Road Network Extraction via Aperiodic Directional Structure Measurement

Yu Zang, Cheng Wang, *Member, IEEE*, Liujuan Cao, *Member, IEEE*,
Yao Yu, and Jonathan Li, *Senior Member, IEEE*

***Abstract*—In this paper, we present a novel aperiodic directional structure measurement (ADSM) toward road network extraction. Based on the observations from Cognitive Psychology regarding the aperiodicity and local directionality, ADSM can well characterize roadlike structures independent of the spectral character and contrast. By exploiting such measurement as guidance, we construct a mask to denote potential road regions. Then, by combining with some common morphology operators, our approach is able to provide robust road centerlines efficiently. We evaluate our approach with data from various satellite sensors and make comprehensive comparisons with previous state-of-the-art methods. Experimental results demonstrate the merit using our ADSM as a metric to identify potential road structures, as well as the effectiveness and efficiency of our road network extraction system.**

***Index Terms*—Road network extraction, road network topology analysis, road structure measurement.**

I. Introduction

ROAD network extraction is a fundamental but long-standing problem in remote sensing image processing. Robust road extraction systems make it possible to save much labor for transportation data acquisition and can also provide valuable prior knowledge for the detection and recognition of vehicles, buildings, or other objects.

Road network extraction has been gaining much attention in recent years, where a common assumption of many solutions is that roads are often characterized as high-contrast regions with low curvature and specific spectral behavior [1]. Therefore, the extraction boils down to determining the appropriate features to describe the "high contrast long and thin regions" [2]. However, as pointed by previous work [3], such an assumption may be limited in two cases. First, the spectral behaviors of data from different sensors appear differently, as shown in Fig. 1(a) and (b), which are two remote sensing images from Pleiades-1A and GeoEye satellite, i.e., the spectral intensity of road regions is bright for (a) and dark for (b). Such difference may lead to an additional recognition step such as simple interactions [4] or leaning-based classification [5]. Second, the road regions are of low spectral contrast, as shown in Fig. 1(c). Most works are able to extract the regular road network with strong spectral intensity contrast; however, for the zoomed-in area (as shown in the right part), when the contrast is not high enough, the detection becomes more challenging. To address these limitations, recent works [4]–[6] have been able to produce impressive results by introducing "assistant information" such as simple interaction [4], 3-D road surface model [6], or training samples [5].

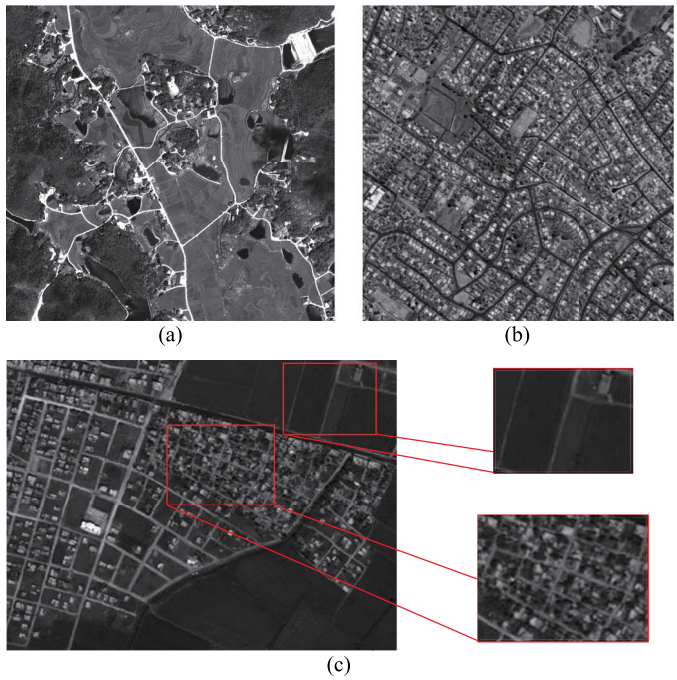


Fig. 1. Remote sensing images from different satellite sensors. (a) and (b) Examples with different spectral performance. (c) Example with low-contrast road regions.

On the other hand, we find that, for most human experts, these roads are easily marked. Such an observation provides an interesting clue that human perception about a road does not *just* rely on spectral behavior or contrast. In view of these problems, inspired by the human cognitive process, this paper attempts to explore a road extraction method without any information beyond the image itself. According to the theory of Cognitive Psychology, in addition to intensity contrast, saliency of prominent structures also relies on organized arrangement and "low social conformity" [7], which means that salient structures do not repeat periodically. Most previous works focus on the first two parts and employ various spectral behavior

Manuscript received June 26, 2015; revised November 7, 2015 and December 18, 2015; accepted December 29, 2015. This work was supported in part by the Natural Science Foundation of China under Grant 61501387, Grant 61371144, and Grant 61401382 and in part by the Postdoctoral Science Foundation of China under Grant 149621. *(Corresponding author: Cheng Wang.)*

Y. Zang, C. Wang, L. Cao, and Y. Yu are with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen 361005, China (e-mail: zangyu7@126.com; cwang@xmu.edu.cn).

J. Li is with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen 361005, China, and also with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2514602

and local directionality descriptors for road detection but pay less attention to the third aspect. Hence, in this approach, road regions are assumed to have the following two characters. First, they always align closely in a local area, i.e., have fine local directionality. Second, they do not repeat periodically often, i.e., have fine aperiodicity.

Based on these assumptions, we present a novel road structure measurement, which not only considers the geometry features but also includes an aperiodicity measurement term to evaluate the "low social conformity" of potential road regions. Such measurement provides valuable guidance for the followed road topology analysis and network extraction. Benefitting from this, based on just the image itself, our system is able to provide spectral and contrast independent road extraction results, efficiently.

The main contribution of this paper is a novel aperiodic directional structure measurement (ADSM)-guided road extraction scheme. Such a measurement provides a cheap but efficient characterization of road structures, whereupon the roadlike structures can be constructed independently with spectral character and contrast. Specifically, the novelty lies in the following two aspects.

1) A robust and efficient structure measurement (ADSM), which can capture potential low-contrast road regions embedded in heavy textures of buildings or residential areas, is established to evaluate not only the geometry features but also the aperiodicity of road structures.
2) A road region mask (a binary mask to denote potential road regions) construction scheme, which enables explicit constraint to guide the road extraction process, is designed to describe the road network topology, thus resulting in high-correctness and high-quality detection results.

II. Related Work

In recent years, works on road network extraction have shown remarkable progress, with many of them investigating the road network extraction problem from different views. Inspired by previous surveys [3], [8]–[10] and latest road extraction works [4], [5], [11], in accordance with whether "assistant information" are employed, we divided current road extraction works into two types.

Works with "assistant information" rely on pretrained classifiers, such as large-scale neural networks [1], locally excitatory globally inhibitory oscillator networks (LEGIONs) [11], higher order conditional random field model [12], spectral–spatial classifiers [5], [13]–[15], or prior guidance such as Geographic Information System (GIS) data [16].

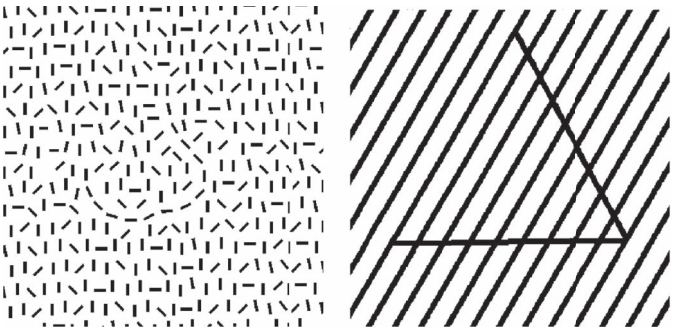
Early works of this kind focus on the classification of specific features (such as spectral or spatial features) of the road regions. By synthetically considering the spectral and spatial statistics of local structures [13], spectral–spatial classification-based road detection methods [14], [15] can acquire better robustness compared with pixelwise classification methods. Inspired by this idea, researchers widely applied mathematical morphology (MM) [17]–[19] to produce image spatial structures. To better cover the curved and rectilinear structures [20], general adaptive neighborhood (GAN) MM (GANMM) [21] is proposed to perform spectral–spatial classification.

Recent works tend to employ large-scale learning networks [1], [12], [16] to directly label the road pixels. Mnih and Hinton [1] first proposed a Graphics Processing Unit (GPU)–based deep learning network with local spatial coherence constraint; such network works well when there are sufficient labeled training samples. Works of Wegner *et al.* [12], which also require large amount of training samples, presented a higher order conditional random field model along with a superpixel-based labeling scheme; then, by the assistance of road network topology analysis, such work significantly improves both the accuracy and the topological correctness of the extracted roads. On the other hand, by the learning of existing road networks from GIS data, Peng *et al.* [16] proposed a statistical method, which is suitable for certain applications such as road map updating.

With the guidance of various learning networks, these methods are able to provide high-accuracy and robust detection results. However, in practical applications, these approaches are limited by two issues. First, a certain training model is often specific to certain kind of roads with similar features, and it may fail for test samples with different appearances. Second, recent large-scale learning networks always require large amount of training samples; sometimes, these samples are difficult to acquire, such as the foreign region. In view of these problems, other kinds of methods are devoted to explore the visual characters of roads.

Works based on the image aim to detect roads by their visual or geometric features, or some defined high-level knowledge. With the assumption that road regions often appear as thin, low–curvature, and high-contrast structures, various methods developed different line detectors for road extraction, such as morphological filters [22], line segment matching [23], Gibbs point [24], directional filters [25], Kalman filtering [26], and line primitive connection [27]–[30]. Knowledge-based methods aim to detect road regions by investigating high-level features. With the observation that low-level road extraction methods are fragmented, Steger *et al.* [31], [32] proposed to construct road network topology according to graph theory. In addition, benefitting from graph representation, Peteri and Ranchin [33] developed a road shape extraction scheme by defining the active contours. Poullis and You [34], employing Gabor filtering and tensor voting for geospatial feature inference classification, followed by orientation-based segmentation, extracted road centerlines to describe the road network. Inspired by this work, Grote *et al.* [35] extracted road networks, by integrating the radiometric and geometric features of road regions, and then, by constructing a subgraph, connected potential road segments to form the detection results. Based on the definition of pixelwised polygonal areas, Hu *et al.* [36] and Zhang *et al.* [37] explored the pixel footprint detector to extract road areas. By utilizing the LEGIONs, Yuan *et al.* [11] extracted the road automatically by clustering the well-aligned pixels according to their directionality.

More recently, image-based approaches tend to employ multistage [4], [5] or multimodel [6] schemes to further improve the detection performance. Ünsalan and Sirmacek [4] utilized spectral, shape, and gradient features to generate rough road primitives and then proposed a graph-based topology analysis scheme to refine the road map. Such approach is able to provide high-quality detection results with just simple interaction. Based on a pretrained spectral–spatial classifier, the latest road detection work [5] developed a road centerline extraction scheme, which is significantly improving the detection robustness, by hybridizing various road characters. Ziems *et al.* [6], by



(a) (b)

Fig. 2. Test examples in Cognitive Psychology. (a) Example to show the local directionality of salient structures. (b) Example to show the aperiodicity of salient structures.

employing different road detection methods, proposed a fusing scheme to combine the results of different models. Benefitting from the multimodel manner, such approach presents impressive robustness and detection performance.

However, as mentioned earlier, to address the limitation of the basic assumption about the road regions, recent works often require some information beyond the image to get high-quality detection results, such as simple interaction [4], 3-D road surface model [6], or pretrained classifiers [5]. Second, the multistage or multimodel scheme may produce more computational burden to reduce the efficiency of the whole system.

On the other hand, recent structure-aware image processing techniques [38]–[41] have already made some efforts for seeking more flexible metrics. These methods exhibit commendable strength, in dealing with structures free of contrast or spectral assumption, and provide interesting heuristics for road extraction.

III. Local Directional Structure Measurement

Despite the fact that spectral behaviors are not actually salient for some of the road regions, most people, as shown in Fig. 1(c), can still figure out these structures. Such an observation implies that, except for spectral behavior or contrast, the roadlike structures do follow some perception rules. Our goal is to extract roads by exploiting their visual saliency in terms of perception, while not relying on any assistant information such as GIS data or training samples.

Fig. 2(a) and (b) shows two famous test images in Cognitive Psychology. For Fig. 2(a), most people can figure out a nose contour, which can be attributed to its fine local directionality, although the short lines are messy [42], [43]. Each line in Fig. 2(b) has the same size and intensity, but the thwartwise and left-slanting lines are much more noticeable, because the perception of the other lines is suppressed by their social conformity [7]. Inspired by these perception observations, we develop the ADSM to describe roadlike structures. Specifically, such measurement is based on the following observations:

1) *Aperiodicity*. As opposed to texturelike elements, roadlike structures present less oscillation of intensity and do not distribute periodically and regularly in a local area.
2) *Local directionality*. Intensity variation of roadlike structures has more consistent local directionality than textures, in a local area.

Then, our ADSM is composed of mainly two parts: the aperiodicity measurement $\boldsymbol{L}_p$, which is designed to suppress texturelike periodic elements, such as residential area, and the directionality measurement $\boldsymbol{D}_p$, which is designed to analyze the local directionality of potential road structures. Next, the input image is preprocessed by Gaussian and bilateral filters. Since the results are not much sensitive to such filters, the parameters are set at fixed values. Specifically, radii are set at 5 for both filters; the σ_G for Gaussian filter is set at 3.0; the σ_s and σ_c for bilateral filter are set at 5.0 and 15.0, separately. With the filtered image, we elaborate on the details for formulating these measurements.

To distinguish dominant roadlike structures from the textures with similar contrast scales, inspired by a previous structure-aware image smoothing method in [39], we first employ the aperiodic measurement as

$$\boldsymbol{L}_p = \left| \sum_{q \in N(p)} w_{p,q} \cdot \nabla_x \boldsymbol{I}_q \right| + \left| \sum_{q \in N(p)} w_{p,q} \cdot \nabla_y \boldsymbol{I}_q \right| \quad (1)$$

where $w_{p,q}$ is defined as

$$w_{p,q} = \exp\left(-\frac{|x_p - x_q|^2 + |y_p - y_q|^2}{2\sigma^2} \right). \quad (2)$$

For pixels p and q, $\nabla_x \boldsymbol{I}_q$ and $\nabla_y \boldsymbol{I}_q$ represent the gradient components along the horizontal and vertical directions, respectively; $w_{p,q}$ is the Gaussian weight; $N(p)$ denotes the neighborhood of pixel p, and in our experiments, radius of $N(p)$ is set at a constant value of 5 (pixels) for images with 4-m resolution, since such value is tested sufficiently large for all the experiments. The periodic structures often show intense gradient oscillations; such measurement essentially suppresses the periodic texturelike elements by allowing counteraction of the gradients.

The local directionality $\boldsymbol{D}_p$ is derived from *structure tensor*, i.e., a sophisticated tool in image processing [44], [45]. It not only implies the intensity distribution at each pixel but also summarizes the predominant directions of the gradient magnitude in a local neighborhood. Formally, a structure tensor at a pixel p is defined as

$$\boldsymbol{S}_p = \begin{pmatrix} \sum\limits_{q \in N(p)} \langle \nabla_x \boldsymbol{I}_q, \nabla_x \boldsymbol{I}_q \rangle & \sum\limits_{q \in N(p)} \langle \nabla_x \boldsymbol{I}_q, \nabla_y \boldsymbol{I}_q \rangle \\ \sum\limits_{q \in N(p)} \langle \nabla_x \boldsymbol{I}_q, \nabla_y \boldsymbol{I}_q \rangle & \sum\limits_{q \in N(p)} \langle \nabla_y \boldsymbol{I}_q, \nabla_y \boldsymbol{I}_q \rangle \end{pmatrix} \quad (3)$$

where $N(p)$ is the neighborhood of a pixel p with a specified radius, $\langle \cdot, \cdot \rangle$ denotes the scalar product, and $\nabla_x(\cdot)$ and $\nabla_y(\cdot)$ are partial derivatives computed by using the Sobel operator. Mathematically, $\boldsymbol{S}_p$ is a positive semidefinite symmetric matrix, which has two eigenvalues $\lambda_{1,p} \geq \lambda_{2,p}$ with the corresponding eigenvectors $\vec{\boldsymbol{\eta}}_p$ and $\vec{\boldsymbol{\xi}}_p$.

The local directionality $\boldsymbol{D}_p$ is computed by the orthogonality distribution of eigenvectors $\vec{\boldsymbol{\xi}}$ corresponding to directions with minimum eigenvalues in a local neighborhood. Concretely, it is defined as

$$\boldsymbol{D}_p = \frac{\sum\limits_{q \in N(p)} \boldsymbol{A}_q \cdot \left| \langle \vec{\boldsymbol{\xi}}_q, \vec{\boldsymbol{\xi}}_p \rangle \right|}{\sum\limits_{q \in N(p)} \boldsymbol{A}_q}. \quad (4)$$

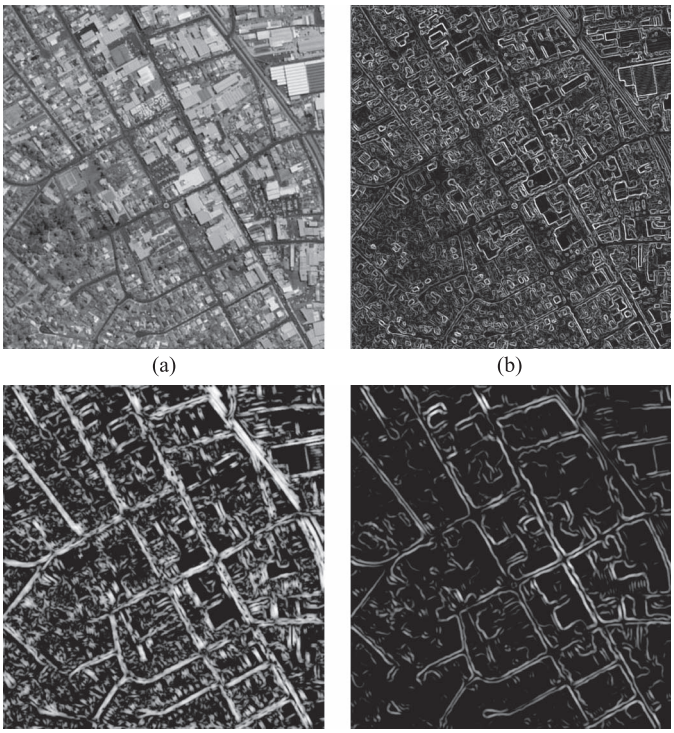


Fig. 3. Responses of gradient magnitude and ADSM. (a) Input image. (b) Gradient magnitude map calculated by the Sobel operator. (c) Visualized result of local directionality measurement. (d) Visualized result of ADSM.

The absolute value of scalar product $|\langle\vec{\boldsymbol{\xi}}_q, \vec{\boldsymbol{\xi}}_p\rangle|$ has a range of [0, 1], and it increases as the two vectors align closely, i.e., the angle between them approaches 0 or π, and equals 0 as the two vectors get orthogonal. To suppress the influence of the flattened area, pixels with high anisotropy are encouraged to contribute more; hence, a measurement $\boldsymbol{A}_q$ is applied as weight to better capture the primary direction. Specifically, it is defined by

$$\boldsymbol{A}_p = \frac{\lambda_{1,p} - \lambda_{2,p}}{\lambda_{1,p} + \lambda_{2,p}}. \tag{5}$$

This anisotropy measurement, which shows potential in capturing local nonuniformity of intensity variance, has been exploited since the pioneering work in [46]. $\boldsymbol{A}_p$ has a range of [0,1]; a larger value indicates structures with a predominant trend of intensity variation at the pixel.

With these definitions, the ADSM can be written as

$$\boldsymbol{M}_p = \boldsymbol{D}_p \cdot \frac{1 + \tanh\left(\lambda \cdot (\boldsymbol{L}_p - \boldsymbol{L}_{\text{mean}})\right)}{2} \tag{6}$$

where parameter λ is designed to tune the sharpness of the curve of function $\tanh(\cdot)$; in our experiments, it is set as the fixed value of 3.3, empirically. L_{mean} is the mean value of the aperiodic measurement map defined in (1). In (6), $\boldsymbol{D}_p$, with the range of [0,1], is designed to indicate potential roadlike structures with high local directionality. The normalized aperiodic measurement $\boldsymbol{L}_p$ is designed to further suppress the texture like areas; thus, regions such as building groups or residential areas would have a low response.

Fig. 3 shows an example of the ADSM, where Fig. 3(a) shows a complex urban remote sensing image; Fig. 3(b) shows the gradient magnitude map of Fig. 3(a) (acquired by the Sobel operator), where identification of the real road regions is difficult; Fig. 3(c) shows the visualized result of the local directionality measurement $\boldsymbol{D}$ defined by (4), where some visible residential areas are also involved; Fig. 3(d) shows the ADSM result defined by (6), where the potential road regions have higher responses while the interferences from residential areas are significantly suppressed.

Compared with previous road detection works, the proposed ADSM relies on the aperiodicity and directionality of road regions implied by perception theories, rather than the use of spectral behavior or intensity contrast. Benefitting from such characteristics, our system, without assistant information such as training samples or user interactions, is able to detect road regions (as highlighted in Fig. 1) with inconspicuous spectral features.

IV. ADSM-Guided Road Network Extraction

ADSM provides the probability that a pixel belongs to the edge of road regions. Guided by such a measurement, the detected road primitives construct a mask to describe the road network topology; then, in combination with some morphology operators, road regions can be extracted. Thus, our road network extraction algorithm includes mainly three parts: road primitive detection, topology description, and network extraction.

A. ADSM-Based Road Primitive Detection

Inspired by the work of Von Gioi *et al.* [47], our line segment detection scheme, by the guidance of ADSM, first generates line- support regions (regions whose pixels share roughly the same minimal gradient change orientation). The line segments are then defined by a fitted rectangle.

The line-support region is generated by the greedy algorithm, with a start point (the left top pixel in the image); the region orientation $\theta(\vec{\boldsymbol{\xi}}_R)$ is set to the weighted mean orientation of the minimal gradient change direction on each pixel, which is written as

$$\theta(\vec{\boldsymbol{\xi}}_R) = \theta\left(\frac{\sum_{p\in R} \boldsymbol{M}(p) \cdot \vec{\boldsymbol{\xi}}_p}{\sum_{p\in R} \boldsymbol{M}(p)}\right) \tag{7}$$

where $\boldsymbol{M}_p$ is the ADSM of a pixel p, and $\vec{\boldsymbol{\xi}}_p$ denotes the minimal intensity change direction at pixel p. Here, we use the orientation of the weighted average vector to denote the region orientation, where the weights are the ADSM values. Such weight is designed to suppress the contribution of the pixels in *smooth* or *texture* area. Then, the four connected pixels adjacent to the seed point are evaluated. The pixels with orientation similar to the mean direction of the region (i.e., the orientation difference $|\langle\vec{\boldsymbol{\xi}}_R, \vec{\boldsymbol{\xi}}_p\rangle|$ larger than a user-specified tolerance τ) are added to the region. The range of τ is [0,1], and a smaller value leads to "more" detected roads. In our experiment, it is set at the fixed value of 0.75; however, such parameter can be manually adjusted for different applications. After each iteration, the mean orientation of the region is updated, and the process is repeated until no new point can be added. We traverse the whole image from the left top pixel until each pixel is assigned a line-support region.

With the roughly defined line-supported regions, a rectangle is employed to fit each region to form the line primitives. Inspired by the idea of Kahn *et al.* [48], we use the weighted

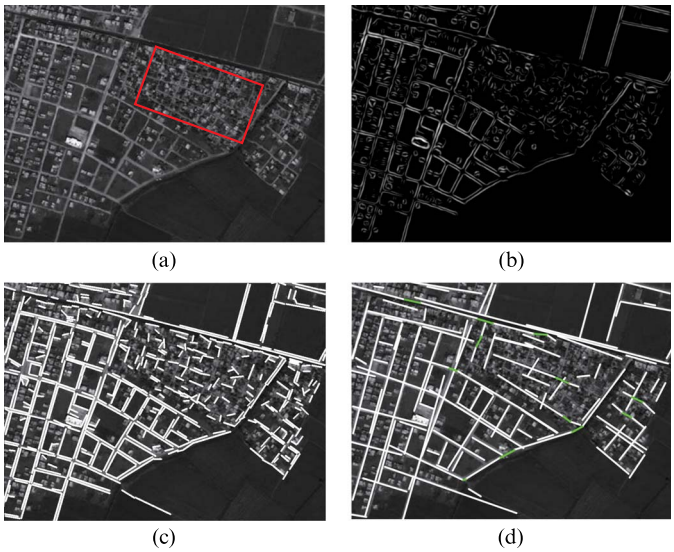


Fig. 4. Examples of the road mask construction. (a) Input image. (b) Response of ADSM. (c) Detected line segments. (d) Road region mask after consolidation and connection.

centroid to define the center of the rectangle, and the ADSM value is taken as the weight. The orientation is calculated by the direction of the mean eigenvector of the region [see (7)]. The length is chosen as the maximum length of the region, and the width is set as a fixed value to cover the line-support region. With the rectangular approximation of the line-support region, the validation scheme proposed in [47] is applied to confirm the line segments. Specifically, for each line segment, the number of aligned points is counted, and the segments which satisfy a threshold criterion based on their length and their number of aligned points are kept as valid detections.

Fig. 4 shows an example of the road mask construction; Fig. 4(a) shows the input remote sensing image where; within the highlighted area, it is a challenge to detect the road regions. Fig. 4(b) shows the visualized result of ADSM. Fig. 4(c) shows the detected line primitives (white lines) that indicate the potential road regions.

B. Road Network Extraction

The road primitives are scattered and messy, and some gaps are also observed. Thus, a road network topology refinement scheme is proposed to form a mask for better description of the entire road network topology. The refinement scheme consists of two steps: consolidation and connection.

Consolidation: Our consolidation scheme aims to reduce the false alarm. Specifically, for line segments, we find a target one with minimum *cost* to them in a predefined neighborhood and then merge them. Specifically, the merging scheme aims to find a pair of points with longest distance between the following four points: the terminal points of the longer line and the projection points from the shorter line to the longer line. Such process is repeated until there is no neighbors for each line segment in the neighborhood.

The definitions of *neighborhood* and *cost* are based on three criteria: 1) direction similarity evaluating whether two line segments are aligned closely in the same direction, which is calculated by the normalized vector of two line segments $\theta_s = \langle \vec{\xi_{l_1}} \cdot \vec{\xi_{l_2}} \rangle$, as shown in Fig. 5(a); 2) the distance perpendicular

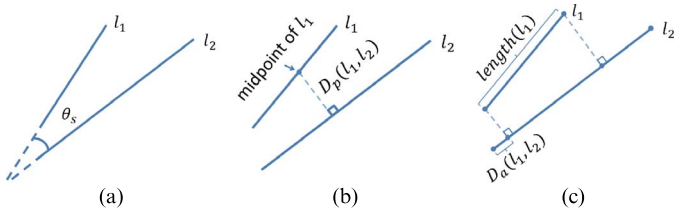


Fig. 5. Illustration of the consolidation criteria. (a) Angle definition between two lines. (b) Perpendicular distance definition between two lines. (c) Horizontal distance definition between two lines.

to the direction of the longer one of the two line segments, denoted by $D_p(l_1, l_2)$, which is calculated by the distance between the midpoint of the shorter line and its projection on the longer line, as shown in Fig. 5(b); 3) the normalized distance along the direction of the long line segment, denoted by $D_n(l_1, l_2) = D_a(l_1, l_2)/\text{length}(l_1)$, which is calculated by the ratio between the distance of the projection point (from the short to the long segments) and its nearest terminal point in the longer line segment (denoted by $D_a(l_1, l_2)$), and the length of the shorter line segment, as shown in Fig. 5(c).

Considering the aforementioned criteria, the neighborhood of the centerline segments is defined by the empirical thresholds $N(l_c) = \{(\theta_s, D_p, D_n) | \theta_s < \tau_s, D_p < \tau_{D_p}, D_n < \tau_{D_n}\}$, and the cost is written as

$$\boldsymbol{E}_{l_1,l_2} = \theta_s \cdot D_p(l_1, l_2) \cdot D_n(l_1, l_2). \tag{8}$$

The cost function is designed to favor the merger of the line segments with roughly the same direction and smaller distances of D_p and D_n. The parameters τ_s, τ_{D_p}, and τ_{D_n} are set at the experimental values $\tau_s = 20$ (degree), $\tau_{D_p} = 20$ (pixel) (with 4-m image resolution), and $\tau_{D_n} = 5$. Such configuration is sufficient for all our experiments. Details of the consolidation scheme are shown in Algorithm 1.

Algorithm 1 Line segments consolidation

Input: A set of line segments $l_k \in L$.
Output: Consolidated road mask lines $\tilde{l}_k \in \tilde{L}$.

```
1 L̃ ← L;
2 mergingFlag = TRUE;
3 while (mergingFlag = TRUE) do{
4   mergingFlag = FALSE
5   foreach line segment l̃_k ∈ L̃ do
6     if (N(l̃_k) ≠ φ) then
7       foreach neighbor l̄_k of l̃_k do
8         find l_n s.t. E_{l_n,l̃_k} = infE_{l̄_k,l̃_k};
9       end
10      mergeTwoLineSegments(l_n, l̃_k)
11      mergingFlag = TRUE
12    end
14  UpdateLineSegmentsSet(L̃)
13  end;
15 }
```

Connection: Inspired by previous tensor-voting-based approaches [49], [50], to fully cover the curving part of the roads with the consolidated line segments, adjacent line primitives are connected to form the road mask. More specifically, two lines are favored to be connected if they are: 1) near each other; 2) aligned closely in the same direction; 3) their terminal points are not connected with any other lines.

According to the aforementioned considerations, to determine which lines should be related, each pair of detected line segments is connected by an edge with a designed cost. Then, the desired connection is defined by a label set $l_e \in \{-1, 1\}$, which indicates whether an edge should be preserved ($l_e = 1$) or not ($l_e = -1$). Our goal is then to solve a binary labeling problem by minimizing the following energy function:

$$\arg\min_{l_{e_p}} \left(\sum_{e_p} \boldsymbol{O}_{l_{e_p}} + \alpha \sum_{e_q \in N(e_p)} \boldsymbol{S}_{l_{e_p}, l_{e_q}} \right) \tag{9}$$

where α, whose typical range is (1.0, 5.0), is the weight for in-between tradeoff of data term $\boldsymbol{O}$ and smoothness term $\boldsymbol{S}$.

The data term $\boldsymbol{O}$ evaluates whether an edge should be preserved, and it is defined by

$$\boldsymbol{O}_{l_{e_p}} = -\text{sign}(l_{e_p}) + \lambda \cdot L(e_p) \Big/ \left(\left| \langle \vec{e_p} \cdot \vec{l_1} \rangle \right| \cdot \left| \langle \vec{e_p} \cdot \vec{l_2} \rangle \right| + \epsilon \right) + C \cdot f \left(2 - \sum_{e_w \in V(e_p)} (l_{e_w} + 1) \right) \tag{10}$$

where $L(e_p)$ is the length of an edge; $\text{sign}(\cdot)$ is the standard sign function, i.e., $\text{sign}(x) = 1$ if $x > 0$, $\text{sign}(x) = -1$ if $x < 0$, and $\text{sign}(x) = 0$ is $x = 0$; $\vec{l_1}$ and $\vec{l_2}$ denote the normalized vectors of line segments connected by edge e_p; ϵ is a sufficient small constant to prevent zero division. λ, with a typical range of (0.005, 0.1), is the user-defined parameter denoting the relaxation of the connection. A smaller value indicates more connections between the road primitives. For the third term, C is a large enough constant, which is set at 999999.0 in our program; $V(e_p)$ is the set of edges sharing the same vertex; function $f(a)$ equals 0 if $a \geq 0$; otherwise, it equals to 1. Such term guarantees that no more than one of the edges sharing the same vertex is allowed to be marked as 1. The data term is designed to encourage shorter edges, which connect two roughly aligned line segments, to be preserved.

The smoothness term $\boldsymbol{S}$ is defined as

$$\boldsymbol{S}_{l_{e_p}, l_{e_q}} = \sum_{e_q \in N(e_p)} \frac{|l_{e_p} - l_{e_q}|}{d(e_p, e_q) + \epsilon} \tag{11}$$

where the function $d(e_p, e_q)$ denotes the distance of two edges, which is calculated by the distance between their middle points; ϵ is a sufficient small constant to prevent zero division; $N(e_p)$ is the K-nearest neighborhood of edge e_p. Of course larger size of neighborhood provides better results but will significantly increase the computational cost. Hence, an 8-nearest neighborhood is employed in our experiment for better balance between quality and efficiency. The smooth term $\boldsymbol{S}_{l_{e_p}, l_{e_q}}$, which encourages an edge and its close neighbors to have the same label, is designed for the spatial coherence.

The objective function of (9), which has the classical form of Markov random field problem, has been demonstrated solvable by the graph cut optimization in previous work [51]. Fig. 4(d) shows an example of the road mask after consolidation and connection, where the scattered road primitives are clustered to a long and continuous road mask, and as indicated by the green lines, gaps are also well filled.

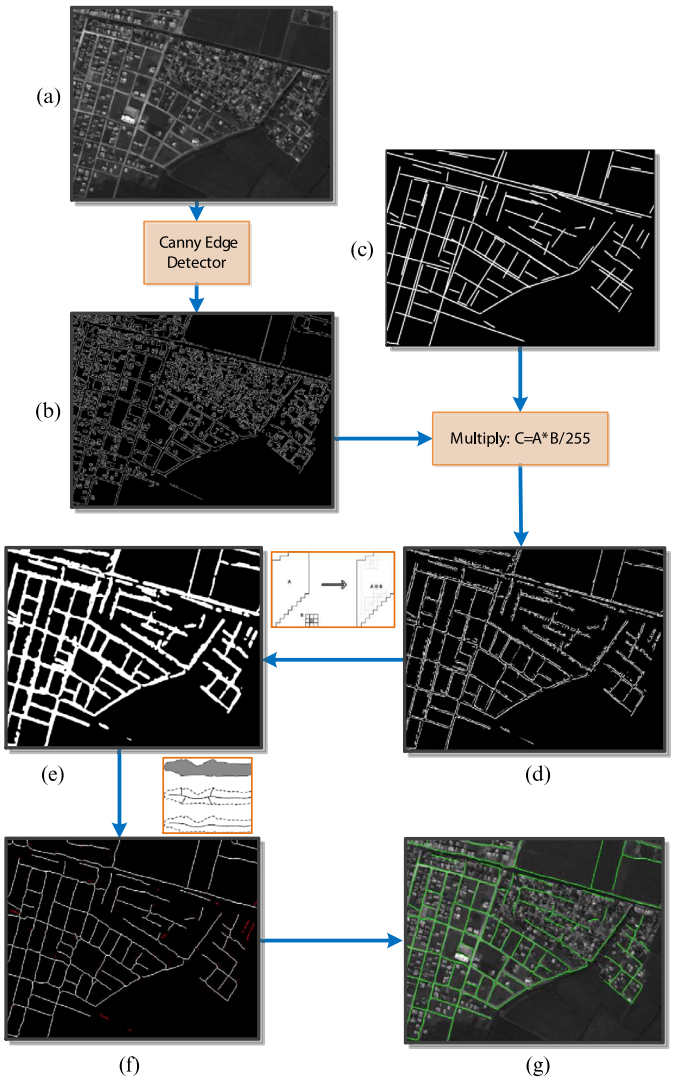


Fig. 6. Pipeline of the centerline extraction according to the road region mask. (a) Input image. (b) Canny edge map. (c) Visualized result of road region mask. (d) Rough road map acquired by overlapping (c) to (b). (e) Dilation version of (d). (f) Road network extraction result. (g) Visualized result of overlapping the network extraction map to the input.

Extraction: With the established road mask, we use some simple image processing operations to get the road information, as indicated by step results (a)–(g) in the pipeline shown in Fig. 6. Specifically, the Canny operator, because of its scalability and robustness [4], is first applied to extract a basic road map. The road network topology is described mainly by the mask; hence, to obtain sufficient potential roads, we can relax the edge detector parameters. In our experiments, the higher and lower thresholds of Canny operator are set at 80 and 50, respectively; the experimental result is shown in Fig. 6(b). The edge image is multiplied by the Gaussian blurred version of the constructed mask, as shown in Fig. 6(c), to form a rough road map, as shown in Fig. 6(d). Then, to extract the centerline, the rough road map is first dilated by the morphology operator, then binarized, as shown in Fig. 6(e), and thinned to form the skeleton, as shown in the white part of Fig. 6(f). Finally, according to connectivity analysis, the isolated short lines are removed, as highlighted in red in Fig. 6(f), to form the final pixelwised binary road map, as shown in Fig. 6(g). Such map

can be easily transformed to vector form by recent commercial software such as ARCGIS or Adobe Illustrator.

V. Results

Our approach is implemented based on C++ programming on a PC with a 3.2-GHz Intel Core i5-3200 CPU and 4-G RAM. To comprehensively evaluate the algorithm, several experiments are designed here. First, images from different satellite sensors, such as Ikonos, GeoEye, QuickBird, and aerial, are applied for testing, where various terrains, such as urban, rural region, mountain, etc., are involved. The corresponding road reference map is generated by a hand drawing method. We also evaluate our approach on the Pleiades-1A remote sensing image of an entire city of China (Shaoshan City), where the reference data are obtained by ground survey. Finally, the proposed method is compared with some of the latest road extraction works in both quality and efficiency, and the reference is provided by previous works.

On the parameter setting phase, to further adjust the results for specific application, values of the sharpness factor λ in (6) and tolerance τ in Section IV-A can be adjusted (lower values for higher completeness and higher values for higher correctness) for different types of results.

A. Evaluation on Images From Various Sensors

The evaluation is first applied on images from four sensors, including GeoEye, QuickBird, Ikonos satellites, and aerial images. Ten samples for each sensor are selected for testing. The resolutions of the GeoEye, QuickBird, Ikonos, and aerial images are 0.5 m, 0.5 m, 1 m, and 1 m, respectively. The high-resolution images show the terrain or objects in great detail, which may interfere in the detection of roads; therefore, similar to the previous work [4], all the test images are first appropriately downsampled to 4-m resolution.

In the evaluation phase, the ground truth data are manually generated; then, followed by the classic accuracy measurements [29], [52], three criteria are applied as follows:

$$\begin{aligned} \text{completeness} &= \frac{\text{TP}}{\text{TP}+\text{FN}} \\ \text{correctness} &= \frac{\text{TP}}{\text{TP}+\text{FP}} \\ \text{quality} &= \frac{\text{TP}}{\text{TP}+\text{FN}+\text{FP}} \end{aligned} \tag{12}$$

where TP, FN, and FP denote true positive, false negative, and false positive, respectively. For each sensor, the average quantitative evaluations over all ten test images are listed first. Then, for better comparison, we test our method on two additional examples, which have been applied in recent road extraction work [4].

Test on GeoEye Satellite Image: We have selected ten GeoEye images (including seven city region images, two rural region images, and one mountain region image, and most of them are about 1000 × 1000 patches) with 13 627 789 pixels, and the length of all the roads is about 480 km. The average detection results are listed in the second row of Table I. We also test our method on two additional examples, which are applied in [4] (corresponding results are listed in the last two rows of Table I); as shown in Figs. 7 and 8, these figures are the step-by-step results of our approach, where (a) is the input image; (b) is the visualized result of the ADSM, in which roadlike regions have higher responses; (c) is the road primitive detection result, in which the roadlike structures are well covered by the short line segments; (d) is the road network mask after *consolidation* and *connection*, where the short and scattered lines are clustered

TABLE I
Statistics of Remote Sensing Image Recording by GeoEye Satellite. The Second Row is the Average Result for Ten Tested Images; the Third and Fourth Rows Are Results for Two Examples in Figs. 7 and 8

Data	Completeness	Correctness	Quality
Average	0.8369	0.7991	0.6915
Geoeye-1	0.8522	0.8538	0.7437
Geoeye-2	0.8514	0.7068	0.6291

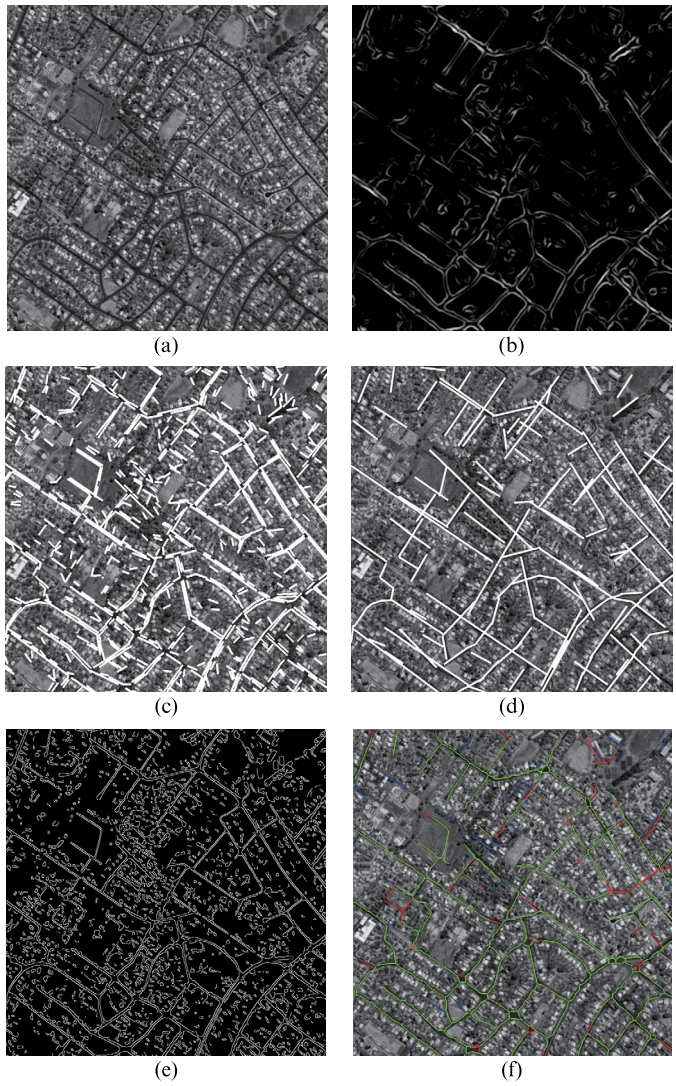


Fig. 7. Road extraction results on GeoEye image. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result. The green, blue, and red lines represent the true positive, false positive, and false negative detections.

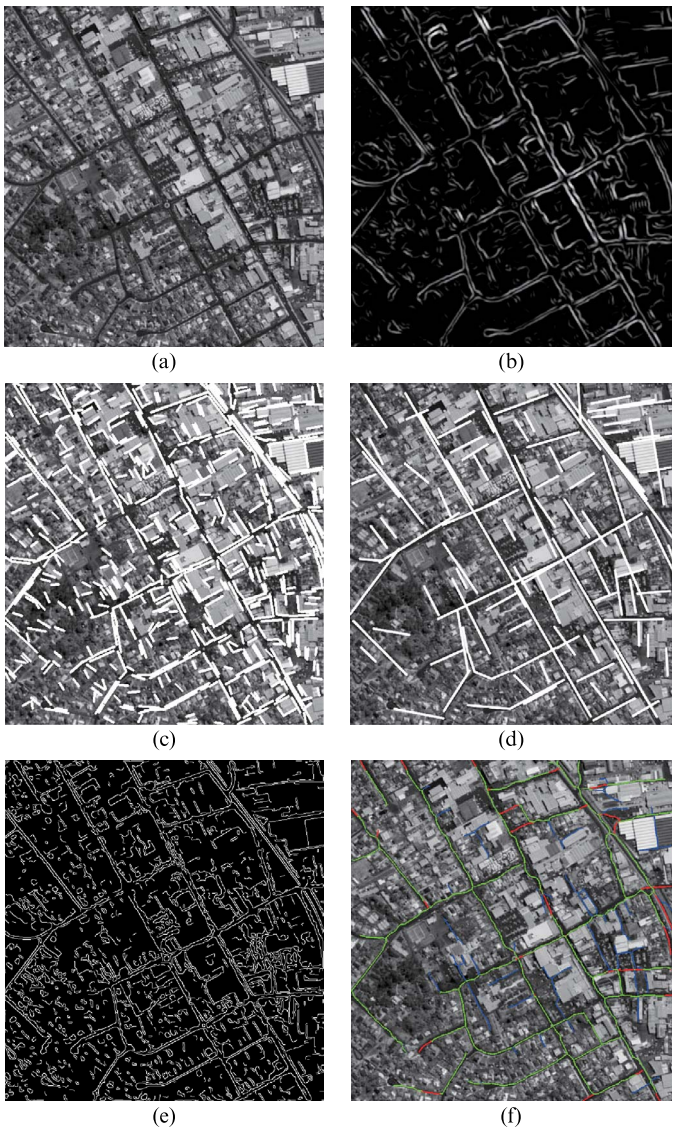


Fig. 8. Road extraction results on GeoEye image. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result. The green, blue, and red lines represent the true positive, false positive, and false negative detections.

TABLE II
STATISTICS OF REMOTE SENSING IMAGE RECORDING BY QUICKBIRD SATELLITE. THE SECOND ROW IS THE AVERAGE RESULT FOR TEN TESTED IMAGES; THE THIRD AND FOURTH ROWS ARE RESULTS FOR TWO EXAMPLES IN FIGS. 9 AND 10

Data	Completeness	Correctness	Quality
Average	0.7837	0.8154	0.6656
QuickBird-1	0.6389	0.9252	0.6076
QuickBird-2	0.9897	0.6638	0.6593

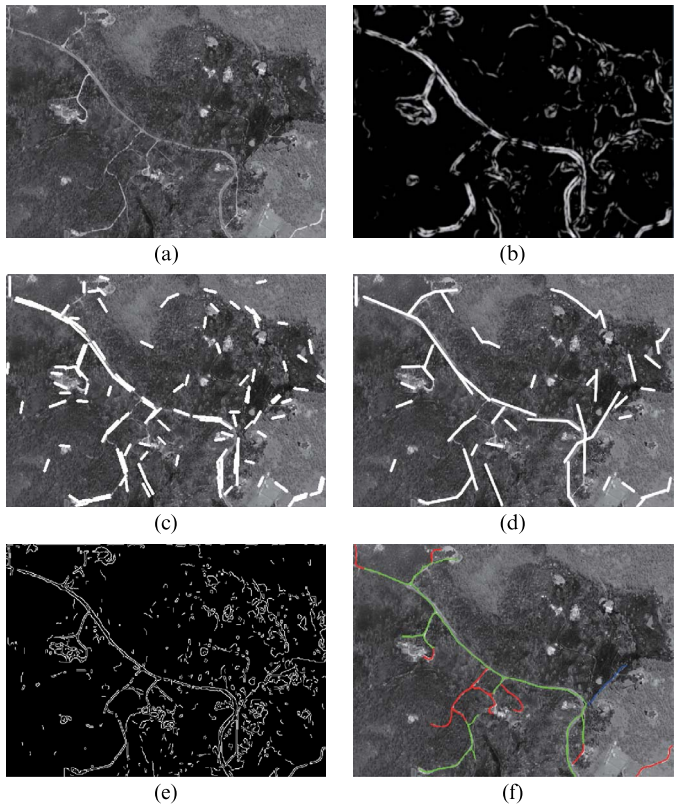


Fig. 9. Road extraction results on QuickBird image. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result. The green, blue, and red lines represent the true positive, false positive, and false negative detections.

and connected to acquire complete road network topology description. Then, combining with the Canny edge detector [as shown in (e)] along with some commonly used image processing operators, we can get the final road network map. As shown in (f), the green, blue, and red lines represent the true positive, false positive, and false negative detections.

Test on QuickBird Satellite Image: The selected ten QuickBird satellite images (including three mountain region images and three city region images) involve 10 516 297 pixels, and the length of all the roads is about 310 km. Corresponding results are listed in Table II. In previous work [4], tests using the QuickBird satellite sensor did not perform well because the images are JPEG compressed. In contrast, because our approach relies on the structures of the road region rather than the image details, the results are not sensitive to JPEG compression. The examples of this sensor are two rural region images; the road is a set of curvelike structures and is rather simple. In Fig. 9, occlusions are observed in some parts of the road, leading to lower *completeness*. In Fig. 10, the terrain boundary is misidentified as road; hence, the *correctness* is not satisfactory. To handle such a situation, some semantic information is helpful.

Test on Ikonos Satellite Image: The selected ten Ikonos satellite images (including six rural region, two city region, and two mountain region images) involve 11 638 971 pixels, and the length of all the roads is about 370 km. Limited by the license of the data, we do not present the comparison with the work in [4]. The *completeness*, *correctness*, and *quality* of this test, which show well robustness, are 0.8109, 0.7866, and 0.6648, separately, and there is not much difference comparing to the results of GeoEye and QuickBird sensors.

Test on Aerial Image: The selected ten aerial images (including five city region images, four rural region images, and one mountain region image) involve 16 948 971 pixels, and the

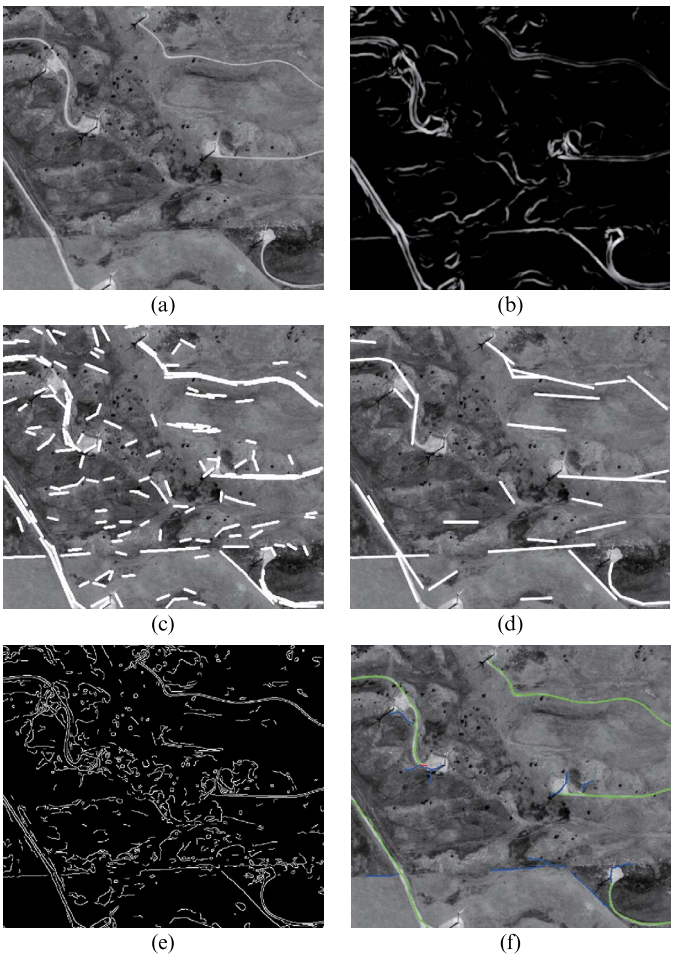


Fig. 10. Road extraction results on QuickBird image. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result. The green, blue, and red lines represent the true positive, false positive, and false negative detections.

TABLE III
AVERAGE STATISTICS OF DATA FROM DIFFERENT SENSORS

	Completeness	Correctness	Quality
Overall performance	0.8090	0.8078	0.6784
Variance	0.0011	0.0014	0.0004

length of all the roads is about 450 km. In addition, limited by the license of the data, we do not present the comparisons. Statistics of *completeness*, *correctness*, and *quality* are 0.7955, 0.8381, and 0.6895 separately. Because the samples of this sensor contain some rural region images, the roads suffer from more occlusion and shade; hence, the *completeness* is slightly lower than that for the GeoEye and Ikonos sensors, but the overall *quality* is fairly good.

Overall Performance: The average measurements of our approach over all the 44 images are listed in Table III. Generally, the performance of our approach is stable around different sensors and is not sensitive to resolution or the encoding compression. The overall variance is also satisfactory. The major factor influencing the effect is the occlusion or shading of road regions. To solve this problem, prior semantic information is consider to be involved in our system in the future.

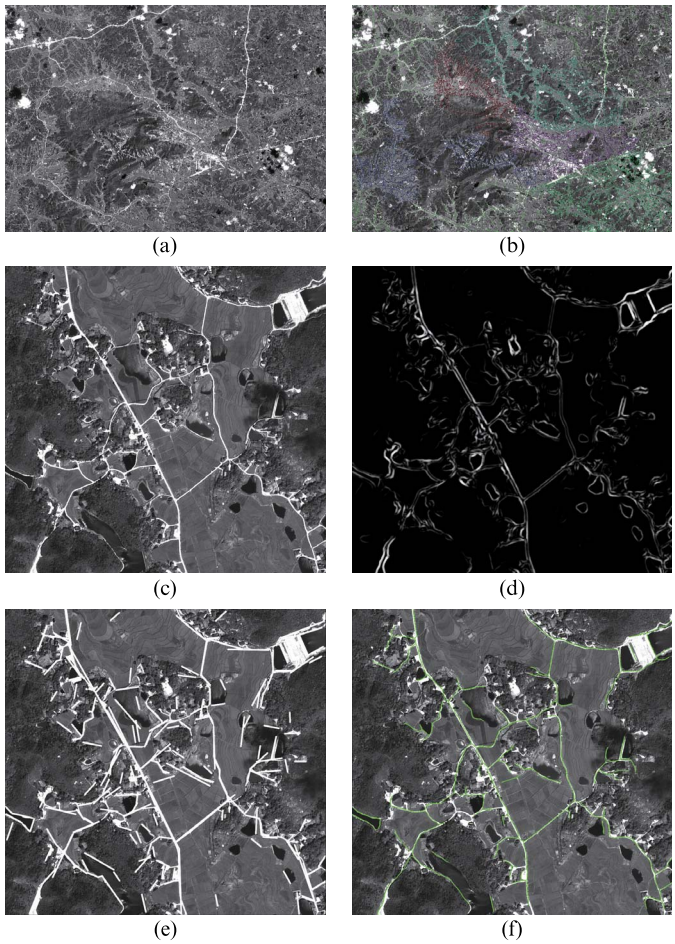


Fig. 11. Road extraction results on Pleiades-A image of Shaoshan City. (a) Input image. (b) Benchmark from ground survey. Different colors denote different regions of Shaoshan City. (c) Selected typical patch. (d) Response of the ADSM. (e) Result of the road region mask. (f) Our detection result.

B. Evaluation on the Remote Sensing Data of Shaoshan City

We also applied our method on the remote sensing image of Shaoshan City recorded by Pleiades-1A satellite with 0.5-m resolution [as shown in Fig. 11(a)]; the technical data of Pleiades-1A satellite are given in Table IV. The experiment are applied in divided image patches. Then, to merge the road network extraction results of different patches(i.e., a binary pixel map), for the edge pixel of the road in each patch, in its neighborhood (radius of 20 pixels for 4-m resolution), if there are other edge pixels in the adjacent patches, we calculate their gradient change direction, and if their directions are roughly the same, these two roads are considered continuous, and we connect them.

As shown in Fig. 11(b), the ground truth is acquired by ground survey. Shaoshan is a classical mountainous city covering about 250 km^2. The size of the satellite image is 28 648 $*$ 37 929 pixels. To test the performance of our method, the image is divided into 1000 $*$ 1000 patches with 30% overlap. Evaluations are applied in each patch; then, the results are combined together. The overall testing time is about 15 518 s, and the corresponding results are listed in the first row of Table V. Compared with the experiments in Section V-A, all of the three measurements decrease by different degrees, as shown in the second row of Table V. The major variation is the extraction

TABLE IV
TECHNICAL PARAMETERS OF PLEIADES-1A SATELLITE

	General features
Imagery Products	50-cm black and white
	50-cm color
	2-meter multispectral
	Bundle: 50-cm B&W and 2-meter multispectral
Spectral Bands	P: 480-830 nm
	Blue: 430-550 nm
	Green: 490-610 nm
	Red: 600-720 nm
	Near Infrared: 750-950 nm
Image Location Accuracy	With ground control points: 1m
	Without ground control points: 3m (CE90)
Imaging Capacity	Daily constellation capacity: 1,000,000 sq.km.
	Strip mapping (mosaic): 100 km x 100 km
	Stereo imaging: 20 km x 280 km
	Max. spots over 100 km x 200 km: 30
Imaging Swath	20 km at nadir
Revisit Interval	Daily

TABLE V
STATISTICS ON THE REMOTE SENSING IMAGE OF SHAOSHAN AREA

Test Data	Completeness	Correctness	Quality
Shaoshan city	0.7786	0.7135	0.5963
Comparison to Table III	-3.5%	-12.2%	-12.0%

correctness, which decreases more than 10%. We guess that the first reason is that, for experiments in Section V-A, we always select images with roads, to guarantee the *correctness* to some degree; however, for some image patches of the Shaoshan City image, there are no roads but only the mountains, thus leading to high false positive rate. On the other hand, the road reference of this test is acquired by ground survey rather than manual creation; hence, some of the hidden roads (due to shading or occlusion) are also involved, and this also slightly affects the performance of our method.

The spectral features of the road regions in this data set are quite good. However, we found that, in the experiments, such features are not that reliable. The spectral performance of the road can differ significantly from data set to data set. Hence, as mentioned earlier, our approach focuses mainly on the structural features of the road to obtain better adaptiveness. We also conducted an experiment, taking the spectral information of the road region into account. Here, a simple intensity threshold is applied to help improve the detection performance. The overall *completeness* increases about 5% (0.8198), and the *correctness* increases about 8% (0.7716).

Due to the licenses of these data, we just choose a typical area to visualize our road extraction results, as shown in Fig. 11(c)–(f), where Fig. 11(c) shows part of an image patch, Fig. 11(d) shows the visualized result of our ADSM measurement, Fig. 11(e) shows the generated road network mask, and Fig. 11(f) shows the extraction result. Limited by the copyrights, we are not allowed to visualized the ground truth in the image. The measurements of *completeness*, *correctness*, and *quality* of this patch are 0.8519, 0.8377, and 0.7311, respectively.

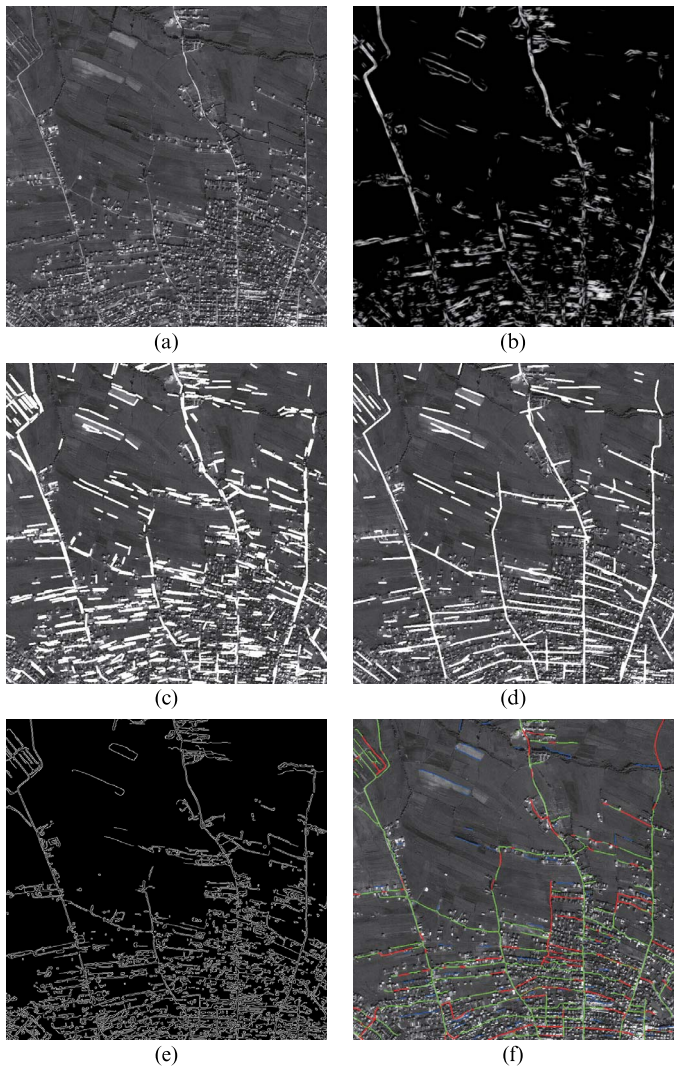


Fig. 12. Our extraction results on Ikonos1-Sub1. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result.

C. Comparison

To compare our road extraction approach with the other related works, evaluations on the aspects of *quality* and *efficiency* are applied.

Quality Evaluation: For the quality evaluation, we tested our methods on three classic Ikonos satellite images provided by Prof. H. Mayer. In their previous work [53], road centerline detection results of six different methods on three Ikonos satellite images (Ikonos1-Sub1, Ikonos3-Sub1, and Ikonos3-Sub2) are carefully evaluated, and the quantitative measurements are also collected. Based on this, many followed road extraction works reported their results on these images; thus, our quality evaluation is tested on the same data set. The version of these images that we obtained is of grayscale with 4-m resolution. Although the resolution is not very high, it is sufficient for our testing.

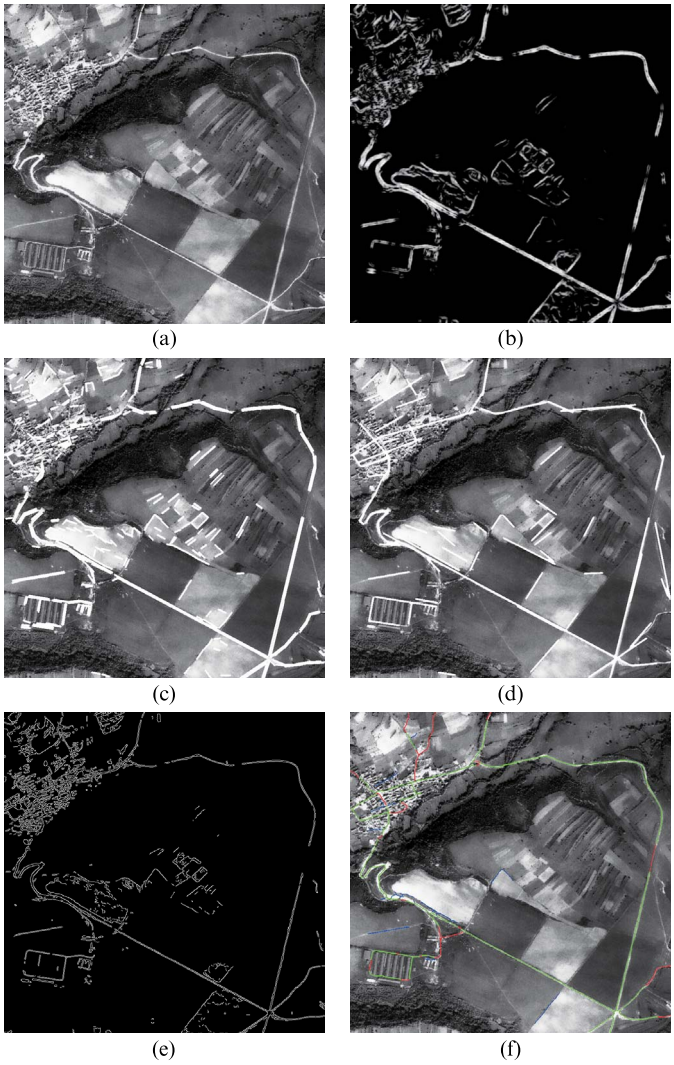


Fig. 13. Our extraction results on Ikonos3-Sub1. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result.

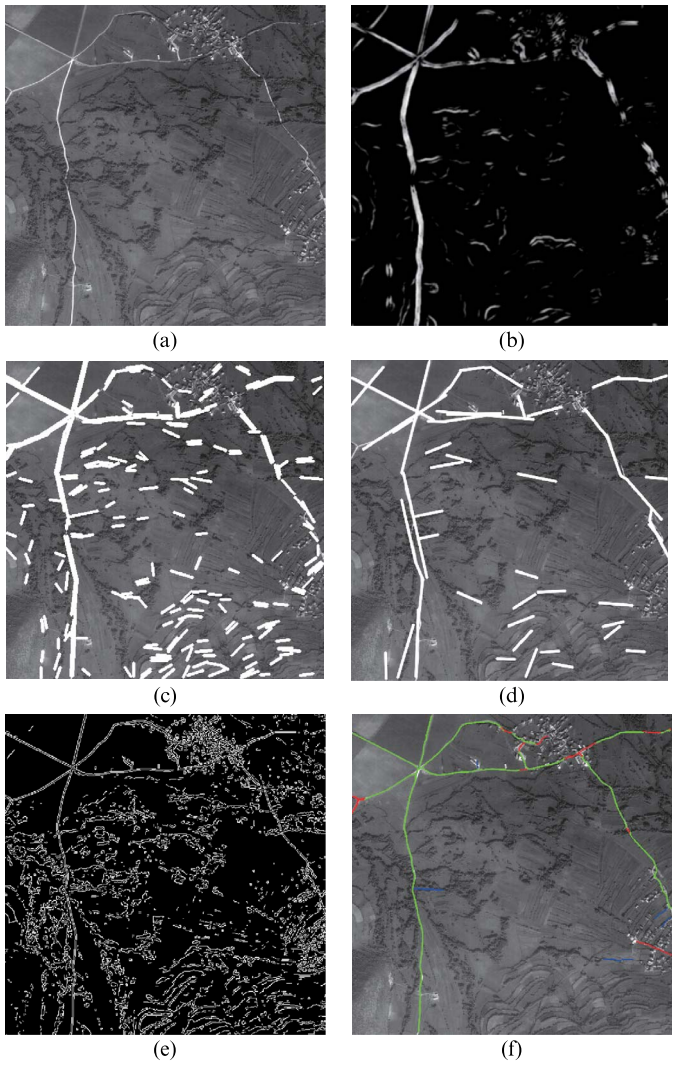


Fig. 14. Our extraction results on Ikonos3-Sub2. (a) Input image. (b) Response of ADSM. (c) Result of the detected line segments. (d) Road region mask. (e) Edge detection by Canny operator. (f) Our extraction result.

Figs. 12–14 show the results of our methods on these images. In each of these three figures, (a) is the input, (b) is the visualized result of the ADSM, (c) is the line segment detection result, (d) is the road network mask, (e) is the Canny edge detector, and our extraction result is shown in (f). For the first example (image Ikonos1-Sub1), the road network is rather complicated, which is a very challenging example for road detection. Our method also misses some of the roads, but the correctness is basically satisfactory. For the other two examples (image Ikonos3-Sub1 and Ikonos3-Sub2), our method performs fairly well.

The quantitative measurement of our approach along with the reports of ten other related works are listed in Table VII, where the highest two values are emphasized. Based on the definition of *quality* in (12), we also calculate such measurement of each method for a comprehensive comparison. Specifically, the *quality* is calculated by the following equation:

$$\text{quality} = \frac{\text{comp} \cdot \text{corr}}{\text{comp} + \text{corr} - \text{comp} \cdot \text{corr}}. \tag{13}$$

The results show that, in most cases, among the ten methods, the performance of our approach is in the top two. We also calculate the relative difference between the best method and ours. Corresponding results are shown in the last row of the table. For image Ikonos1-Sub1, our method has the best performance in either *completeness* or *correctness*. For image Ikonos3-Sub1, our *correctness* and *quality*, which ranked second, are slightly lower than those in the work of Bacher [54]. For image Ikonos3-Sub2, we have the third highest *completeness* and the highest *correctness* and *quality*. In general, the performance of our approach is fairly stable and has balanced *completeness* and *correctness*. In Mayer's work [53], values of 60% and 75% for completeness and correctness are recommended as the baseline for practice. In this respect, our method passes both the thresholds for all the images.

TABLE VI
TIME PERFORMANCE COMPARISON OF DIFFERENT METHODS

Method	Implementation Environment	Time (s)
Unsalan et al. [4]	3.2GHz Intel Core i5-3200 CPU	113.68
Shi et al. [5]	3.2GHz Intel Core i5-3200 CPU	78.92
Our approach	3.2GHz Intel Core i5-3200 CPU	13.17

TABLE VII
PERFORMANCE COMPARISON OF DIFFERENT METHODS

Method	Ikonos1-Sub1			Ikonos3-Sub1			Ikonos3-Sub2		
	Completeness	Correctness	Quality	Completeness	Correctness	Quality	Completeness	Correctness	Quality
Bacher [54]	0.34	0.66	0.29	0.81	**0.87**	**0.72**	0.86	0.89	**0.78**
Beumier [55]	0.48	0.69	0.39						
Gerke_W [28]	0.27	0.41	0.19	0.80	0.67	0.57	0.75	0.52	0.44
Gerke_WB [56]	0.19	0.49	0.16	0.68	0.75	0.55	0.71	0.84	0.63
Hedman [28]	0.31	0.51	0.24	0.77	0.78	0.63	0.85	**0.91**	**0.78**
Malpica [57]	0.25	**0.74**	0.23	0.60	0.79	0.52	0.60	0.89	0.56
Zhang [58]	0.56	0.41	0.31	0.72	0.35	0.31	0.70	0.34	0.30
Yuan [11]	0.59	0.65	**0.45**	**0.83**	0.71	0.62	**0.89**	0.62	0.58
Ünsalan [4]	**0.63**	0.40	0.32	0.73	0.81	0.62	0.73	0.90	0.68
Shi [5]	0.34	0.63	0.28	0.77	0.65	0.54	**0.94**	0.62	0.60
Proposed	**0.67**	**0.78**	**0.56**	**0.81**	**0.82**	**0.69**	0.85	**0.91**	**0.78**
Difference	+6.3%	+5.4%	+24.4%	-2.4%	-6.1%	-4.3%	-11.7%	0	0

Efficiency Evaluation: Our implementation is based on C++ programming. For an image size of 1000×1000, our approach takes about 6 s for whole processing. The major bottleneck of our algorithm lies in the step of line segment consolidation and connection Section IV-B, which consumes almost 50% of the processing time. The other parts of our work are very efficient and have a linear complexity $\boldsymbol{O}(kn)$. We also compare our time performance with the latest road extraction works [4], [5], where all the methods are programmed with MATLAB 2013b (for a fair comparison, we also prepare a MATLAB version code for testing) and tested on ten images with sizes of $1000 * 1000$. The average running times along with the implementation environments are listed in Table VI. Time cost of our approach is about 16.7% and 11.6% of previous works [4] and [5].

VI. CONCLUSION

We have presented a novel road network extraction approach, which is able to produce stable and efficient detection results on data from various satellite sensors. The qualified adaptivity stems from a robust local directional aperiodic road structure measurement, which is inspired by Cognitive Psychology. Such a measurement combines the main properties of the local directionality and aperiodicity for spectral-behavior- and contrast-independent road structure characterization. Guided by such measurement, we designed a road region mask construction scheme to analyze the road network topology. Then, in combination with some common morphology operators, road centerlines are able to be extracted. Experimental results demonstrate the effectiveness and efficiency of our approach to produce high-quality road extraction results (see Table VII).

On the other hand, in our experiments, it is found that most of the errors occurred in road regions with shade or occlusion, since both the spectral or geometric features were affected in different degrees. Such a case can be difficult for purely computer-vision-based solutions. Thus, in the future, with the aid of semantic information or training samples, we plan to specifically focus on the analysis of these cases.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] V. Mnih and G. E. Hinton, "Learning to detect roads in high-resolution aerial images," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 210–223.
[2] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 5, pp. 530–549, May 2004.
[3] A. Baumgartner, C. Steger, H. Mayer, and W. Eckstein Lehrstuhl Fur, "Multi-resolution, semantic objects, and context for road extraction," in *Semantic Modeling for the Acquisition of Topographic Information From Images and Maps*. Basel, Switzerland: Birkhauser, pp. 140–156, 1997.
[4] C. Ünsalan and B. Sirmacek, "Road network detection using probabilistic and graph theoretical methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4441–4453, Nov. 2012.
[5] W. Z. Shi, Z. L. Miao, and J. Debayle, "An integrated method for urban main-road centerline extraction from optical remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3359–3372, Jun. 2014.
[6] M. Ziems, U. Breitkopf, C. Heipke, and F. Rottensteiner, "Multiple-model based verification of road data," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. I-3, pp. 329–334, 2012.
[7] G. Kanizsa, *Organization in Vision: Essays on Gestalt Perception*. New York, NY, USA: Praeger, 1979.
[8] J. B. Mena, "State of the art on automatic road extraction for GIS update: A novel classification," *Pattern Recognit. Lett.*, vol. 24, no. 16, pp. 3037–3058, 2003.
[9] C. Unsalan and K. L. Boyer, "A system to detect houses and residential street networks in multispectral satellite images," *Comput. Vis. Image Understand.*, vol. 98, no. 3, pp. 423–461, 2005.
[10] S. Das, T. T. Mirnalinee, and K. Varghese, "Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3906–3931, Oct. 2011.
[11] J. Yuan, D. Wang, B. Wu, L. Yan, and R. Li, "LEGION-based automatic road extraction from satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4528–4538, Nov. 2011.
[12] J. D. Wegner, J. A. Montoya-Zegarra, and K. Schindler, "A higher-order CRF model for road network extraction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 1698–1705.
[13] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. 1, pp. S110–S122, 2009.
[14] W. Su *et al.*, "Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery," *Int. J. Remote Sens.*, vol. 29, no. 11, pp. 3105–3117, 2008.

[15] C. W. Emerson, N. S.-N. Lam, and D. A. Quattrochi, "A comparison of local variance, fractal dimension, and Moran's I as aids to multispectral image classification," *Int. J. Remote Sens.*, vol. 26, no. 8, pp. 1575–1588, 2005.

[16] T. Peng, I. H. Jermyn, V. Prinet, and J. Zerubia, "Incorporating generic and specific prior knowledge in a multiscale phase field model for road extraction from VHR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 1, no. 2, pp. 139–146, Jun. 2008.

[17] R. Bellens *et al.*, "Improved classification of VHR images of urban areas using directional morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2803–2813, Oct. 2008.

[18] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

[19] D. Tuia, F. Pacifici, M. Kanevski, and W. J. Emery, "Classification of very high spatial resolution imagery using mathematical morphology and support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3866–3879, Nov. 2009.

[20] C. L. L. Hendriks, "Constrained and dimensionality-independent path openings," *IEEE Trans. Image Process. A Pub. IEEE Signal Process. Soc.*, vol. 19, no. 6, pp. 1587–1595, Jun. 2010.

[21] J. Debayle and J.-C. Pinoli, "General adaptive neighborhood image processing," *J. Math. Imag. Vis.*, vol. 25, no. 2, pp. 245–266, 2006.

[22] A. Katartzis, H. Sahli, V. Pizurica, and J. Cornelis, "A model-based approach to the automatic extraction of linear features from airborne images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 9, pp. 2073–2079, Sep. 2001.

[23] W. Shi and C. Zhu, "The line segment match method for extracting road network from high-resolution satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 2, pp. 511–514, Feb. 2002.

[24] R. Stoica, X. Descombes, and J. Zerubia, "A Gibbs point process for road extraction from remotely sensed images," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 121–136, May 2004.

[25] P. Gamba, F. Dell'Acqua, and G. Lisini, "Improving urban road extraction in high-resolution images exploiting directional filtering, perceptual grouping, and simple topological concepts," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 387–391, Jul. 2006.

[26] S. Movaghati, A. Moghaddamjoo, and A. Tavakoli, "Road extraction from satellite images using particle filtering and extended Kalman filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2807–2817, Jul. 2010.

[27] J. Yang and R. S. Wang, "Classified road detection from satellite images based on perceptual organization," *Int. J. Remote Sens.*, vol. 28, no. 20, pp. 4653–4669, Oct. 2007.

[28] C. Wiedemann and S. Hinz, "Automatic extraction and evaluation of road networks from satellite imagery," *Int. Archives Photogramm. Remote Sens.*, pp. 95–100, 1999.

[29] C. Wiedeman and H. Ebner, "Automatic completion and evaluation of road networks," *ISPRS Int. Archives of Photogramm., Remote Sens. Spatial Inf.*, vol. 33, no. B3/2, pp. 976–986, 2000.

[30] S. Hinz and C. Wiedemann, "Increasing efficiency of road extraction by self-diagnosis," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 12, pp. 1457–1466, 2004.

[31] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 2, pp. 113–125, Feb. 1998.

[32] C. Steger, H. Mayer, and B. Radig, "The role of grouping for road extraction," *Monte Verit*, 1997, to be published.

[33] R. Peteri and T. Ranchin, "Automated road network extraction using collaborative linear and surface models," in *Proc. MAPPS/ASPRS*, 2006, pp. 1–7.

[34] C. Poullis and S. You, "Delineation and geometric modeling of road networks," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 2, pp. 165–181, 2010.

[35] A. Grote, C. Heipke, and F. Rottensteiner, "Road network extraction in suburban areas," *Photogramm. Rec.*, vol. 27, no. 137, pp. 8–28, 2012.

[36] J. Hu, A. Razdan, J. C. Femiani, M. Cui, and P. Wonka, "Road network extraction and intersection detection from aerial images by tracking road footprints," *Int. J. Remote Sens.*, vol. 45, no. 12, pp. 4144–4157, 2007.

[37] J. Zhang, X. Lin, Z. Liu, and J. Shen, "Semi-automatic road tracking by template matching and distance transformation in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 23, pp. 8331–8347, May 2011.

[38] K. Subr, C. Soler, and F. Durand, "Edge-preserving multiscale image decomposition based on local extrema," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 147:1–147:9, Dec. 2009.

[39] L. Xu, Q. Yan, Y. Xia, and J. Jia, "Structure extraction from texture via relative total variation," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 139:1–139:10, Nov. 2012.

[40] L. Karacan, E. Erdem, and A. Erdem, "Structure-preserving image smoothing via region covariances," *ACM Trans. Graph.*, vol. 32, no. 6, p. 176, 2013.

[41] Y. Zang, H. Huang, and L. Zhang, "Efficient structure-aware image smoothing by local extrema on space-filling curve," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 9, pp. 1253–1265, Sep. 2014.

[42] N. Petkov and M. A. Westenberg, "Suppression of contour perception by band-limited noise and its relation to nonclassical receptive field inhibition," *Biol. Cybern.*, vol. 88, no. 3, pp. 236–246, 2003.

[43] Y. Zang, H. Huang, and L. Zhang, "Guided adaptive image smoothing via directional anisotropic structure measurement," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 2, pp. 1015–1027, Sep. 2015.

[44] J. E. Kyprianidis and H. Kang, "Image and video abstraction by coherence-enhancing filtering," in *Comput. Graph. Forum*, vol. 30, no. 2, pp. 593–602, 2011.

[45] H. Kang, S. Lee, and C. K. Chui, "Flow-based image abstraction," *IEEE Trans. Vis. Comput. Graph.*, vol. 15, no. 1, pp. 62–76, Jan./Feb. 2009.

[46] W. Förstner and E. Gülch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *Proc. ISPRS Intercommission Conf. Fast Process. Photogramm. Data*, 1987, pp. 281–305.

[47] R. G. Von Gioi, J. Jakubowicz, J. M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, Apr. 2010.

[48] P. Kahn, L. Kitchen, and E. M. Riseman, "A fast line finder for vision-guided robot navigation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 11, pp. 1098–1102, Nov. 1990.

[49] C. Poullis, "Tensor-cuts: A simultaneous multi-type feature extractor and classifier and its application to road extraction from satellite images," *ISPRS J. Photogramm. Remote Sens.*, vol. 95, no. 9, pp. 93–108, 2014.

[50] T. P. Wu, S. K. Yeung, J. Jia, C. K. Tang, and G. Medioni, "A closed-form solution to tensor voting: Theory and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1482–1495, Aug. 2012.

[51] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 109–131, Nov. 2006.

[52] C. Wiedemann, C. Heipke, H. Mayer, and O. Jamet, "Empirical evaluation of automatically extracted road axes," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 1998, pp. 172–187.

[53] H. Mayer *et al.*, "A test of automatic road extraction approaches," *Remote Sens. Spatial Inf. Sci. Int. Archives Photogramm.*, vol. 36, no. 3, pp. 209–214, 2006.

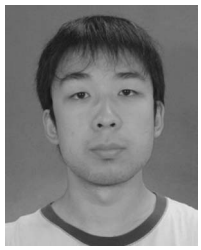
[54] U. Bacher and H. Mayer, "Automatic road extraction from multispectral high resolution satellite images," *Remote Sens. Spatial Inf. Sci. Int. Archives Photogramm.*, vol. 36, no. 3, pp. 29–34, 2005.

[55] V. Lacroix and M. Acheroy, "Feature extraction using the constrained gradient," *ISPRS J. Photogramm. Remote Sens.*, vol. 53, no. 2, pp. 85–94, 1998.

[56] A. Baumgarner, C. Steger, H. Mayer, W. Eckstein, and H. Ebner, "Automatic road extraction based on multi-scale, grouping, and context," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 2, pp. 777–785, 1999.

[57] J. B. Mena and J. A. Malpica, "An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery," *Pattern Recognit. Lett.*, vol. 26, pp. 1201–1220, 2005.

[58] Q. Zhang and I. Couloigner, "Benefit of the angular texture signature for the separation of parking lots and roads on high resolution multi-spectral imagery," *Pattern Recognit. Lett.*, vol. 27, pp. 937–946, 2006.



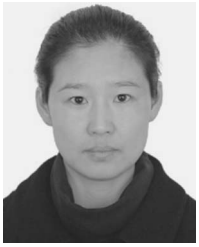
Yu Zang received the B.S. and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 2008 and 2014, respectively.

He is currently a Research Assistant Professor with the School of Information Science and Technology, Xiamen University, Xiamen, China. His main researches include remote sensing image processing, computer vision and graphics, and mobile LiDAR data analysis.

Cheng Wang (M'07) received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2002.

He is currently a Professor with Xiamen University, Xiamen, China, where he is also the Associate Dean of the School of Information Science and Technology. He has authored more than 80 papers. His current research interests include remote sensing image processing, mobile LiDAR data analysis, and multisensor fusion.

Dr. Wang is a Council Member of the Chinese Society of Image and Graphics and a member of Society of Photo-Optical Instrumentation Engineers (SPIE) and IEEE Geoscience and Remote Sensing Society (GRSS). He is the Cochair of International Society of Photogrammetry and Remote Sensing Work Group I/3 (ISPRS WG I/3).



Liujuan Cao (M'11) received the Ph.D. degree in computer applied technology from the Harbin Engineering University, Harbin, China, in 2013.

She is currently an Assistant Professor with the School of Information Science and Technology, Xiamen University, Xiamen, China. Her research interests include digital vector data security, large-scale image retrieval, and remote sensing image processing.

Yao Yu received the B.E. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2013. She is currently working toward the Ph.D. degree with the School of Information Science and Technology, Xiamen University, Xiamen, China.

Her major research direction includes remote sensing image processing and road network extraction.

Jonathan Li (M'00–SM'11) received the Ph.D. degree in geomatics engineering from the University of Cape Town, Cape Town, South Africa.

He is currently a Professor with the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of the Ministry of Education, School of Information Science and Engineering, Xiamen University, China. He is also a Professor with the Faculty of Environment, University of Waterloo, Waterloo, ON, Canada, where he is the Head of the GeoSTARS Laboratory. He has coauthored more than 300 publications, over 100 of which were published in refereed journals, including IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (TITS), IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (GRSL), *ISPRS Journal of Photogrammetry and Remote Sensing (JPRS)*, *IJRS*, *Photogrammetric Engineering & Remote Sensing (PE&RS)*, and *Remote Sensing of Environment (RSE)*. His current research interests include information extraction from mobile LiDAR point clouds and from earth observation images. Dr. Li is the Chair of the International Society of Photogrammetry and Remote Sensing Work Group I/Va (ISPRS WG I/Va) on Mobile Scanning and Imaging Systems for 2012–2016, the Vice Chair of the International Cartographic Association (ICA) Commission on Mapping from Remote Sensor Imagery for 2011–2015, and the Vice Chair of the International Federation of Surveyors (FIG) Commission on Hydrography for 2015–2018.