An Integrated Method for Urban Main-Road Centerline Extraction From Optical Remotely Sensed Imagery

Wenzhong Shi, Zelang Miao, and Johan Debayle

***Abstract*—Road information has a fundamental role in modern society. Road extraction from optical satellite images is an economic and efficient way to obtain and update a transportation database. This paper presents an integrated method to extract urban main-road centerlines from satellite optical images. The proposed method has four main steps. First, general adaptive neighborhood is introduced to implement spectral–spatial classification to segment the images into two categories: road and nonroad groups. Second, road groups and homogeneous property, measured by local Geary's C, are fused to improve road-group accuracy. Third, road shape features are used to extract reliable road segments. Finally, local linear kernel smoothing regression is performed to extract smooth road centerlines. Road networks are then generated using tensor voting. The proposed method is tested and subsequently validated using a large set of multispectral high-resolution images. A comparison with several existing methods shows that the proposed method is more suitable for urban main-road centerline extraction.**

***Index Terms*—General adaptive neighborhood (GAN), local Geary's C, local linear kernel smoothing regression, optical remotely sensed images, shape feature, spectral–spatial classification, tensor voting, urban main-road centerline extraction.**

I. Introduction

SUCCESSFUL urban design, navigation, and image registration owe much to road information collected from optical remotely sensed images. With the aid of computers, it is possible to reduce a lot of labor and time when obtaining transportation data. Various road extraction algorithms have been proposed over the past decades. Mena [1] and Das *et al.* [2] presented overviews of road detection methods in this area. Quackenbush [3] gave a review of linear feature extraction from imagery which can be used for road extraction. Automatic road detection method tests were devised by Mayer *et al.* [4]. Based on the level of road knowledge used, Poullis and You [5] classified road detection methods into three categories: 1) *pixel-based*; 2) *region-based*; and 3) *knowledge-based*.

The *pixel-based* methods depend on the information obtained from the pixels. Line [6], [7], edge [8], and ridge [9], [10] detectors are used to extract potential road points. Road points are then connected to produce road segments and also used as input to a higher level processing phase. Other pixel-level information, such as directional adaptive filters [11], magnitude, and orientation [5], is also explored for road extraction. The *region-based* methods first segment or classify the imagery into regions, followed by a rule to refine extracted road networks. In this field, Shi and Zhu [12] proposed a line segment match rule to extract urban road networks. Their method first segmented imagery into a binary map by simple thresholding. The road network was then extracted by the line segment match rule and morphological processing. Song and Civco [13] used a two-step-based method to detect road networks. In their method, a support vector machine (SVM) was used to classify the imagery. Yuan *et al.* [14] proposed automatic extraction of road networks from satellite imagery based on locally excitatory globally inhibitory oscillator networks.

A multistage framework was designed to extract road networks based on probabilistic SVMs and salient features [2]. A semiautomated road detection method, which includes directional morphological enhancement, directional segmentation, and thinning, was proposed by Chaudhuri *et al.* [15]. Huang and Zhang [16] proposed the detection of road centerlines from high-resolution images by integrating multiscale structural features and SVM. The *knowledge-based* methods use higher level information to detect roads. Poullis and You [5] presented a vision-based system for automatic road detection. To achieve this, Gabor filtering and tensor voting were used for geospatial feature inference classification; orientation-based segmentation, using graph cuts, was then used to extract road features, followed by road centerline extraction and tracking to produce a road network. Similarly, Grote *et al.* [17] integrated radiometric and geometric features to extract road networks. The normalized cut algorithm was used to segment the images; the segments were then grouped. Road segments were extracted from the groups and then generated subgraph. Context information was also used to eliminate false connections in the subgraphs.

Manuscript received April 29, 2013; revised June 27, 2013; accepted July 1, 2013. Date of publication July 24, 2013; date of current version March 3, 2014. This work was supported in part by the National Natural Science Foundation of China under Grants 41201451 and 40901214, by the Ministry of Science and Technology of China under Projects 2012BAJ15B04 and 2012AA12A305, by the Research Grants Council, Hong Kong, under Grant PolyU 5249/12E, and by The Hong Kong Polytechnic University under Projects 1-ZVBA, G-U753, G-YK75, and G-YJ75.

W. Shi is with the Joint Research Laboratory on Spatial Information, The Hong Kong Polytechnic University, Hong Kong/Wuhan University, Wuhan 430079, China.

Z. Miao is with the Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, and also with the School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China (e-mail: cumtzlmiao@gmail.com).

J. Debayle is with the Laboratoire Georges Friedel (Unité Mixte de Recherche Centre National de la Recherche Scientifique 5307), Centre Ingénierie et Santé, Ecole Nationale Supérieure des Mines, 42023 Saint-Etienne, France.

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.2013.2272593

Another set of studies focused on road centerline extraction in classified imagery. The thinning algorithm [18] is usually used to extract road centerlines from classified road imagery. However, the road centerlines extracted by this method have many spurs which reduce the centerline smoothness and correctness. To extract accurate centerlines, Zhang and Couloginer [19] proposed the use of Radon transform to extract centerlines from classified road imagery. This method shows a good performance when dealing with straight road segments. However, it is not suitable for short and curvilinear lines. Poullis and You [5] presented the use of Hough transform to extract road centerlines. This approach uses a set of Gaussian-based filters to compute the magnitude of road pixel. Hough transform is then iteratively implemented to extract road centerlines. Although this approach reports a good performance, some false positives still exist due to Hough transform limitations. A self-organized clustering method was developed by Doucette *et al.* [20] for road centerline extraction from classified imagery. This approach cannot extract centerlines located at the end of the road segments. As indicated previously, this study examines how to extract smooth and accurate road centerlines from classified images.

Several authors also proposed road extraction methods from different viewpoints. Based on homogeneous polygonal areas around each pixel, Hu *et al.* [21] defined the pixel footprint to extract road areas. Similarly, Zhang *et al.* [22] applied this detector to extract roads in urban areas. Movaghati *et al.* [23] applied particle filtering and Kalman filtering to extract road networks. Road intersection extraction from remotely sensed images was also studied [24]–[26]. Although road intersection extraction alone cannot be a complete road network generation, it is useful for understanding road network topologies and higher processing.

Most popular and successful region-based methods rely on one or two methods of processing: 1) classification and 2) road geometrical property [2]. Consequently, it is possible to improve road extraction accuracy by the following methods: 1) using classifiers with higher accuracy and 2) integrating road geometrical features. Recently, spectral–spatial classification has become one of the advanced techniques used for remotely sensed imagery [27]. Spectral–spatial classification methods use spectral and spatial information simultaneously, resulting in higher robustness and accuracy than that of traditional pixelwise classification methods which use spectral information alone. Local spatial statistics [28], [29], which measures image local autocorrelation, has been studied as a means to implement spectral–spatial classification. Mathematical morphology (MM) is also one of the commonly used methods to produce image spatial structure [30]–[32]. A classical MM depends on a fixed-shape structuring element. Consequently, it is not sufficiently flexible to simultaneously detect curved and rectilinear structures [33]. As indicated earlier, to solve this problem, this paper investigates general adaptive neighborhood (GAN) MM (GANMM) [34] to perform spectral–spatial classification. Other spectral–spatial classification methods can be referenced [37].

In other studies, the road's geometrical features are also explored to improve road extraction accuracy. In general, integrating the road's shape features can contribute to a more accurate road network extraction. It has been found that road characteristics and topology, especially road direction distribution in urban areas can be used to improve road extraction accuracy [11]. Two shape features, i.e., smoothness and compactness, were integrated with pixelwise SVM classification result to extract the road image [13]. By including geographical and topological road information, an automatic method was developed to extract roads from high-resolution satellite imagery [38]. Although integrating shape features gives a good performance in road extraction, it is difficult to obtain a universal linear feature extraction method for any situation [13], and further studies are needed.

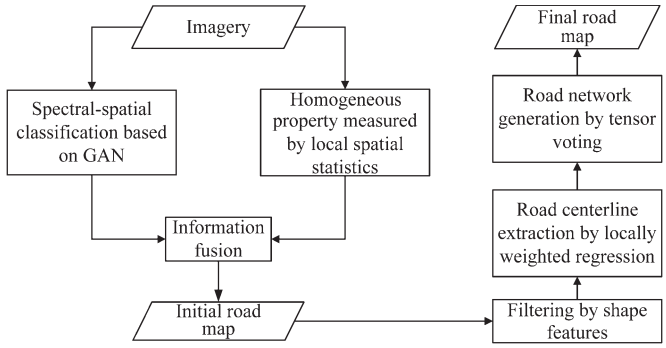


Fig. 1. Flowchart of the proposed method.

This paper presents a new approach for urban main-road centerline extraction from optical remotely sensed images integrating spectral–spatial classification, local Geary's C, road shape features, locally weighted regression, and tensor voting. This method has three main stages: 1) initial road network extraction by fusion of spectral–spatial classification based on GANMM and homogeneous property measured by local Geary's C; 2) false road removal based on road shape feature to extract reliable roads; and 3) road centerline extraction based on locally weighted regression and road network generation using tensor voting.

The remainder of this paper is organized as follows. The new approach is presented in the next section. Experimental results are given in Section III. Finally, conclusions are drawn in Section IV.

II. Main-Road Centerline Extraction Algorithm

The proposed method is presented in this section. The organization of this method is shown in Fig. 1. The details of the aforementioned steps follow in detail the following.

A. Spectral–Spatial Classification

The purpose of the spectral–spatial classification step is to segment the imagery into two groups: road groups and nonroad groups. In this paper, GANMM is investigated for spectral–spatial classification.

The GANMM has been introduced by Debayle and Pinoli [34]–[36]. The central idea of GANMM is to substitute traditional fixed-shape structural elements (SEs) by adaptive SEs. Suppose that $D \subseteq R^2$ is the spatial domain of an image and

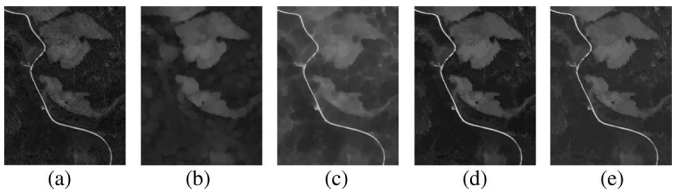


Fig. 2. Comparison of classical MM and GANMM. (a) Original image. (b) Classical opening. (c) Classical closing. (d) Adaptive opening. (e) Adaptive closing. In this example, the homogeneity tolerance value is 30.

I is the natural set of image mappings from D into $\mathbb{R}$. In this paper, the GAN sets $V_m^f(x)$ are defined as

$$V_m^f(x) = C_{f^{-1}([f(x)-m, f(x)+m])}(x) \tag{1}$$

where f in I is the image to be analyzed, $m \in \mathbb{R}^+$ is a homogeneity tolerance, and x is a pixel within the spatial support D of the image $f \in I$.

The set $V_m^f(x)$ is characterized by the following two properties.

1) $\forall y \in V_m^f(x) | f(y) - f(x)| \leq m$.
2) The set is connected.

Based on the definition of adaptive neighborhood, the adaptive SEs are defined as

$$\forall (m, f, x) \in \mathbb{R}^+ \times I \times D \; R_m^f(x) = \underset{z \in D}{\cup} \left\{ V_m^f(z) | x \in V_m^f(z) \right\}. \tag{2}$$

By applying the adaptive SEs (which are symmetric), the adaptive dilation and adaptive erosion are then defined as

$$D_m^f(f)(x) = \sup_{\omega \in R_m^f(x)} f(\omega) \tag{3}$$

$$E_m^f(f)(x) = \inf_{\omega \in R_m^f(x)} f(\omega). \tag{4}$$

Therefore, the GAN closing and the GAN opening are given as

$$C_m^f(f)(x) = E_m^f \circ D_m^f(f)(x) \tag{5}$$

$$O_m^f(f)(x) = D_m^f \circ E_m^f(f)(x). \tag{6}$$

Fig. 2 presents a comparison of classical and GAN-based MMs. As can be seen, results of classical MM are blurred, and the image spatial structure is also damaged by using classical MM. However, GANMM does not produce blurred effects; hence, the image spatial structure is well retained. The results also indicate that GANMM produces smooth results which reduce the local spectral variance caused by image noise.

In the next phase, morphological profiles (MPs) are constructed based on the aforementioned GANMM method. MPs, proposed by Pesaresi and Benediktsson [39], are representations of the size and shape information of objects in the image. The profiles are composed of multiple morphological processes using increasing sizes of an SE. In this paper, the adaptive MPs (using GANMM) at a pixel x of an image f are computed as

$$\mathrm{MP}(x) = \{C_{m_n}(x), \ldots, I(x), \ldots, O_{m_n}(x)\} \tag{7}$$

where m_n denotes the nth homogeneous tolerance.

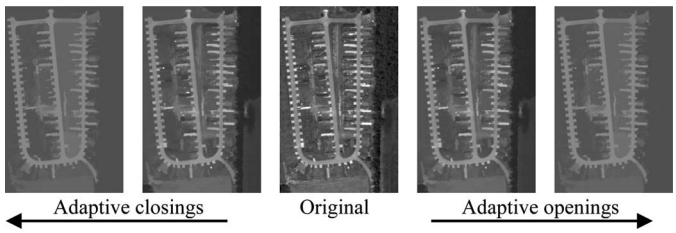


Fig. 3. Construction of MPs based on adaptive openings and closings. In this example, the homogeneity tolerances are 30 and 50.

For the multispectral imagery with B-bands, (7) will be rewritten as

$$\mathrm{MP}(x) = \left\{ \mathrm{MP}^1(x), \mathrm{MP}^2(x), \ldots, \mathrm{MP}^B(x) \right\}.$$

As can be seen from (7), MPs contain spectral information as well as the image spatial information. Fig. 3 shows an example of MPs obtained by applying a series of homogeneity tolerance.

Once the adaptive MPs are obtained, SVM [40] classification is applied to segment the imagery. The result of SVM is defined as

$$S(x) = \begin{cases} 1, & \text{if } x \text{ is classified as road} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

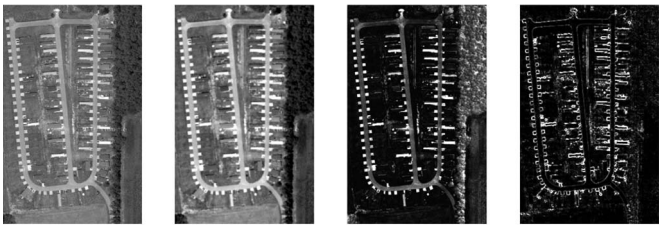
After the construction of MPs, the pixelwise classifier, such as Bayes classifier, SVM, and neural networks, is then selected to classify the imagery. In this paper, SVM is selected for two main reasons. First, previous studies show that the SVM performance is as good as or significantly better than other competing methods in most cases [41]. Second, SVM is a supervised classification method, which can still work when road materials and intensity change. The SVM classification algorithm is given in Algorithm 1.

Algorithm 1 The SVM classification algorithm

1) For each pixel, MPs are constructed as the input features.
2) For each class, 5% of ground truth data are selected as the training samples.
3) Select Gaussian kernel, and the parameters, namely, penalty parameter C and kernel parameter γ, are selected by using fivefold cross-validation.
4) Use the best parameters selected in step 3) to train the SVM classifier.
5) Classify the whole imagery using the classifier trained in step 4).

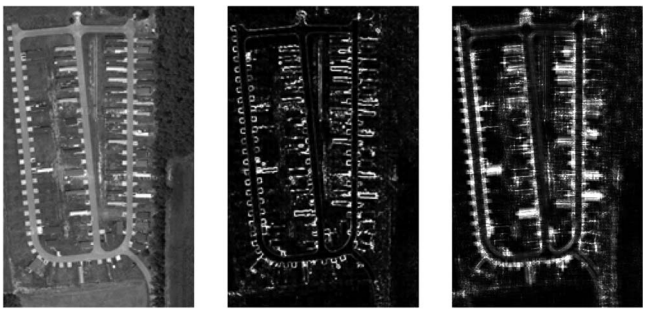
B. Local Homogeneous Property

Due to the complexity of urban images, other land-cover types, such as buildings, park lots, and bare soil, tend to be misclassified as roads. It is difficult to completely extract roads using classification alone, and hence, there is large room to improve road extraction accuracy. To overcome this problem, more road properties should be studied and integrated with classification.



(a) (b) (c) (d)

Fig. 4. (a) Original image. (b) Getis-Ord G_i map. (c) Local Moran's I map. (d) Local Geary's C map.



(a) (b) (c)

Fig. 5. (a) Original image. (b) Local Geary's C result (window size = 2 pixels). (c) Local Geary's C result (window size = 6 pixels).

Roads are continuous features. Based on this property, the inference is that roads should be located in homogeneous regions. In this paper, local Geary's C [42] is used to measure image local homogeneity properties. Local Geary's C, which is one of the local spatial statistics indicators, measures the autocorrelation between a pixel and its neighboring pixels.

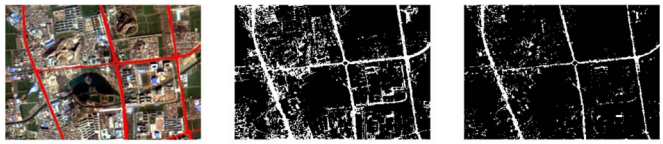
Local Geary's C is defined as

$$c_i = \frac{1}{\frac{1}{n}\sum_{j=1}^{n}(y_j-\bar{y})^2}\sum_{j=1}^{n} w_{ij}\left[(y_i-\bar{y})-(y_j-\bar{y})\right]^2 = \frac{1}{\frac{1}{n}\sum_{j=1}^{n}(y_j-\bar{y})^2}\sum_{j=1}^{n} w_{ij}(y_i-y_j)^2 \tag{9}$$

where n is the number of georeferenced observations, y_i is the value of the observation at the ith location, $\bar{y}$ is the mean of the observations, and w_{ij} is the weight of spatial relations between points i and j. For ease of computation, w_{ij} is set as one in this study.

Local Geary's C index identifies areas of high variability between a pixel value and its neighboring pixels. It is useful for detecting edge areas between clusters and other areas with dissimilar neighboring values [43]. Fig. 4 gives a comparison result of three local spatial statistics. As can be seen, local Geary's C is more suitable to measure the local homogeneity of the gray values.

In this paper, the window size is 2 pixels, as suggested by experimental results. Fig. 5 shows the results of local Geary's C on an image with different window sizes. As shown in Fig. 5(c), local Geary's C computed from large size windows damages the homogeneous property of small areas.



(a) (b) (c)

Fig. 6. (a) Original image. The reference road is shown in red. (b) Classification result. (c) Information fusion result.

After the computation of local Geary's C of every spectral band, a binary SVM classification is then applied to extract homogeneous regions. The binary classification result is defined as

$$L(x) = \begin{cases} 1, & \text{if } x \text{ is in the homogeneous region} \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

C. Information Fusion

The information fusion purpose is to improve the robustness and the accuracy of extracted roads by incorporating different road features. After Step A and Step B, the classification and local Geary's C binary result are combined using the following rule:

$$F(x) = \begin{cases} 1, & \text{if } S(x)=1 \text{ and } L(x)=1 \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

A comparison of classification and information fusion results is presented in Fig. 6. As can be seen, the information fusion not only removes misclassified roads to some extent but also disconnects most misclassified and true roads. After disconnecting misclassified and true roads, it can be seen from Fig. 6(c) that misclassified and true roads have different geometrical properties. This opens a door to filter information fusion result using shape features to further improve road extraction accuracy.

D. Filtering by Shape Feature

Misclassified roads will be removed to some extent by information fusion processing. However, false roads still exist, and further processing is necessary to improve the reliability of road extraction. In general, roads have a unique geometric property, which is quite different from other land-cover features. Roads are always elongated with small changes of curvature. Hence, road shape information can be used to filter false segments. In this paper, image moments are used to measure road shape feature. The key advantage of using image moments is that the filtering threshold can be set automatically via its distribution histogram.

For an $m \times n$ image, its $(p+q)$ order moment [44] is defined as

$$m_{pq} = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1}(x)^p \cdot (y)^q I(x,y) \tag{12}$$

where $I(x,y)$ is the gray value at the point (x,y), p and q are nonnegative integers, and $(p+q)$ is called the *order* of the moment.

The corresponding central moment μ_{pq} is then given as

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y) \tag{13}$$

where the coordinates (x_c, y_c) denote the centroid of $I(x, y)$, $x_c = m_{10}/m_{00}$, and $y_c = m_{01}/m_{00}$.

Based on (13), the normalized moment η_{pq} is described as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{14}$$

where $\gamma = (p + q + 2)/2$.

The sum of second-order moments is computed as

$$M_1 = \eta_{02} + \eta_{20}. \tag{15}$$

Previous studies [45] show that roads have large values of M_1 as they are complex and irregular, and hence, segments with values of M_1 less than a threshold T_{M_1} are removed. In this paper, the T_{M_1} threshold is automatically defined with Otsu's adaptive method [46]. Connected component analysis is used to label each disjoint segment and compute its value of M_1. The detailed filtering steps by image moments are given in Algorithm 2.

Algorithm 2 Steps of filtering roads by image moments

1. Label the connected segments.
2. Obtain second moment (M_1) value of each connected segment.
3. Select the T_{M_1} threshold by using Otsu's method.
4. For each component

if $(M_1 \leq T_{M_1})$ **then**
delete that segment
end if

E. Road Centerline Extraction

After road extraction, morphological thinning algorithms are a commonly used method to extract road centerlines. The advantage of the thinning algorithm is that it is both fast and easy to perform. However, road centerlines extracted by this method always produce many spurs which reduce the smoothness and correctness of road network. To solve this problem, the local linear kernel regression method to extract the road centerlines is used in this study.

Local linear kernel regression [47] solves a separate weighted least squares problem at each target point x_0

$$\min_{\alpha(x_0), \beta(x_0)} \sum_{i=1}^{N} K_\lambda(x_0, x_i) \left[y_i - \alpha(x_0) - \beta(x_0) x_i \right]^2 \tag{16}$$

where N is the number of points, $K_\lambda(x_0, x_i) = D(|x_0 - x_i|/\lambda)$ is the kernel function which determines local weights, λ is the smoothing parameter of the kernel function, and $\alpha(x_0)$ and $\beta(x_0)$ are the solutions to the weighted least squares problem.

After solving (16), the estimate is then expressed as

$$\hat{f}(x_0) = \hat{\alpha}(x_0) + \hat{\beta}(x_0) x_0. \tag{17}$$

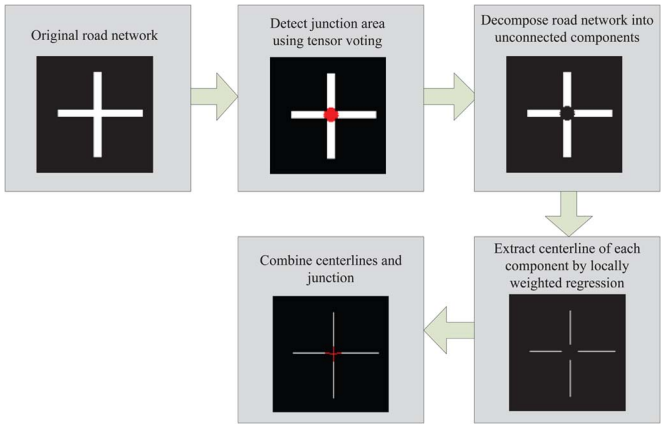


Fig. 7. Road centerline extraction using local linear kernel smoothing regression.

The matrix form of (18) is given as

$$\begin{aligned} \hat{f}(x_0) &= b(x_0)^{\mathrm{T}} \left(\mathbf{B}^{\mathrm{T}} \mathbf{W}(x_0) \mathbf{B} \right)^{-1} \mathbf{B}^{\mathrm{T}} \mathbf{W}(x_0) y \\ &= \sum_{i=1}^{N} l_i(x_0) y_i \end{aligned} \tag{18}$$

where

$$b(x)^{\mathrm{T}} = (1, x) \qquad \mathbf{B} = \begin{bmatrix} b(x_1)^{\mathrm{T}} \\ b(x_2)^{\mathrm{T}} \\ \vdots \\ b(x_N)^{\mathrm{T}} \end{bmatrix}_{N \times 2}$$

$$\mathbf{W} = \begin{bmatrix} K_\lambda(x_0, x_1) & & \\ & \ddots & \\ & & K_\lambda(x_0, x_N) \end{bmatrix}_{N \times N}.$$

In real-world applications, road networks are complex and generally have many junctions. Hence, road centerlines cannot be directly extracted by regression. To overcome this problem, a road network decomposition method is proposed in this study. Fig. 7 gives a summary of the centerline extraction from classified roads using a locally weighted regression method. First, junction areas of the road network are detected by the tensor voting method [48], [49]. This method is introduced in the next section. Second, the road network is decomposed to unconnected parts using the tensor voting method. Third, centerlines are extracted by applying the locally weighted regression to every branch. Finally, all isolated centerlines are combined with junctions.

F. Road Network Generation

Extracted roads always have discontinuities caused by many factors, such as the image noise, shadows, and the classifier's limitations. The purpose of road network generation is to eliminate road gaps to generate a complete road network. In this paper, tensor voting is introduced to link unconnected roads.

Tensor voting is a perceptual grouping and segmentation framework introduced by Medioni *et al.* [48], [49]. In 2-D, a second-order symmetric tensor T is defined as

$$\begin{aligned} T &= \begin{bmatrix} \vec{e_1} & \vec{e_2} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \vec{e_1} \\ \vec{e_2} \end{bmatrix} \\ &= \lambda_1 \vec{e_1} \vec{e_1}^{\mathrm{T}} + \lambda_2 \vec{e_2} \vec{e_2}^{\mathrm{T}} \end{aligned} \tag{19}$$

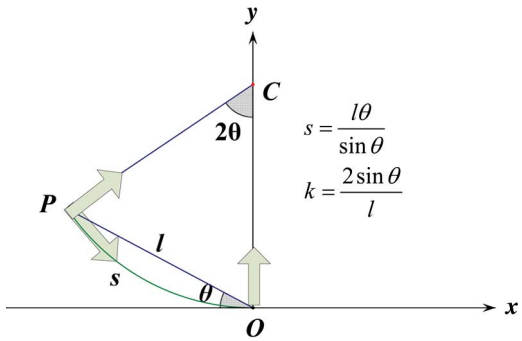


Fig. 8. Votes cast by a stick tensor located at the origin O. C is the center of the osculating circle passing through points P and O.

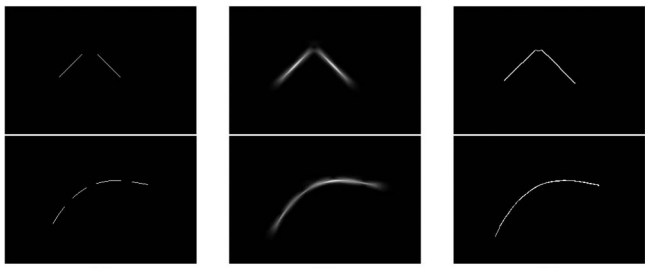


Fig. 9. Road connections using tensor voting. (a) Input image. (b) Stick saliency map. (c) Connection result.

where λ_i denotes the eigenvalues and $\vec{e_i}$ denotes corresponding eigenvectors. The tensor T can be decomposed as follows:

$$T = (\lambda_1 - \lambda_2)\vec{e_1}\vec{e_1}^{\mathrm{T}} + \lambda_2\left(\vec{e_1}\vec{e_1}^{\mathrm{T}} + \vec{e_2}\vec{e_2}^{\mathrm{T}}\right) \tag{20}$$

where $\vec{e_1}\vec{e_1}^{\mathrm{T}}$ is a stick tensor with associated saliency $(\lambda_1 - \lambda_2)$ which indicates an elementary curve and $(\vec{e_1}\vec{e_1}^{\mathrm{T}} + \vec{e_2}\vec{e_2}^{\mathrm{T}})$ is a stick tensor with associated saliency λ_2 that describes a structure which has no orientation preference to a location where multiple orientations coexist.

The saliency decay function has the following form:

$$\mathrm{DF}(s, \kappa, \sigma) = e^{-\left(\frac{s^2 + c\kappa^2}{\sigma^2}\right)} \tag{21}$$

where s is the arc length OP (Fig. 8), κ is the curvature, c controls the degree of decay with curvature, and σ is the scale factor which determines the effective neighborhood size.

After tensor voting, vote analysis is performed to analyze the point feature. The following cases have to be considered.

1) A point with $(\lambda_1 - \lambda_2) > \lambda_2$ is classified as a curve point.
2) A point with $\lambda_1 \approx \lambda_2 > 0$ is classified as a region or junction point.
3) A point with low values of λ_1 and λ_2 is classified as an outlier.

An important advantage of tensor voting is that it can recognize global context refinement automatically and the connection hypothesis does not need to be set in advance. Another advantage of tensor voting is that it has only one parameter, i.e., scale factor σ, to be set by users. These advantages of tensor voting make it an ideal choice to eliminate road discontinuities. Fig. 9 shows two examples of road connection results. It can be seen that discontinuities are successfully removed.

TABLE I
ZIYUAN-3 SATELLITE TECHNICAL DATA

	General features	
Launch date	January 11, 2012	
Launch vehicle	Long March 4B	
Nominal lifetime	5 years	
Spectral bands and resolution	Panchromatic (front-facing and rear-facing): 3.5 *m* Panchromatic (ground-facing):2.1 *m* Multi-spectral (ground-facing):6.0 *m*	
Swatch	Panchromatic (front-facing and rear-facing): 52.3 *m* Panchromatic (ground-facing): 51.1 *m* Multi-spectral (ground-facing): 51.0 *m*	
Wavelength range	Panchromatic: 450 *nm* -800 *nm*	
	Multi-spectral	Blue: 450 *nm* -520 *nm* Green: 520 *nm* -590 *nm* Red: 630 *nm* -690 *nm* Near infrared: 770 *nm* -890 *nm*

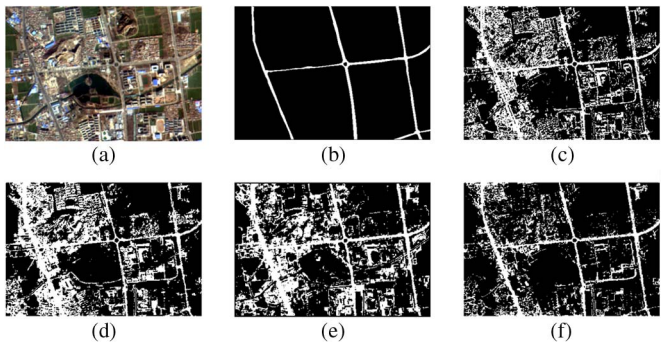


Fig. 10. (a) Original image. (b) Reference road map. (c) Road extraction result by pixelwise SVM [40]. (d) Road extraction by Tarabalka's method [50]. (e) Road extraction by Fauvel's method [31]. (f) Road extraction result by the proposed method.

III. EXPERIMENTAL RESULTS AND DISCUSSION

This paper focuses on main-road extraction from urban remotely sensed imagery. In this paper, the definition of "main road" depends on road width. The road whose width is 4–6 pixels is defined as "main road." The corresponding road reference map is generated by a hand drawing method.

In the following section, several experiments which test the proposed method are described. The proposed method is also compared with other methods in literature to show the advantages and disadvantages of the proposed method.

A. Tests on Different Spectral–Spatial Classification Methods

The initial road network of the proposed method is extracted using the spectral–spatial classification. In this experiment, different spectral–spatial classification methods were used to test the performance of road segmentation accuracy.

The experimental area is a part of the *Xuzhou City* image recorded by an infrared multispectral spectrometer carried by the Ziyuan-3 satellite, a Chinese Earth Observation satellite. The Ziyuan-3 satellite technical data are given in Table I. Fig. 10(a) gives the image of the data set. The hand-drawn reference road map is given in Fig. 10(b). Fig. 10(c) shows the road extraction result by pixelwise SVM classification [40],

TABLE II
COMPARISON OF DETECTED ROAD ACCURACY BY DIFFERENT SPECTRAL–SPATIAL CLASSIFICATION METHODS

Method	Classification accuracy (%)	Error rate (%)
Pixel-wise SVM [40]	74.21	25.79
Tarabalka's method [48]	70.87	29.13
Fauvel's method [49]	72.71	27.29
Proposed method	**84.83**	**15.17**

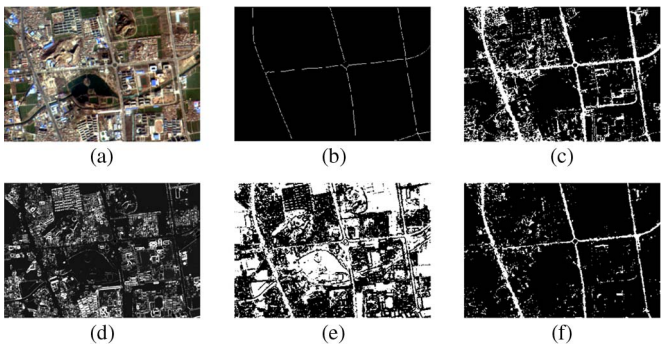


Fig. 11. (a) Study area image. (b) Hand-drawn reference road map. (c) Classification result using SVM. (d) Example of local Geary's C results. (e) Binary classification of local Geary's C map. (f) Fusion result.

and Fig. 10(d) shows the road extraction result by the algorithm proposed by Tarabalka *et al.* [50], using the expectation–maximization clustering approach. The result by the algorithm proposed by Fauvel *et al.* [31] is illustrated in Fig. 10(e), and Fig. 10(f) shows the roadmap detected by the method proposed in this study.

The classification accuracy and error rate [51] are used to evaluate different spectral–spatial classification methods. The results are presented in Table II. As can be seen, Tarabalka's method [50] and Fauvel's method [31] achieve lower classification accuracy compared to pixelwise SVM, which uses only spectral information. This is because many nonroad pixels are misclassified as roads by using these two methods. The results indicate that GAN integration, together with the spectral classification results, substantially improves road classification accuracy. Compared to classification using only spectral information, there is a 10% improvement in the proposed classification method. A vision comparison reveals that the completeness and correctness of the proposed method are superior to those of the other three methods. Hence, spectral–spatial classification based on GAN is used in the proposed method.

B. Study Area I

In this experiment, an image recorded by the Ziyuan-3 satellite is used to verify the performance of the proposed method. The image is 322 by 472 pixels. Fig. 11(a) shows the study area in image form. The manually formed ground truth data are shown in Fig. 11(b). The remote sensing image used in this study has a spatial resolution of 6 m/pixel. At this resolution, smaller roads are unclear; hence, only "main roads" with widths of about 5–6 pixels are able to be extracted.

First, the MP was constructed by applying a series of GAN openings and with increasing homogeneity tolerance. In this paper, the homogeneity tolerance is set to 10, 20, ..., 40.

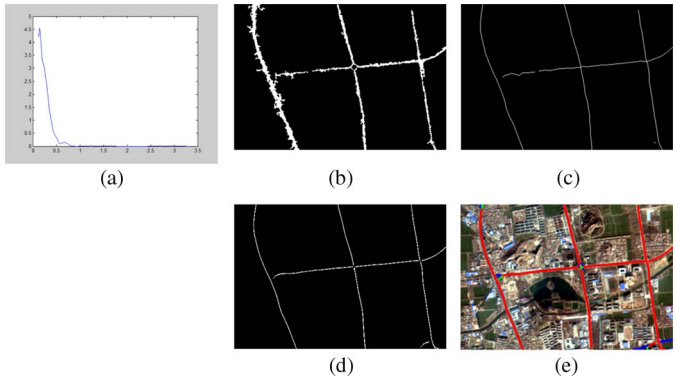


Fig. 12. (a) Second-order moment distribution of information result. (b) Filtering result. (c) Road centerline extraction result. (d) Road network generation result by tensor voting. (e) Road network and original image superposition. True positives are shown in red, false positives in green, and false negatives in blue.

SVM was then trained to classify the whole image. In this paper, Libsvm library [52] was used to train the SVM classifier. The optimal parameters were selected as $C = 128$ and $\gamma = 0.00313$. The SVM classification result is shown in Fig. 11(c). The road features are shown in white, and the nonroad features are shown in black. As can be seen from this figure, road features are extracted well by spectral–spatial classification. However, due to spectral similarity, false roads such as parks, buildings, and bare soil were misclassified as roads, and further processing was needed.

After spectral–spatial classification, local Geary's C of each band was computed. Fig. 11(d) shows an example of the local Geary's C result. The result shows that the homogeneous areas have low intensity and edge pixels have large intensity. By applying supervised binary classification, local Geary's C map was segmented into two parts: homogeneous regions and edge pixels, as shown in Fig. 11(e), where homogeneous regions are shown in white and edge pixels are shown in black.

The classification result and local Geary's C binary result were fused by applying logical "AND" operation [see Fig. 11(f)]. As can be seen, fusion of classification and local Geary's C result reduced misclassified roads to some extent. Some misclassified roads which connected with roads were also disconnected after applying information fusion, which is convenient to remove by using a shape feature in the following step. Based on these results, it is concluded that information fusion leads to improvement of the robustness of the extracted road.

After fusion processing, the second moment value of each object in road imagery was computed. Fig. 12(a) shows the second moment distribution. As can be seen in Fig. 12(a), the second moment difference between road feature and nonroad feature is obvious. In this experiment, the second moment threshold is set as 0.33. The objects whose second moment values are smaller than 0.33 were removed. The filtering result is shown in Fig. 12(b). The locally weighted regression was then performed to extract road centerline and is shown in Fig. 12(b). The road centerline map is shown in Fig. 12(c). To eliminate discontinuity, tensor voting was performed. In this paper, the scale parameter was set as 20. The road network generation result is shown in Fig. 12(d). The superposition result of the original image and extracted road centerline is shown in Fig. 12(e).

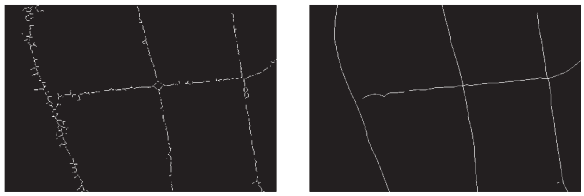


Fig. 13. (a) Thinning algorithm extraction result. (b) Proposed method extraction result.

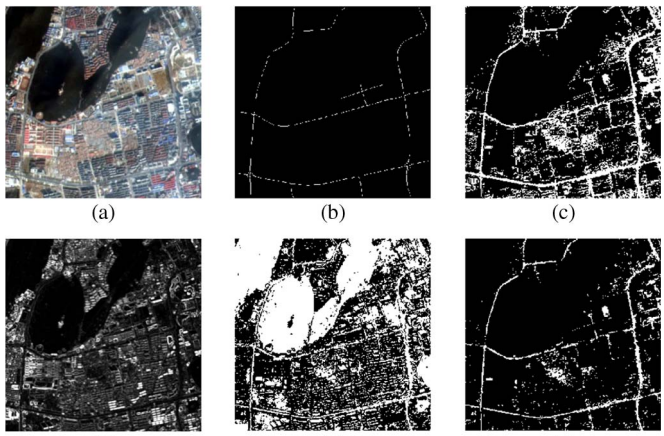


Fig. 14. (a) Image of the study area. (b) Hand-drawn reference road map. (c) Classification result from SVM. (d) Example of local Geary's C results. (e) Binary classification of local Geary's C map. (f) Fusion result.

Fig. 13 shows the comparison results produced by the thinning algorithm and the proposed method. It is clear from the results that the road centerline extracted by the thinning algorithm has many spurs, which reduces the smoothness and accuracy of the result. As seen from the results, the centerline extracted by the proposed method does not produce spur and retains the road's smoothness. However, the result also indicates that the proposed method fails to handle the complicated junctions, such as the crossing circle.

C. Study Area II

In the second experiment, the developed method was tested on another image of an urban area in Xuzhou. The image has a spatial dimension of 469 $\times$ 477. An image of the data set and the hand-drawn reference map are presented in Fig. 14(a) and (b), respectively. In this experiment, the optimal SVM parameters were selected as $C=512$ and $\gamma=0.5$. The spectral–spatial classification result is illustrated in Fig. 14(c). Local Geary's C map and its binary map are given in Fig. 14(d) and Fig. 14(e), respectively. The fusion result of spectral–spatial classification and local Geary's C result is given in Fig. 14(f).

As for the previous experiment, the second-order moment of each object was computed after performing information fusion. The second-order moment distribution is shown in Fig. 15(a). From Fig. 15(a), it is seen that the second-order moment values of road and nonroad features are easily separated. In this experiment, the second-order moment threshold is 0.45. The filtering algorithm with a threshold of 0.45 gave the result shown in Fig. 15(b). The locally weighted regression algorithm was performed to extract the road centerline, resulting in Fig. 15(c).

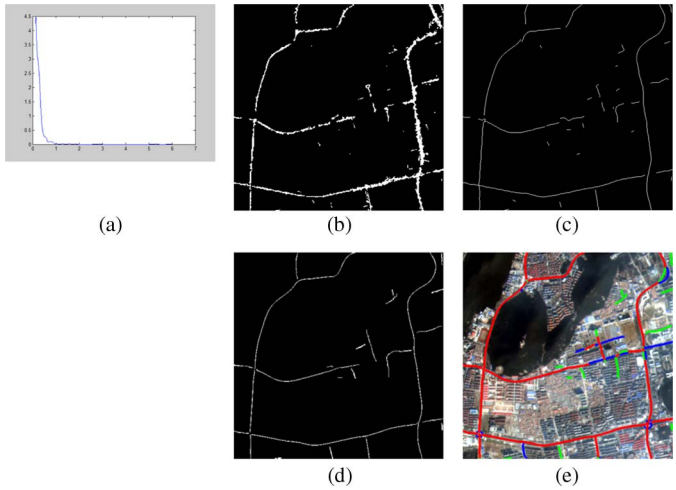


Fig. 15. (a) Second-order moment distribution. (b) Filtering result. (c) Road centerline extraction result. (d) Road network generation result by tensor voting. (e) Road network and original image superposition. True positives are shown in red, false positives in green, and false negatives in blue.

TABLE III
PERFORMANCE OF THE PROPOSED METHOD

Experiment	E_1 (%)	E_2 (%)	E_3 (%)
1	94.03	98.10	92.34
2	85.24	83.38	72.86

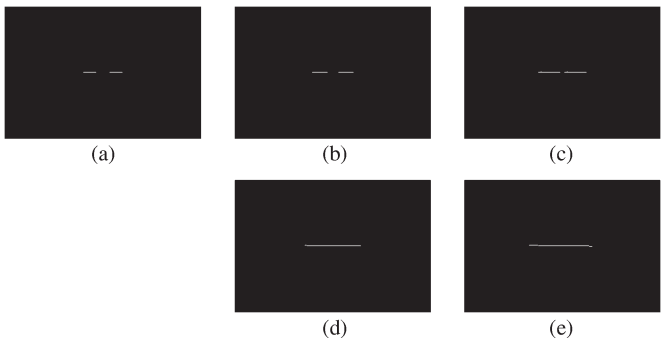


Fig. 16. Tests on scale parameter selection. (a) Original image: The gap in the image is 20 pixels. (b) Connection result with $\sigma = 1$. (c) Connection result with $\sigma = 5$. (d) Connection result with $\sigma = 10$. (e) Connection result with $\sigma = 20$.

The road network was then connected by the tensor voting method to eliminate gaps between the road segments. Fig. 15(d) shows the road network generation result. The superposition result of the original image and extracted road centerline is given in Fig. 15(e).

To evaluate the proposed method, the following three accuracy measures proposed by Wiedemmann *et al.* [53] were used in this study:

$$E_1 = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{22}$$

$$E_2 = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{23}$$

$$E_3 = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \tag{24}$$

where E_1, E_2, and E_3 denote completeness, correctness, and quality, respectively, and TP, FN, and FP represent true positive, false negative, and false positive, respectively.

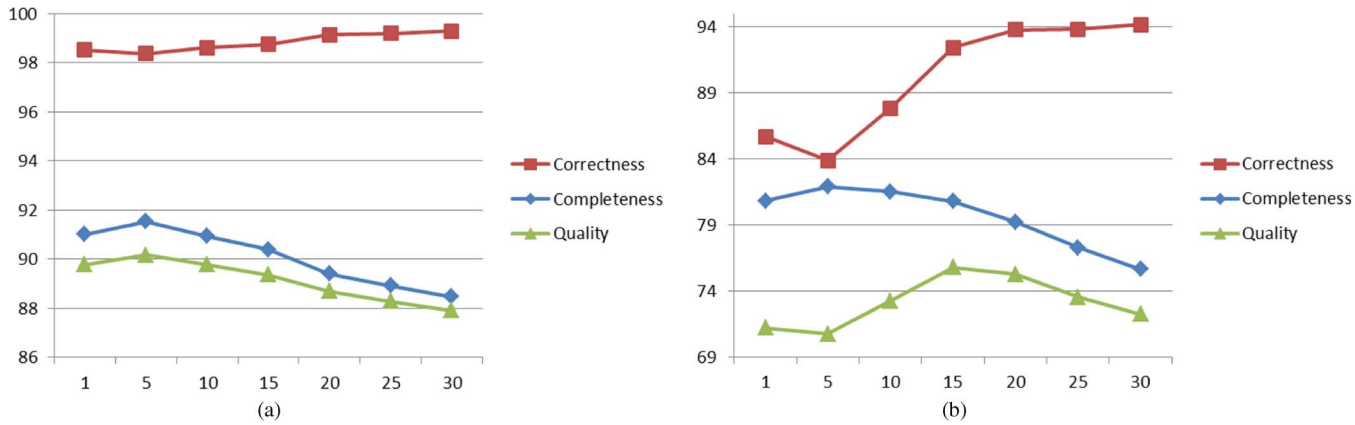


Fig. 17. Scale parameter influences on road network generation accuracy test. (a) Experiment I. (b) Experiment II.

Table III shows the proposed method performance. Mayer *et al.* [4] claim that the achievement of a completeness of at least 60% and a correctness of at least 75% is the absolute minimum for road extraction results, before they can be considered useful in practice. From Table III, it is seen that the proposed method in this study achieves these goals.

D. Tests on Scale Parameter

Tensor voting is used in the proposed method to eliminate the road discontinuity caused by image noise, such as shadows, trees, and other obstructions. Only one parameter needs to be set by the users: scale parameter σ. In this section, the sensitivity of the proposed method to variations of the scale parameter is tested, and the method and results are given in this chapter.

First, a simulation image was used to test the influence of the scale parameter setting on the connection result. Fig. 16(a) shows an image with two disconnected line segments with a 20-pixel gap. Fig. 16(b)–(e) shows the connection results using different scale parameters. As seen from the results, the gap fails to connect when σ is too small (e.g., $\sigma = 1$). When σ is progressively larger, the gap becomes increasingly smaller (e.g., $\sigma = 5$) until it ceases to exist (e.g., $\sigma = 10$). However, it should be noted that some undesired segments are produced when σ is very large (e.g., $\sigma = 20$). This is mainly because input tokens begin to heavily "cross talk" [46].

Second, two previous experimental road centerlines, extracted by locally weighted regression, were chosen to test the scale parameter setting's influence on connection result. Fig. 17 gives a quantitative evaluation result. As can be seen, the scale parameter is linearly related to correctness. This is because, as indicated earlier, the large-scale parameter will eliminate most gaps and hence produce a more correct road network. It is also seen that completeness and quality are first improved when σ becomes larger (e.g., $1 \leq \sigma \leq 10$) and then decreases (e.g., $10 \leq \sigma \leq 30$). Again, as indicated earlier, this is because many undesired road segments are produced when σ becomes larger. This phenomenon leads to a decrease of completeness and quality. Thus, in conclusion, a moderate-scale parameter σ should be carefully selected to achieve a balance between correctness, completeness, and quality.

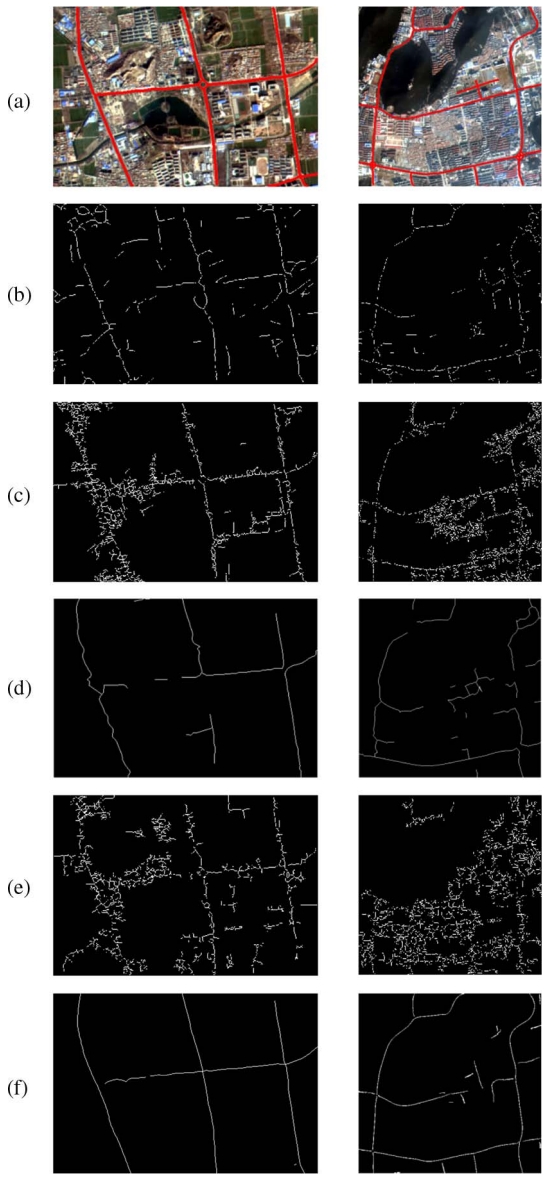


Fig. 18. Comparison of different road extraction methods on Ziyuan-3 images. (a) Input images. The hand-drawn road reference map is shown in red. (b) Results of the method proposed by Shi and Zhu [12]. (c) Results of Song and Civco's method [13]. (d) Results of the method proposed by Miao *et al.* [54]. (e) Results of Huang and Zhang's method [16]. (f) Results of the proposed method.

TABLE IV
PERFORMANCE OF DIFFERENT ROAD EXTRACTION METHODS ON ZIYUAN-3 DATA SETS

Sensors: ZIYUAN-3						
Method	Study Area I			Study Area II		
	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)
Shi	79.55	37.56	34.26	67.01	45.58	37.23
Song	97.30	32.57	32.28	94.78	24.47	24.14
Miao	76.08	86.97	68.30	63.01	67.39	48.28
Huang	88.80	27.54	26.61	81.02	16.80	16.16
Proposed	94.03	98.10	92.34	85.24	83.38	72.86

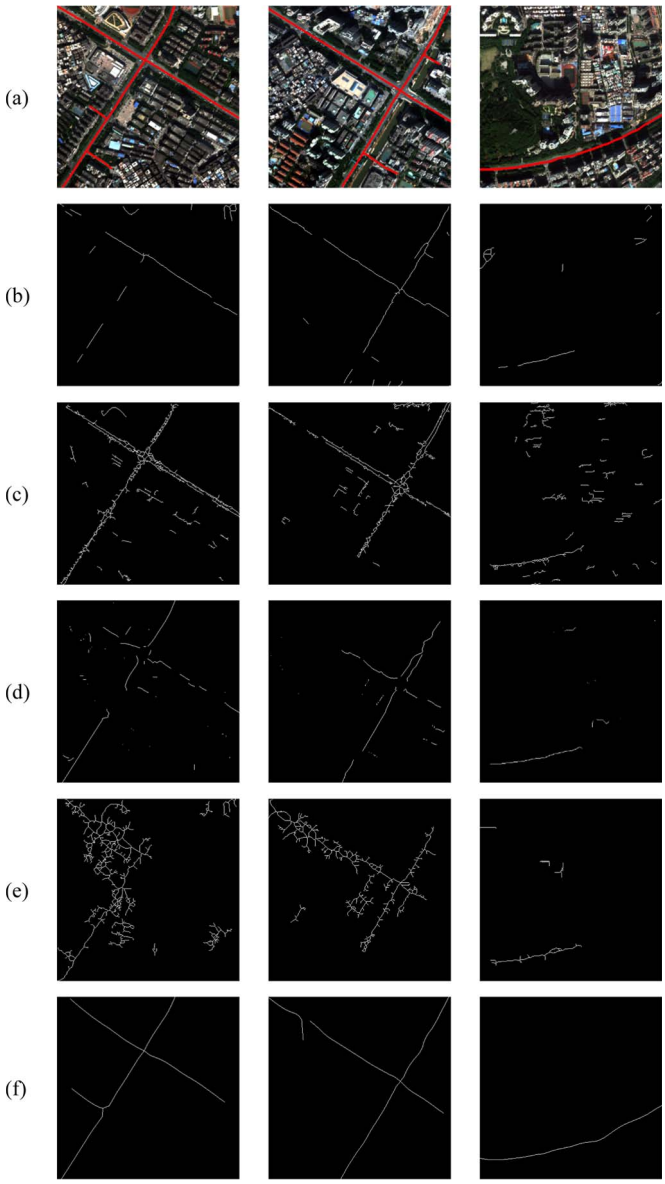


Fig. 19. Comparison of different road extraction methods on WorldView-2 images. (a) Three input images of size (512 × 512). The road reference map is shown in red. (b) Results of the method proposed by Shi and Zhu [12]. (c) Results of Song and Civco's method [13]. (d) Results of the method proposed by Miao *et al.* [54]. (e) Results of Huang and Zhang's method [16]. (f) Results of the proposed method.

E. Comparisons With the Existing Methods

In this section, the proposed method is compared with four existing road extraction methods from the literature. These four methods are introduced by Shi and Zhu [12], Song and Civco [13], Miao *et al.* [54], and Huang and Zhang [16]. All methods were programmed and performed in MATLAB [55]. Fig. 18 gives the comparison results of different road extraction methods. To evaluate these methods quantitatively, the completeness, correctness, and quality of each method are computed. The results are presented in Table IV.

As is seen in Table IV, in terms of the completeness criterion, the proposed method gives the second best result after Song and Civco's method. Of all these methods, the proposed method achieves the best performance in terms of correctness and quality indices. From Fig. 18, it is seen that the thinning algorithm used by Song and Civco's method and Huang and Zhang's method leads to low values of correctness and quality. Shi's method depends on a binary map obtained by simple

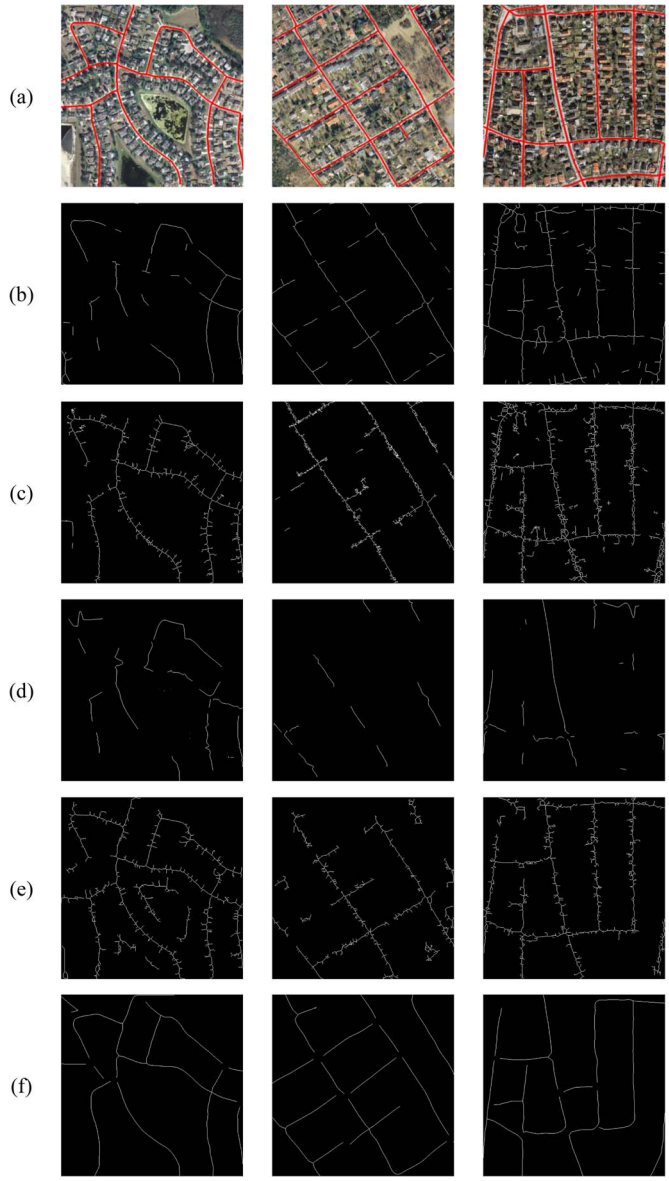


Fig. 20. Comparison of different road extraction methods on QuickBird images. (a) Input images. The road reference map is shown in red. (b) Results of the method proposed by Shi and Zhu [12]. (c) Results of Song and Civco's method [13]. (d) Results of the method proposed by Miao *et al.* [54]. (e) Results of Huang and Zhang's method [16]. (f) Results of the proposed method.

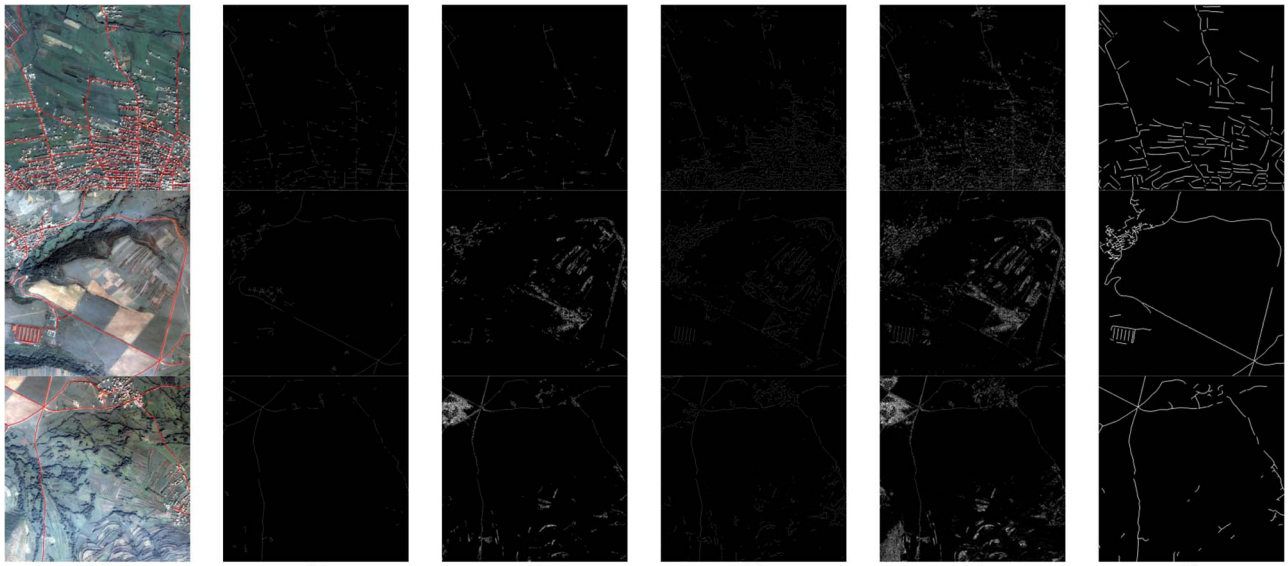


Fig. 21. Comparison of different road extraction methods on Ikonos images. (a) Input images. The road reference map is shown in red. The first, second, and third rows were using Ikonos1-Sub1, Ikonos3-Sub1, and Ikonos3-Sub2, respectively. (b) Results of the method proposed by Shi and Zhu [12]. (c) Results of Song and Civco's method [13]. (d) Results of the method proposed by Miao *et al.* [54]. (e) Results of Huang and Zhang's method [16]. (f) Results of the proposed method.

TABLE V
PERFORMANCE OF DIFFERENT ROAD EXTRACTION METHODS ON WORLDVIEW-2 DATA SETS

Sensors: WorldView-2									
Method	Study Area I			Study Area II			Study Area III		
	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)
Shi	39.47	54.02	29.54	80.05	**81.79**	**67.94**	30.26	31.22	18.15
Song	**93.34**	48.60	46.97	**81.95**	41.88	38.34	59.01	18.05	16.04
Miao	47.52	55.84	34.54	40.72	55.91	30.82	51.01	73.46	43.07
Huang	52.00	14.79	13.01	68.92	30.42	26.75	50.00	48.71	32.75
Proposed	70.00	**76.54**	**57.63**	69.51	77.22	57.68	**81.45**	**81.25**	**68.57**

TABLE VI
PERFORMANCE OF DIFFERENT ROAD EXTRACTION METHODS ON QUICKBIRD DATA SETS

Sensors: QuickBird									
Method	Study Area I			Study Area II			Study Area III		
	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)
Shi	64.84	84.11	57.77	70.65	93.79	**67.50**	84.42	63.36	56.73
Song	90.73	57.75	54.54	**72.64**	71.91	56.59	**91.30**	50.62	48.29
Miao	55.51	**84.36**	50.33	22.06	**96.91**	21.90	33.67	**87.40**	32.11
Huang	**94.94**	50.07	48.77	70.04	64.01	50.25	83.98	61.89	55.35
Proposed	75.76	75.04	**60.51**	55.45	68.31	44.10	67.70	83.03	**59.47**

thresholding. However, as urban imagery's complexity, such as image noise, material change, shadow, it is difficult to set a suitable threshold value. This leads to low accuracy in Shi's method.

The proposed method has also been tested on several very high resolution remotely sensed imageries, such as WorldView-2, QuickBird, and Ikonos. The QuickBird data set can be downloaded from Wikimapia [56]. The comparison results of the proposed method and several existing methods in the literature are presented in Figs. 19–21. Figs. 19 and 20 show the results obtained using the road extraction methods on satellite images of developed urban areas, whereas Fig. 21 shows the results for urban/suburban areas. The quantitative evaluation results are presented in Tables V–VII. Clearly, it can be claimed that the proposed method has some advantages compared to the existing methods in this field, such as Shi and Zhu's method [12], Song and Civco's method [13], and Huang and Zhang's method [16]. The proposed method was also compared with other approaches shown in Mayer *et al.* [4]. The quantitative evaluation results are presented in Table VIII. As can be seen from Table VIII, the proposed method yields the moderate accuracy, and the accuracy of the proposed method is close to the best.

TABLE VII
PERFORMANCE OF DIFFERENT ROAD EXTRACTION METHODS ON IKONOS DATA SETS

Sensors: Ikonos									
Method	Study Area I			Study Area II			Study Area III		
	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)	E_1 (%)	E_2 (%)	E_3 (%)
Shi	32.79	46.67	23.86	46.76	56.98	34.56	48.49	**74.96**	41.73
Song	17.92	**68.19**	16.54	11.62	3.87	2.99	81.23	16.68	16.06
Miao	30.22	23.91	15.41	11.73	7.11	4.63	56.99	30.47	24.77
Huang	**68.84**	23.55	21.28	21.47	3.69	3.25	90.37	8.83	8.75
Proposed	34.29	63.39	**28.63**	**76.87**	**64.68**	**54.14**	**93.65**	61.54	**59.08**

TABLE VIII
RESULTS OF THE QUANTITATIVE EVALUATION

Sensors: Ikonos						
Method	Ikonos1-Sub1		Ikonos3-Sub1		Ikonos3-Sub2	
	Completeness	Correctness	Completeness	Correctness	Completeness	Correctness
Bacher	0.34	0.66	**0.81**	0.87	0.86	0.89
Beumier	0.48	0.69	/	/	/	/
Gerke_W	0.27	0.41	0.8	0.65	0.75	0.52
Gerke_WB	0.19	0.49	0.68	0.75	0.71	0.84
Hedman	0.31	0.51	0.77	0.78	0.85	**0.91**
Malpica	0.25	**0.74**	0.6	**0.79**	0.6	0.89
Zhang	**0.56**	0.41	/	/	/	/
Proposed	0.34	0.63	0.77	0.65	**0.94**	0.62

IV. CONCLUSION

In this paper, a framework for accurate and reliable road centerline extraction from urban remotely sensed imagery has been presented. Uniquely, this framework is an integrated method which incorporates spectral–spatial classification, local Geary's C, shape features, locally weighted regression, and tensor voting.

First, the MPs obtained from GANMM opening and closing have been used to implement spectral–spatial classification to extract initial road network. The experiments indicate that GANMM-based spectral–spatial classification extracts initial road network with higher accuracy compared with other spectral–spatial classification methods. The classification and local Geary's C results are then fused in such a way that it removes some misclassified roads and disconnects some misclassified roads from true roads. The fusion approach improves the robustness and accuracy of the roads extracted.

Second, a second-order moment, as it can effectively separate road and other objects, is used to measure road shape features. The advantage of using a second-order moment to measure road shape feature is that the threshold is more easily determined from the histogram.

Third, locally weighted regression is used to extract road centerlines from classified imagery. A major advantage of using regression is that it does not produce spurs like thinning algorithm, and hence, it retains the smoothness of road centerlines.

Finally, tensor voting is used to solve certain problems of discontinuity. The advantage of using tensor voting is that it does not need to set the possible links between road segments in advance. Tensor voting can infer possible connections between two unconnected segments as it has a strong geospatial feature inference. Another benefit of using tensor voting is that it only has one parameter to be set by the user.

The proposed method has two parameters which need to be set by users: 1) homogeneity tolerance value and 2) scale parameter. The experimental results have been presented on two urban images to evaluate the proposed method, which shows that the proposed method achieves a good performance for urban main-road network extraction. However, experimental results show that the proposed method fails to handle complicated road junctions (i.e., circle junction). This is mainly because of the simplification of road centerline extraction model based on locally weighted regression. A possible method to overcome this problem would be to use a snake method to handle junctions. Another limitation of the proposed method is that it is not suitable for low-resolution imagery (less than 8-m resolution). This is because main roads just have 2–3 pixels and the local homogeneity of the gray values cannot be obtained. The third limitation of the proposed method is that it currently needs to train the SVM for each input image, which limits the applicability of the method in practice. A possible solution to this limitation is to collect massive amounts of road training samples to perform the SVM training offline [2], which does not require the interaction from users.

ACKNOWLEDGMENT

The authors would like to thank Prof. H. Mayer from the Universität der Bundeswehr München for providing the Ikonos data sets and Prof. P. Gamba from the University of Pavia for the kind discussions on road junction detection. The authors would also like to thank the Editor and three anonymous reviewers whose insightful suggestions have significantly improved this paper.

REFERENCES

[1] 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, Dec. 2003.

[2] 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.

[3] L. J. Quackenbush, "A review of techniques for extracting linear features from imagery," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 12, pp. 1383–1392, Dec. 2004.

[4] H. Mayer, S. Hinz, U. Bacher, and E. Baltsavias, "A test of automatic road extraction approaches," *Int. Archives Photogramm., Remote Sens., Spatial Inf. Sci.*, vol. 36, no. 3, pp. 209–214, 2006.

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

[6] A. Gruen and H. Li, "Semi-automatic linear feature extraction by dynamic programming and LSB-snakes," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 8, pp. 985–995, Aug. 1997.

[7] F. Dell'Acqua and P. Gamba, "Detection of urban structures in SAR images by robust fuzzy clustering algorithms: The example of street tracking," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2287–2297, Oct. 2001.

[8] C. Unsalan 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.

[9] R. Nevatia and K. Babu, "Linear feature extraction and description," *Comput. Graph. Image Process.*, vol. 13, no. 3, pp. 257–269, Jul. 1980.

[10] K. Treash and K. Amaratunga, "Automatic road detection in gray scale aerial images," *ASCE J. Comput. Civil Eng.*, vol. 14, no. 1, pp. 60–69, 2000.

[11] 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.

[12] W. Z. Shi and C. Q. 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.

[13] M. J. Song and D. Civco, "Road extraction using SVM and image segmentation," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 12, pp. 1365–1371, Dec. 2004.

[14] 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.

[15] D. Chaudhuri, N. K. Kushwaha, and A. Samal, "Semi-automated road detection from high resolution satellite image by directional morphological enhancement and segmentation techniques," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1538–1544, Oct. 2012.

[16] X. Huang and L. Zhang, "Road centerline extraction from high-resolution imagery based on multiscale structural features and support vector machines," *Int. J. Remote Sens.*, vol. 30, no. 8, pp. 1977–1987, Apr. 2009.

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

[18] P. Soille, *Morphological Image Analysis: Principles and Applications*. New York, NY, USA: Springer-Verlag, 1998.

[19] Q. Zhang and I. Couloigner, "Accurate centerline detection and line width estimation of thick lines using the radon transform," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 310–316, Feb. 2007.

[20] P. Doucette, P. Agouris, A. Stefanidis, and M. Musavi, "Self-organized clustering for road extraction in classified imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 55, no. 5/6, pp. 347–358, Mar. 2001.

[21] 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," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 4144–4157, Dec. 2007.

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

[23] 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.

[24] A. Barsi and C. Heipke, "Artificial neural networks for the detection of road junctions in aerial images," *ISPRS Arch.*, vol. 34, pt. 3/W8, pp. 113–118, Sep. 2003.

[25] M. Ravanbakhsh, C. Heipke, and K. Pakzad, "Road junction extraction from high resolution aerial imagery," *Photogramm. Rec.*, vol. 23, no. 124, pp. 405–423, Dec. 2008.

[26] J. J. Ruiz, T. J. Rubio, and M. A. Urena, "Automatic extraction of road intersections from images based on texture characterisation," *Survey Rev.*, vol. 43, no. 321, pp. 212–225, Jul. 2011.

[27] A. Plazaa, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. 1, pp. S110–S122, Sep. 2009.

[28] W. Su, J. Li, Y. H. Chen, Z. Liu, J. Zhang, T. M. Low, I. Suppiah, S. Atikah, and M. Hashim, "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, Jun. 10, 2008.

[29] 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, Apr. 2005.

[30] R. Bellens, S. Gautama, L. Martinez-Fonte, W. Philips, J. C. Chan, and F. Canters, "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.

[31] 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.

[32] 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.

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

[34] J. Debayle and J. C. Pinoli, "General adaptive neighborhood image processing Part I: Introduction and theoretical aspects," *J. Math. Imaging Vis.*, vol. 25, no. 2, pp. 245–266, Sep. 2006.

[35] J. C. Pinoli and J. Debayle, "General adaptive neighborhood mathematical morphology," in *Proc. IEEE Int. Conf. Image Process.*, Cairo, Egypt, Nov. 7–10, 2009, pp. 2249–2252.

[36] J. C. Pinoli and J. Debayle, "Spatially and intensity adaptive morphology," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 7, pp. 820–829, Nov. 2012.

[37] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral–spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[38] 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, no. 9, pp. 1201–1220, Jul. 2005.

[39] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.

[40] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[41] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, Jun. 1998.

[42] L. Anselin, "Local indicators of spatial association—LISA," *Geograph. Anal.*, vol. 27, no. 2, pp. 93–115, Apr. 1995.

[43] *ENVI Help*, Exelis Visual Information Solutions, Boulder, CO, USA, 2013.

[44] J. Flusser, "Moment invariants in image analysis," in *Proc. World Acad. Sci. Eng. Technol.*, Prague, Czech Republic, 2006, pp. 196–201.

[45] B. Lu, Y. Ku, and H. Wang, "Automatic road extraction method based on level set and shape analysis," in *Proc. 2nd Int. Conf. Intell. Comput. Technol. Autom.*, Changsha, China, 2009, pp. 511–514.

[46] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jun. 1979.

[47] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Berlin, Germany: Springer-Verlag, 2008.

[48] G. Medioni, M. S. Lee, and C. K. Tang, *A Computational Framework for Segmentation and Grouping*. New York, NY, USA: Elsevier, 2000.

[49] P. Mordohai and G. Medion, *Tensor Voting: A Perceptual Organization Approach to Computer Vision and Machine Learning*. San Rafael, CA, USA: Morgan and Claypool, 2006.

[50] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral–spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.

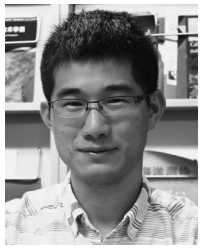
[51] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

[52] C. Chang and C. Lin, LIBSVM—A Library for Support Vector Machines 2013. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm
[53] C. Wiedemmann, C. Heipke, and H. Mayer, "Empirical evaluation of automatically extracted road axes," in *Proc. CVPR*, 1998, pp. 172–187.
[54] Z. Miao, W. Shi, H. Zhang, and X. Wang, "Road centerline extraction from high-resolution imagery based on shape features and multivariate adaptive regression splines," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 583–587, May 2013.
[55] *MATLAB*, The MathWorks, Natick, MA, USA, 2013.
[56] VPLab, Downloads 2013. [Online]. Available: http://www.cse.iitm.ac.in/~sdas/vplab/satellite.html

Wenzhong Shi received the Ph.D. degree from the University of Osnabrück, Vechta, Germany, in 1994.

He is currently a Professor in geographic information system (GIS) and remote sensing with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong. He has published some 400 research papers (including over 80 science citation index (SCI) papers) and 10 books. His current research interests include GIS and remote sensing, uncertainty and spatial data quality control, and image processing for high-resolution satellite images.

Dr. Shi was the recipient of the State Natural Science Award from the State Council of China in 2007 and The Wang Zhizhuo Award from the International Society for Photogrammetry and Remote Sensing in 2012.



Zelang Miao received the B.S. degree in survey engineering from the China University of Mining and Technology, Xuzhou, China, in 2009. He is currently working toward the Ph.D. degree in the Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong.

He is also with School of Environmental Science and Spatial Informatics, China University of Mining and Technology. His research interests include image classification and segmentation, pattern recognition, and computer vision methods with applications in remote sensing.

Johan Debayle received the M.Sc. degree in images from Université Jean Monnet, Saint-Etienne, France, in 2002, and the Ph.D. degree in Image, Vision and Signal from Ecole Nationale Supérieure des Mines, Saint-Etienne, France, in 2005.

In the beginning of 2006, he joined the French National Institute for Research in Computer Science and Control (INRIA) as a Postdoctoral Fellow in the field of biomedical image analysis. Since 2007, he has been an Associate Professor with the Mathematical Imaging and Pattern Analysis Group, Laboratoire Georges Friedel (Unité Mixte de Recherche Centre National de la Recherche Scientifique 5307), Centre Ingénierie et Santé, Ecole Nationale Supérieure des Mines, Saint-Etienne. His research interests include adaptive processing and analysis of images and patterns.