World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:3, No:6, 2009

Automatic Extraction of Roads from High Resolution Aerial and Satellite Images with Heavy Noise

Yan Li and Ronald Briggs

***Abstract*—**Aerial and satellite images are information rich. They are also complex to analyze. For GIS systems, many features require fast and reliable extraction of roads and intersections. In this paper, we study efficient and reliable automatic extraction algorithms to address some difficult issues that are commonly seen in high resolution aerial and satellite images, nonetheless not well addressed in existing solutions, such as blurring, broken or missing road boundaries, lack of road profiles, heavy shadows, and interfering surrounding objects. The new scheme is based on a new method, namely *reference circle*, to properly identify the pixels that belong to the same road and use this information to recover the whole road network. This feature is invariable to the shape and direction of roads and tolerates heavy noise and disturbances. Road extraction based on reference circles is much more noise tolerant and flexible than the previous edge-detection based algorithms. The scheme is able to extract roads reliably from images with complex contents and heavy obstructions, such as the high resolution aerial/satellite images available from Google maps.

***Keywords*—**Automatic road extraction; Image processing; Feature extraction; GIS update; Remote sensing; Geo-referencing.

I. Introduction

EXTRACTION of road network from raster images is a very important part of many GIS features such as GIS updating, geo-referencing [1,2], and geospatial data integration. However, extracting road networks from a raster image is a time-consuming operation when performed manually, especially when the image is complex. Automatic road extraction is critical and essential to the fast and effective processing of large number of raster images in various formats, complexities, and conditions.

How roads are extracted properly from raster images depends heavily on how roads appear in raster images. In this paper, we study the road extraction from high resolution images, e.g., satellite image, aerial photos. A high resolution satellite image typically has a resolution of 0.5 to 1 m. Under such high resolution, a road is not a thin line any more, instead, objects such as cars and trees are easily identifiable. This class of images contains very rich information and when fused with vector map can provide a comprehensive view of a geographical area. Google, Yahoo, and Virtual Earth maps are good examples to demonstrate the power of such high resolution images. However, high resolution images pose great challenges for automatic feature extraction due to the inherent complexities. First, a typical aerial photo captures everything in the area such as buildings, cars, trees, etc. Second, different objects are not isolated, but mixed and interfere with each other, e.g., the shadows of trees on the road, building tops with similar materials. Third, roads may even look quite differently within the same image, due to their respective physical properties. Assuming all roads have the same characteristics will fail to extract all the roads. In addition, the light and weather conditions have big impact over images. Therefore, it is impossible to predict what and where objects are, and how they look like in a raster image. All these uncertainties and complexities make the automatic extraction very difficult. Due to its importance, much effort has been devoted to this problem [3, 4]. Unfortunately, there are still no existing methods that can deal with all these problems effectively and reliably. Some typical high resolution images are shown in Figure 1 and they show huge difference among them in terms of the color spectrum and noise level.

Most of the existing extraction methods for high resolution images rely on the road boundaries as key hints for road extraction. For roads with straight and clean edges, these edges can be detected using edge detection algorithms. Then certain navigation algorithm is used to follow the edges and extract the roads. As we can see from the examples in Figure 1, roads in high resolution images typically don't have straight and clean edges. There are numerous factors that can distort the edges, including but not limited to blocking objects such as trees and shadows, surrounding objects in similar colors such as roof tops. As a matter of fact, the result of edge detection is as complicated as the image itself. Edges of roads are either missing or broken and straight edges correspond to buildings instead of roads, as shown in Figure 2. Therefore, edge-based extraction schemes will all fail to produce reliable results under such circumstances.

In this paper, we develop an integrated scheme for automatic extraction that exploits the inherent nature of roads. Instead of relying on the edges to detect roads, it tries to find

Y. Li is with the Geospatial Information Sciences, The University of Texas at Dallas, Dallas, TX, 75080 USA (phone: 469-878-8858; fax: 214-670-5341; e-mail: yanli@utdallas.edu).

R. Briggs is with the Geospatial Information Sciences, The University of Texas at Dallas, Dallas, TX, 75080 USA (email: briggs@utdallas.edu).

the pixels that belong to the same road region based on how they are related visually and geometrically. Studies have shown that the visual characteristic of a road is heavily influenced by its physical characteristics such as material, surface condition. It is impossible to define a common pattern just based on color or spectrum to uniquely identify roads. In our scheme, we consider a road as a group of "similar" pixels. The similarity is defined in the overall shape of the region they belong to, the spectrum they share, and the geometric property of the region. Different from edge-based extraction schemes, the new scheme first examines the visual and geometric properties of pixels using a new method called reference circle. Then, the key pixels, namely central pixels, are identified to represent each region. All the regions are verified based on the general visual and geometric constraints associated with a road. Therefore, the roof top of a building or a long strip of grass is not misidentified as a road segment. There is also no need to assume or guess the color spectrum of roads, which varies greatly from image to image according to the example images. As illustrated by examples in Figure 1, a road is not always a contiguous region of regular linear shape, but a series of segments with no constant shapes. This is because each segment may include pixels of surrounding objects in similar colors or miss some of its pixels due to interference of surrounding objects. A reliable extraction scheme must be able to deal with such issues. In the following sections, we will discuss how to capture the essence of "similarity" and translate them into segments and finally turn them into complete contiguous roads.

II. Previous Works

The existing approaches for road extraction cover a wide variety of strategies, using different resolution aerial or satellite images. A quite extensive overview of such approaches is given in [3, 4]. Overall, schemes can be divided into two groups: semi-automatic and automatic. Semi-automatic schemes require human interaction to provide some prior knowledge during the process of extraction, such as identifying the areas of roads. Based on the information provided by users, roads are then extracted by methods such as profile matching [5], cooperative algorithms [6], and dynamic programming or LSB-Snakes [7]. Automatic methods usually try to extract some hypotheses for road segments through edge and line detection and then establish connections between road segments to form road networks. Data from multiple sources may be combined [8] to improve the reliability. Depending on the type of images, in some schemes, contextual information is used to guide the extraction of roads [9]. For images that are not cluttered, reducing the resolution may help identify roads as lines [11]. However, most of the proposed methods share the common assumptions of relatively simplistic road models and require roads to be easily identified in images, such as constant intensity or straight and smooth road edges. As a result, they are very sensitive to interferences such as cars, shadows or occlusions, as shown in Figure 1, and do not always provide consistent and reliable results.

III. Automatic Extraction Scheme

The goal of the proposed system is to develop a complete and practical automatic solution to the road extraction problem for high resolution aerial/satellite images. It consists of two similar major stages. The first stage extracts roads that are relatively easier to identify such as major roads and the second stage deals with roads that are harder to identify. The reason for such design is a balance between reliability and efficiency. Some roads are easier to identify because they are more identifiable and contain relatively less noise. Since roads in the same image share some common visual characteristics, the information from the already extracted roads and other objects, such as spectrum, can be used to simplify the process of identifying roads that are less visible or heavily impacted by surrounding objects. Otherwise, such roads are not easily distinguishable from patterns formed by other objects. For example, a set of collinear small blocks may correspond to a road or a group of buildings (houses) from the same block. The second stage also serves an important purpose to fill the big gaps for road extracted in stage one. Under some severe noise, part of the road may be disqualified as valid road segments and hence missed in stage one, leaving some major gaps. With the additional spectrum information, these missed areas can be easily identified to complete the road extraction. Therefore, the two stage process eliminates the need to assume or guess the color spectrum of roads and allow road extraction to be much more complete.

Each major stage consists of the following three major steps: filtering, segmentation, and grouping and optimization, as shown in Figure 3. The details of each step will be discussed in the following section.

IV. Algorithms

In this paper, we assume images to satisfy the following two general assumptions. These two assumptions are derived based on the minimum conditions for roads to be identifiable and therefore are easily met by most images.

- Visual constraint: majority of the pixels from the same road have similar spectrum that is distinguishable from most of the surrounding areas;
- Geometric constraint: a road is a region that is relatively long and narrow, compared with other objects in the image;

These two constraints are different from the assumption made in most previous works. The visual constraint does not require a road region to have single color or constant intensity. It only requires roads to look visually different from surrounding objects in most parts. The geometric constraint does not require a smooth edge, only the overall shape to be a long narrow strip. So these conditions are much weaker and a lot more practical. As we can see, these assumptions can accommodate all the difficult issues very well, including blurring, broken or missing road boundaries, lack of road profiles, heavy shadows, and interfering surrounding objects.

A. Filtering

The step of filtering is to identify the key pixels that will help determine if the region they belong to is likely a road segment.

Based on the assumption of visual constraint, it is possible to establish an image segmentation using methods of edge detection. Notice that such separation of regions is not required to be precise and normally contains quite a lot of noise. In the best, the boundaries between regions are a set of line segments for most images, as is the case shown in Figure 2. It certainly does not tell which region corresponds to a road and which is not based on the extracted edges. As a matter of fact, most of the regions are not completely separated by edges and are still interconnected based on 4-connect path or 8-connect path. In order to fully identify and separate road regions from the rest of image, we propose a new method which involves two key concepts: *reference circle* and *central pixel*.

For a pixel p, its reference circle C(p) is the largest circle centered at p that does not contain edge points. Visually, the radius of the reference circle is the maximum distance from the pixel to the closet edge point. A pixel is a central pixel if it has the maximum reference circle among its neighboring pixels, which is a 3x3 area. Both concepts are illustrated in Figure 4. Figure 4 shows the reference circles of two points P and Q. Between them, Q is central pixel. The motivation of these two concepts is, under ideal situation, all the central pixels form the center line of a road. The incomplete and inaccurate edges will affect the location of central pixels to some degree. But central pixels have high tolerance over such error, especially when they are grouped together statistically, as illustrated in Figure 4. Therefore, all the central pixels can be considered as anchoring points for the road. For each central pixel, the set of all central pixels inside its reference circle is defined as its central pixel group, namely G(p). Again, ideally, the G(p) is the part of central line within the reference circle. In general case, we fit a line L(p) to the G(p) using the least squares method. L(p) is named as the axis for central pixel p and the direction of L(p) is called the direction of p. Figure 4 also shows the axis and direction for Q and S. To summarize, each central pixel has a local maximum reference circle, an axis, and a direction. The output of the filtering step is a set of central pixels for each region.

The central pixels can be found using distance transform based on the result of edge detection using either Sobel or Canny operator. Each central pixel corresponds to a local maximum in a 3x3 window in the distance transformed image. The radius of its reference circle is also the distance at the local maximum. The complete flow chart of the filtering algorithm is illustrated in Figure 5. Although the filtering algorithms still uses edge detection, it is not used to directly extract roads. Instead, it is used to locate reference circles and central pixels. There is no need to navigate along the broken edges.

At the filtering step of second stage, objects such as cars and trees in the middle of the roads can be removed using morphological reconstruction because the road spectrum is known after the first stage.

B. Segmentation

The step of segmentation is to verify which region is a possible road region based on the central pixels. Central pixels contain not only the centerline information of a region, but also the information of its overall geometric shape. For example, a perfect square will only have one central pixel at its center. A long narrow strip region will have large number of central pixels. Therefore, the ratio of total number of central pixels to the average radius of their reference circles is a good indicator for the shape of region. Only regions with ratios above certain thresholds are considered to be candidate regions.

In order to filter out interference as much as possible for reliable extraction during the first major stage, a minimum region width can be imposed. This will effectively remove most of the random objects from the image. However, such width constraint will be removed during the second major step as improper regions can also be filtered out based on the color spectrum information obtained from stage one. Therefore, small regions with close spectrum are examined for possible small roads in the second stage.

Another condition to be checked in segmentation is the number of collinear central points based on their directions. This check is done thru Hough transformation. Because a road might be broken into small segments by trees, shadows, etc, in some cases a road segment does not look much different from a long building top or open area if just based on its central pixels. However, the central pixels of a road across broken neighboring segments should look collinear most of the time depending on the curvature. This characteristic can be used to distinguish a road segment from isolated big objects such as buildings.

In addition to the geometry constraint, some general visual information can also be applied to filter out obvious non-road regions. For example, if the image is a color image, most of the tree and grass areas are greenish. Also tree areas usually contain much richer textures than normal smooth road surfaces. The intensity transformation and spatial and frequency filtering can be used to filter out such areas. The minimum assumption of the proposed scheme does not exclude the use of additional visual information to further improve the quality of extraction if they are available.

C. Grouping and Optimization

As the result of segmentation, the road segments are typically incomplete and disconnected due to heavy noise. The purpose of this step is to group corresponding road segments together in order to find the optimal results for road extraction. The optimization part may be skipped during stage 1 if enough information is available to determine the road spectrum. Optimization is better applied after all the segments are identified at stage two.

There are two issues to be addressed in this step: (1) which segments are related; (2) if related, how to group them together into one.

For the first issue, the algorithm checks the following properties based on visual and geometric constraints:

- Collinear central pixels
- Similar spectrum
- Similar average width
- Proximity
- Connectivity

Therefore, if there is a small gap between two very similar collinear segments the algorithm assumes they are grouped. Also, if the two collinear segments are joined by another segment which has similar color spectrum, they are grouped. This is likely the case when certain part of the road is heavily blocked by certain objects.

For the second issue, for each group of road segments, its centerline is calculated with best fitting algorithms using lines or splines depending on the curvature based on all the central pixels. It is further optimized using pixel-based group and optimization methods discussed in [13]. In summary, for the existing central pixels in each segment, they are optimized by removing sharp angles between points. An angle is considered to be sharp if it is above pi/8. To interconnect the segments over the gap, a linked path is calculated to balance the angles with both ends. Once a centerline is found, the contour of road can be derived based on the statistical estimation of the road width according to all the central pixels.

V. Experiments

For example 1(b), the results after major steps are shown in Figure 6(a)-(f). The threshold for road width is set to be 10 pixels during the extraction. Comparing 6(d) and 6(e), we can see that the spectrum information derived from stage 1 helps add a lot more detailed information to the previously extracted roads as well as the newly extracted roads.

One way to measure the quality of extraction is the Hausdorff fraction [14], which is defined from a point set $A = \{a_1, ..., a_p\}$ to a point set $B = \{b_1, ..., b_p\}$ as follows

$$hf^{\tau}(A,B) = \frac{|\{a \in A \mid \min_{b \in B} \| a - b \| < \tau\}|}{|A|}$$

This represents the faction of points in A, which are close to a point in B, within a threshold distance τ . The A and B are the sets of extracted roads central points and real central points respectively. Obviously, one can also define the Hausdorff fraction from B to A, where sets A and B switch roles.

In our experiments, we use road intersections as the set of measuring points in Hausdorff fraction, because they are used in automated georeferencing [1, 2]. We compare the results of automatic extraction with road intersections determined manually. Ten high resolution images are chosen randomly from different areas, which have noise level similar to the example shown in this paper. The results showed that the automatic extraction algorithm is robust and resilient to heavy noise. It is able to identify partially and sometimes fully obscured intersections correctly based on the nearby road information. On average, when the threshold distance is 5 pixels, the Hausdoff fraction is 83%.

The performance of extraction can also be measured using completeness and correctness [4]. They are defined as follows:

$$\text{completeness:} \quad \text{Cp}= \frac{l_{TM}}{l_I}$$

$$\text{correctness:} \quad \text{Cr} = \frac{l_{TM}}{l_E}$$

with l_{TM} standing for true matched roads length, l_E for extracted roads length, and l_I for ideal roads length. On average, the results of our experiments show Cp and Cr are 85% and 97% respectively.

VI. Conclusion

In this paper, we proposed a new automatic system for extracting roads and intersections from high resolution aerial and satellite images. The main contribution of the proposed system is to address the major issues that have caused all existing extraction approaches to fail, such as blurring boundaries, interfering objects, inconsistent road profiles, heavy shadows, etc. To address all these difficult issues, we develop a new method, namely reference circle and central pixel, to capture the essence of both visual and geometric characteristics of roads. The reference circle and central pixels are much less susceptible to the distortions that projected over road regions by other objects. The central pixels play a key role throughout the extraction process including filtering, segmentation, and grouping and optimization. The two stage process eliminates the need to assume or guess the color spectrum of roads. The proposed approach is efficient, reliable, and assumes no prior knowledge about the road conditions and surrounding objects. It is able to process complicated aerial/satellite images from a variety of sources including aerial photos from Google and Yahoo online maps.

References

[1] Y. Li, R. Briggs. Scalable and error tolerant automated georeferencing under affine transformations. IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, July 2008.

[2] Y. Li, R. Briggs. Automated georeferencing based on topological point pattern matching. The International Symposium on Automated Cartography (AutoCarto), Vancouver, WA, June 2006.

[3] J.B. Mena. State of the Art on Automatic Road Extraction for GIS Update: a Novel Classification. Pattern Recognition Letters, 24(16):3037-3058, 2003.

[4] M.-F. Auclair-Fortier, D. Ziou, C. Armenakis, and S. Wang. Survey of Work on Road Extraction in Aerial and Satellite Images. Technical Report 241, Département de mathématiques et d'informatique, Université de Sherbrooke, 1999.

[5] G. Vosselman, J.D. Knecht. Road Tracking by Profile Matching and Kalman Filtering. Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, pages 265-274, 1995.

[6] D.M. Mckeown, J.L. Denlinger. Cooperative Methods for Road Tracking in Aerial Imagery. Workshop Computer Vision Pattern Recognition, pages 662-672, 1988.

[7] A. Gruen, H. Li. Semi-automatic Linear Feature Extraction by Dynamic Programming and LSB-Snakes. Photogrammet Eng. Remote Sensing 63, pages 985-995, 1997.

[8] S.B. Hinz, A. Ebner. Modeling Contextual Knowledge for Controlling Road Extraction in Urban Areas. IEEE/ISPRS Joint Workshop Remote Sensing Data Fusion over Urban Areas, 2001.

[9] T. Ohlhof, T. Emge, W. Reinhardt, K. Leukert, C. Heipke, K. Pakzad. Generation and Update of VMAP data using satellite and airbone imagery. Remote Sensing, vol 33, pages 763-768, 2000.

[10] A. Baumgartner, C.T. Steger, H. Mayer, and W. Eckstein. Multi-Resolution, Semantic Objects, and Context for Road Extraction. In Wolfgang Förstner and Lutz Plümer, editors, Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, pages 140–156, Basel, Switzerland, 1997. Birkhäuser Verlag.

[11] A. Baumgartner, C.T. Steger, C. Wiedemann, H. Mayer, W. Eckstein, and H. Ebner. Update of Roads in GIS from Aerial Imagery: Verification and Multi-Resolution Extraction. Proceedings of International Archives of Photogrammetry and Remote Sensing, XXXI B3/III:53–58, 1996.

[12] D. Geman and B. Jedynak. An Active Testing Model for Tracking Roads in Satellite Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(1):1–14, January 1996.

[13] E. Christophe, J. Inglada. Robust Road Extraction for High Resolution Satellite Images. ICIP 2007, pages 437-440, October 2007.

[14] W.J. Rucklidge. Efficient Visual Recongnition Using Hausdorff Distance. Lecture Notes in Computer Science. Springer-Verlag, 1996.

[15] G. Lisini, C. Tison, F. Tupin, P. Gamba. Feature Fusion to Improve Road Network Extraction in High Resolution SAR Images. IEEE Geoscience Remote Sensing Letter, Vol 3, No. 2, pages 217-221, 2006.

[16] X. Jin, C.H. Davie. An Integrated System for Automatic Road Mapping from High Resolution Multispectural Satellite Imagery by Information Fusion. Information Fusion. Vol 6(4), pages 257-273, 2005.

Fig. 1(a) An aerial photo for an area in the City of Dallas

Fig. 1(b) High resolution image of a residential area

Fig. 1(c) Google map satellite image

Fig. 1 Examples of high resolution raster images

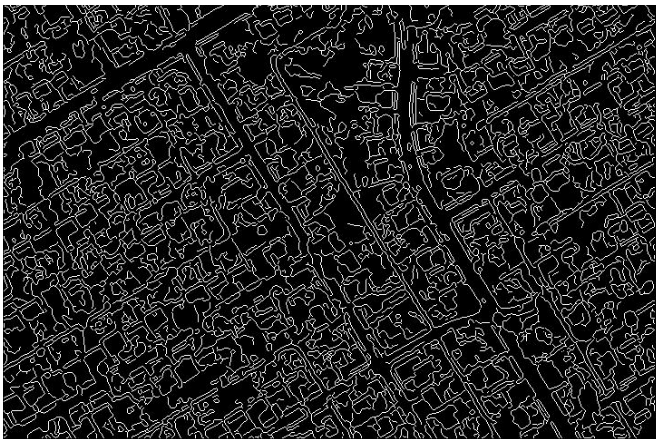

Fig. 2 Edge extraction for example 1(b)

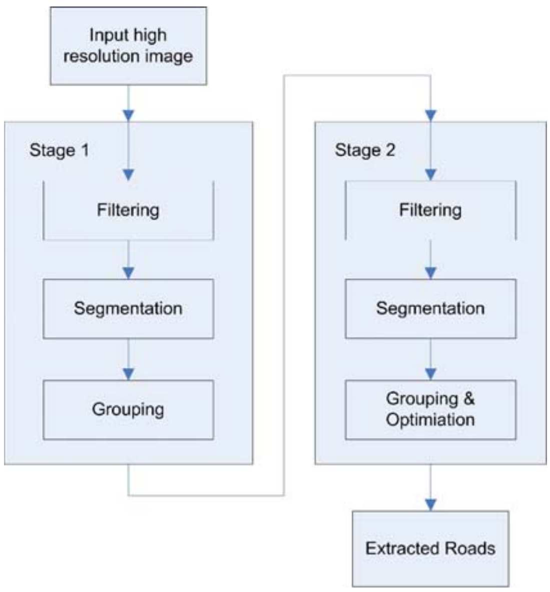

Fig. 3 Automatic extraction scheme

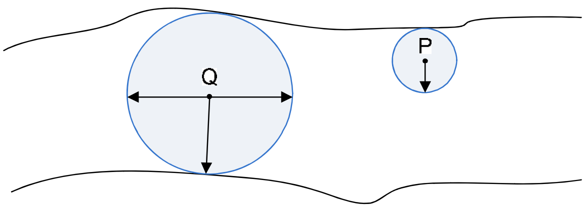

Fig. 4 Reference circle and central pixel

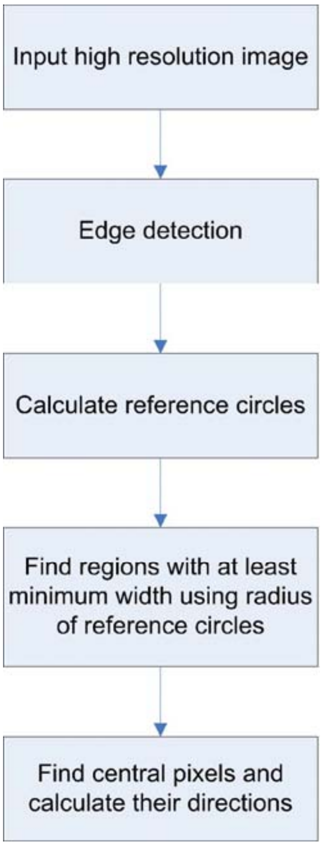

Fig. 5 Algorithm of filtering

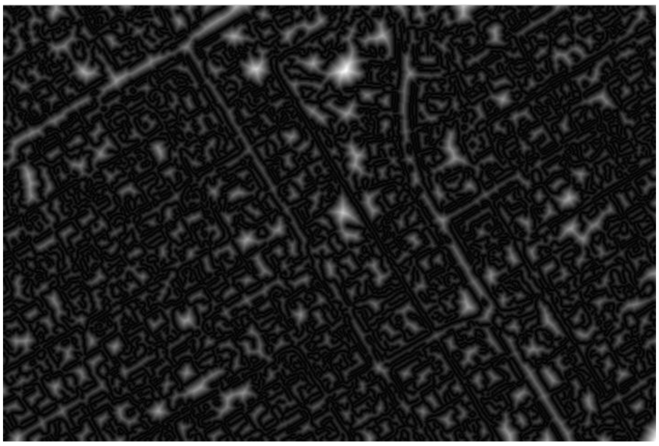

Fig 6(a) Reference circles

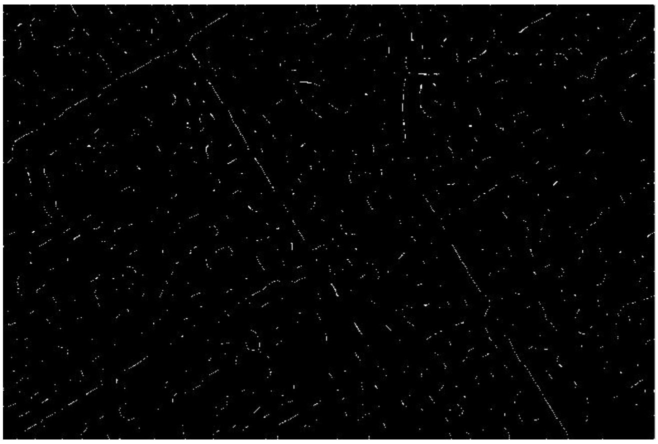

Fig. 6(b) Central pixels

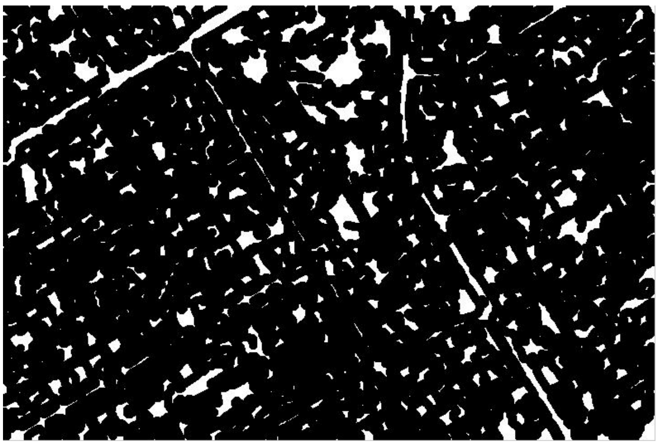

Fig. 6(c) Regions corresponding to the central pixels

Fig. 6(d) Result of segmentation of stage 1

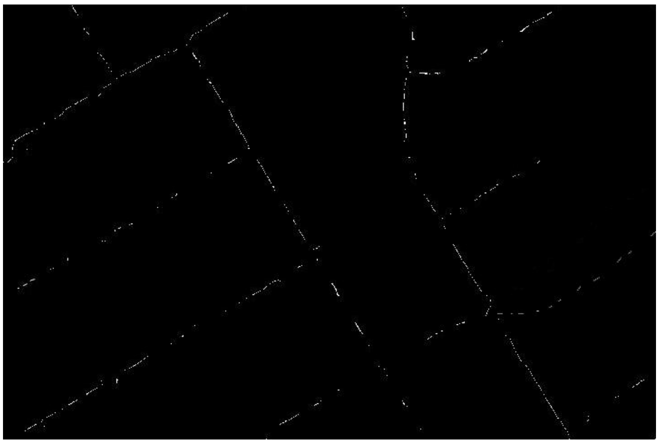

Fig. 6(e) Result of segmentation of stage 2

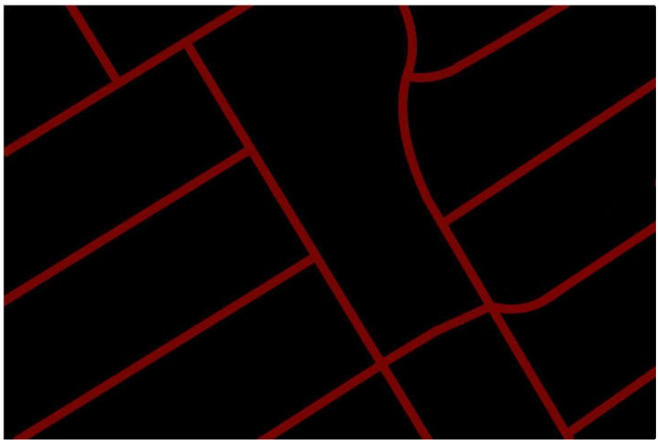

Fig 6(f) Result of grouping and optimization

Fig. 6 Experiment results