adaptive template.
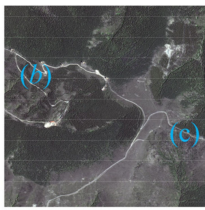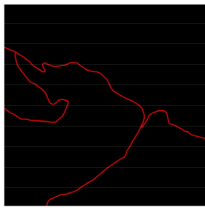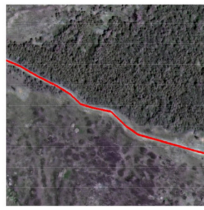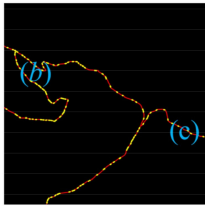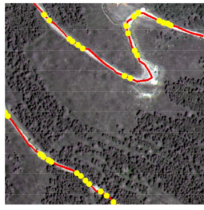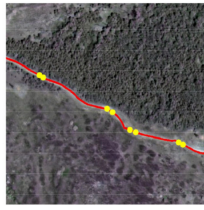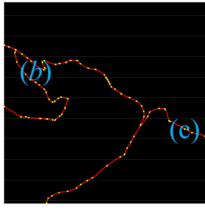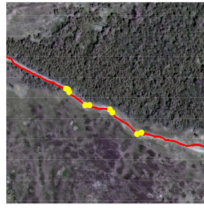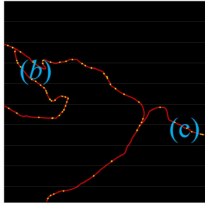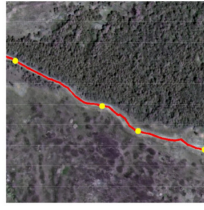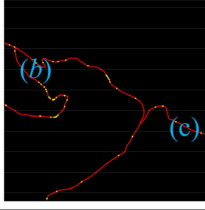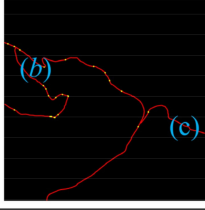

Figure 8. Road extraction results of the five methods analysed in Experiment 2. (**a**) Global graph. (**b**) Local magnification of region b. (**c**) Local magnification of region c.

Figure 9. Road extraction results of the five methods analysed in Experiment 3. (**a**) Global graph. (**b**) Local magnification of region b. (**c**) Local magnification of region c.

The experimental area shown in Figure 9c comprises an area of aggregated housing, so some roads in this region are covered by shadows of houses and vegetation. The circular template method and T-shaped template method mainly rely on texture information for

road matching, so large numbers of manually input points are required in these methods to solve the problem of occlusion. The sector template method and multi-feature combination method can rely on the direction provided by the MLSOH descriptor to avoid directional interference to some extent. However, in areas where the roads are partially occluded by shadows, although the matching template located in the centre of the road meets the texture constraint conditions, the texture measurement value of other matching templates lacking shadow interference may be larger. In this case, the road would be tracked outside the actual road area, so a small number of points need to be manually added to avoid tracking errors. In the method proposed in this paper, the sum of the geometric measurement value and texture measurement value is calculated. After adding a large geometric measurement value to the matching template in the road centre, the template can become the best-matching template, and the interference of some shadows can be eliminated.

4. Discussion

In Table 2, the accuracy, recall rate, quality, input points and time required for the five road extraction methods are calculated. Overall, the five road extraction methods compared are all template matching methods, so the values of accuracy, recall rate and F1 score are all above 95%, and most of them are above 98%.

Table 2. Comparative results of the five methods analysed in the experiment.

		Precision (%)	Recall (%)	IoU (%)	F1 (%)	Seed Points	Time (s)
	T-shape method	97.90	99.77	97.68	98.82	874	714
	Circle method	97.19	99.87	97.07	98.51	751	487
Experiment 1	Sector method	96.75	95.65	92.67	96.20	195	361
	Multifeature method	97.20	95.12	92.59	96.15	135	453
	Proposed method	99.02	96.43	95.52	97.71	56	225
	T-shape method	98.77	99.65	98.43	99.21	250	362
	Circle method	97.50	99.61	97.13	98.55	166	109
Experiment 2	Sector method	97.96	99.23	97.23	98.59	63	242
	Multifeature method	98.60	99.66	98.27	99.13	46	232
	Proposed method	99.07	99.47	98.55	99.27	25	104
	T-shape method	99.37	99.38	98.75	99.37	546	516
	Circle method	99.29	99.66	98.95	99.47	328	199
Experiment 3	Sector method	98.83	99.34	98.19	99.09	56	247
	Multifeature method	98.39	99.35	97.77	98.87	47	238
	Proposed method	99.32	99.38	98.71	99.35	22	181

Compare the five experimental methods. The T-shaped template method and the circular template method lack road direction information, so a large number of artificial points are added to improve the accuracy of the method. By adding the input points and time results of the three experiments in Table 2, it can be seen that the T-shaped template method and the circular template method need to input 1670 points and 1245 points, respectively, for processing the three images, requiring 1592 s and 795 s, respectively, far exceeding the other three methods. This proves that the efficiency of the T-shaped template method and circular template method is very low, and there is not much comparative value. The accuracies of the sector template method and the multifeature combination method are slightly lower. However, due to the MLSOH descriptor for direction prediction, fewer manual points are required, and the efficiency is greatly improved compared with the first two methods. The processing of the three images required 314 points and 228 points, which required 850 s and 923 s of processing time, respectively. In this paper, road extraction was carried out with geometric texture combination constraints. The six evaluation indexes of the experiment were all higher than those of the sector and multifeature combination methods, and the input points and time were significantly reduced. Only 103 points and 510 s of processing time were needed to process three images. It is proven that the method of extracting rural roads with geometric texture combined with measured values in this

paper can ensure certain road extraction accuracy and improve the efficiency of road extraction.

This paper selects three rural road images in different spaces. It shows the scene corresponding to the continuous change of curvature of rural roads and the similarity between roads and surrounding features. These problems are important factors affecting the degree of automation of the template matching method. The T-shaped template method and circular template method rely on the two most basic template matching principles of road internal homogeneity and road internal and external heterogeneity characteristics, and can only match simple roads. In the case of continuous curvature change and road texture similar to surrounding ground objects, a large number of artificial points need to be added to intervene. The sector template method adds the MLSOH descriptor on the basis of template matching. Road tracking has a certain direction basis and adapts to some problems of curvature continuity of rural roads, but it is difficult to solve the problem of texture similarity. On the one hand, the multifeature combination method improves the MLSOH descriptor and makes the road direction information more reliable. On the other hand, the interactive matching model of panchromatic and HSV space is used to solve some problems of similar texture. In this paper, the adaptive road width extraction model and the tracking direction prediction model are proposed to improve the accuracy of initial road point and adapt to the continuous change of rural road curvature. In addition, the matching model under the constraint of geometric texture is proposed to further solve the texture similarity problem between rural roads and their surroundings. Experiments in three different areas prove the universality of the proposed method.

5. Conclusions

In this paper, we propose a semi-automatic rural road extraction method that combines road geometric and texture constraints. The adaptive road width extraction model is used to improve the quality of initial road points. The road tracking direction prediction model adapts to the characteristics of continuous changes in rural road curvature. The geometric texture matching model is used to solve the matching problems that arise when the road and background characteristics are similar. Finally, different types of data are used to verify the effectiveness of the proposed method. The three sets of experiments conducted in this paper show that the proposed method can ensure road extraction accuracy, with a recall rate of more than 95%, while improving the degree of automation to a certain extent.

However, the method in this paper still has some shortcomings: (1) the applicability of the method proposed in this paper to long-distance shadow occlusion or shadow occlusion at road curves is poor and (2) further research is needed to determine how to extract road regions containing neither geometric information nor obvious texture information.

Author Contributions: Conceptualization, Hai Tan, Zimo Shen, and Jiguang Dai; methodology, Hai Tan and Jiguang Dai; software, Zimo Shen; validation, Hai Tan, Zimo Shen, and Jiguang Dai; formal analysis, Hai Tan and Zimo Shen; investigation, Hai Tan and Jiguang Dai; resources, Hai Tan; data curation, Zimo Shen and Jiguang Dai; writing—original draft preparation, Hai Tan and Zimo Shen; writing—review and editing, Hai Tan, Zimo Shen, and Jiguang Dai; visualization, Hai Tan and Jiguang Dai; supervision, Zimo Shen and Jiguang Dai; project administration, Hai Tan; funding acquisition, Hai Tan and Jiguang Dai. All authors have read and agreed to the published version of the manuscript.

References

1. Shi, W.; Miao, Z.; Debayle, J. An Integrated Method for Urban Main-Road Centerline Extraction From Optical Remotely Sensed Imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3359–3372. [CrossRef]
2. Frizzelle, B.G.; Evenson, K.R.; Rodriguez, D.A.; Laraia, B.A. The importance of accurate road data for spatial applications in public health: Customizing a road network. *Int. J. Health Geogr.* **2009**, *8*, 24. [CrossRef] [PubMed]
3. Tunde, A.M.; Adeniyi, E.E. Impact of Road Transport on Agricultural Development: A Nigerian Example. *Ethiop. J. Environ. Stud. Manag.* **2012**, *5*, 232–238. [CrossRef]
4. Hong, M.; Guo, J.; Dai, Y.; Yin, Z. A novel FMH model for road extraction from high-resolution remote sensing images in urban areas. *Procedia Comput. Sci.* **2019**, *147*, 49–55. [CrossRef]
5. Dai, J.; Wang, Y.; Du, Y.; Zhu, T.; Xie, S.; Li, C.; Fang, X. Development and prospect of road extraction method for optical remote sensing image. *J. Remote Sens.* **2020**, *24*, 804–823. (In Chinese) [CrossRef]
6. Hedman, K.; Stilla, U.; Lisini, G.; Gamba, P. Road Network Extraction in VHR SAR Images of Urban and Suburban Areas by Means of Class-Aided Feature-Level Fusion. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1294–1296. [CrossRef]
7. Baltsavias, E.P. Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps towards operational systems. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 129–151. [CrossRef]
8. Grote, A.; Heipke, C.; Rottensteiner, F. Road Network Extraction in Suburban Areas. *Photogramm Rec.* **2012**, *27*, 8–28. [CrossRef]
9. Lei, X.; Wang, W. A Method of Road Extraction from High-resolution Remote Sensing Images Based on Shape Features. *Acta Geod. Cartogr. Sin.* **2009**, *38*, 457–465. [CrossRef]
10. Kumar, M.; Singh, R.K.; Raju, P.L.N.; Krishnamurthy, Y.V.N. Road Network Extraction from High Resolution Multispectral Satellite Imagery Based on Object Oriented Techniques. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *8*, 107. [CrossRef]
11. Bakhtiari, H.R.R.; Abdollahi, A.; Rezaeian, H. Semi automatic road extraction from digital images. *Egypt. J. Remote Sens.* **2017**, *20*, 117–123. [CrossRef]
12. Cao, Y.; Wang, Z.; Li, S.; Xiao, X.; Yang, L. Fusion of Pixel-based and Object-based Features for Road Centerline Extraction from High-resolution Satellite Imagery. *Acta Geod. Cartogr. Sin.* **2016**, *45*, 1231–1249. [CrossRef]
13. Heermann, P.D.; Khazenie, N. Classification of Multispectral Remote Sensing Data Using a Backpropagation Neural Network. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 81–88. [CrossRef]
14. Lu, X.; Zhong, Y.; Zheng, Z.; Zhang, L. GAMSNet: Globally aware road detection network with multi-scale residual learning. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 340–352. [CrossRef]
15. Zhou, M.; Sui, H.; Chen, S.; Wang, J.; Chen, X. BT-RoadNet: A boundary and topologically-aware neural network for road extraction from high-resolution remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *168*, 288–306. [CrossRef]
16. Shao, Z.; Zhou, Z.; Huang, X.; Zhang, Y. MRENet: Simultaneous Extraction of Road Surface and Road Centerline in Complex Urban Scenes from Very High-Resolution Images. *Remote Sens.* **2021**, *13*, 239. [CrossRef]
17. Xu, Z.; Shen, Z.; Li, Y.; Xia, L.; Wang, H.; Li, S.; Jiao, S.; Lei, Y. Road Extraction in Mountainous Regions from High-Resolution Images Based on DSDNet and Terrain Optimization. *Remote Sens.* **2021**, *13*, 90. [CrossRef]
18. Gong, J.; Ji, S. Photogrammetry and Deep Learning. *Acta Geod. Cartogr. Sin.* **2018**, *47*, 693–704. [CrossRef]
19. Zhu, Q.; Zhang, Y.; Wang, L.; Zhong, Y.; Guan, Q.; Lu, X.; Zhang, L.; Li, D. A Global Context-aware and Batch-independent Network for road extraction from VHR satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 353–365. [CrossRef]
20. Wei, Y.; Zhang, K.; Ji, S. Simultaneous Road Surface and Centerline Extraction From Large-Scale Remote Sensing Images Using CNN-Based Segmentation and Tracing. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8919–8931. [CrossRef]
21. Wang, M.; Zhang, S. Road extraction from high-spatial-resolution remotely sensed imagery by combining multi-profile analysis and extended Snakes model. *Int. J. Remote Sens.* **2011**, *32*, 6349–6365. [CrossRef]
22. Leninisha, S.; Vani, K. Water flow based geometric active deformable model for road network. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 140–147. [CrossRef]
23. Lian, R.; Wang, W.; Li, J. Road Extraction from High-resolution Remote Sensing Images Based on Adaptive Circular Template and Saliency Map. *Acta Geod. Cartogr. Sin.* **2018**, *47*, 950–958. [CrossRef]
24. Lin, X.; Zhang, J.; Liu, Z.; Jing, S.; Duan, M. Semi-automatic extraction of road networks by least squares interlaced template matching in urban areas. *Int. J. Remote Sens.* **2011**, *32*, 4943–4959. [CrossRef]
25. Dai, J.; Zhu, T.; Zhang, Y.; Ma, R.; Li, W. Lane-Level Road Extraction from High-Resolution Optical Satellite Images. *Remote Sens.* **2019**, *11*, 2672. [CrossRef]
26. Sun, C.Y.; Zhou, T.G.; Chen, S.B.; Shen, J.W.; Wang, J.F.; Yang, H.; University, J. Research of a Semi-Automatic Extraction Method for Linear Features Based on Rectangular Template Matching. *J. Southwest* **2015**, *37*, 155–160. [CrossRef]
27. Dai, J.; Ma, R.; Ai, H. Semi-automatic Extraction of Rural Roads From High-Resolution Remote Sensing Images Based on a Multifeature Combination. *IEEE Geosci. Remote Sens. Lett.* **2020**, *PP*, 1–5. [CrossRef]
28. Dai, J.; Ma, R.; Gong, L.; Shen, Z.; Wu, J. A Model-Driven-to-Sample-Driven Method for Rural Road Extraction. *Remote Sens.* **2021**, *13*, 1417. [CrossRef]
29. Li, X.; Lu, C.; Yi, X.; Jia, J. Image Smoothing via L0Gradient Minimization. *ACM Trans. Graph.* **2011**, *30*, 1–12. [CrossRef]

30. Sujatha, C.; Selvathi, D. Connected component-based technique for automatic extraction of road centerline in high resolution satellite images. *Eurasip J. Image Video* **2015**, *2015*, 8. [CrossRef]
31. Ekim, B.; Sertel, E.; Kabadayı, M.E. Automatic Road Extraction from Historical Maps Using Deep Learning Techniques: A Regional Case Study of Turkey in a German World War II Map. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 492. [CrossRef]